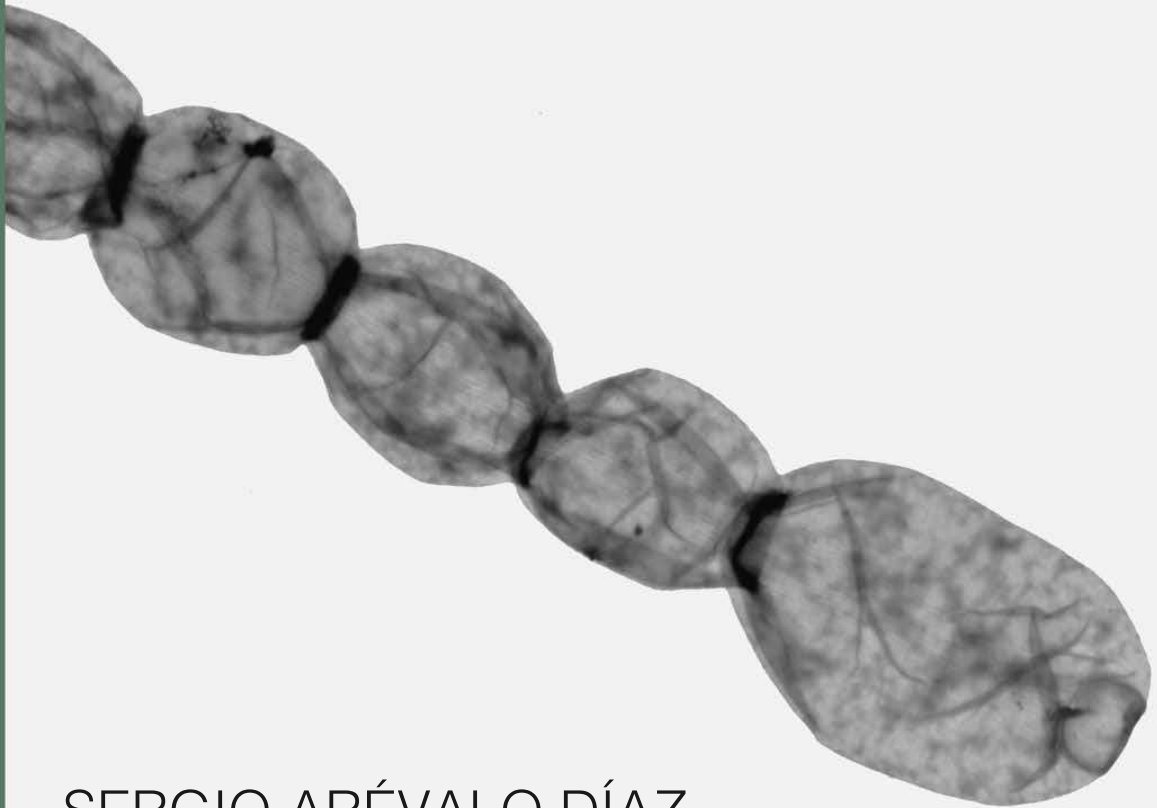


# PROTEINS INFLUENCING THE SEPTAL JUNCTIONS IN HETEROCYSTOUS CYANOBACTERIA



SERGIO ARÉVALO DÍAZ

Sevilla, 2020

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e2000022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular



Cover: Micrograph of the peptidoglycan sacculus of part of a filament of *Anabaena* grown in BG11<sub>0</sub> medium. The peptidoglycan sacculus was stained with uranyl acetate and visualized by TEM.

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**





# PROTEINS INFLUENCING THE SEPTAL JUNCTIONS IN HETEROCYSTOUS CYANOBACTERIA

Trabajo presentado para  
optar al grado de Doctor en  
Biología por el Graduado  
Sergio Arévalo Díaz

Sevilla, mayo de 2020

Director

Tutor

Dr. Enrique Flores García  
Profesor de Investigación del CSIC

Dr. José María Romero Rodríguez  
Catedrático de Bioquímica y Biología Molecular

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e200022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

## Agradecimientos

Esta tesis ha sido realizada en el Instituto de Bioquímica Vegetal y Fotosíntesis de Sevilla, Centro Mixto del Consejo Superior de Investigaciones Científicas y la Universidad de Sevilla, gracias a una beca de Formación de Personal Investigador (BES-2015-071939) asociada al proyecto de investigación BFU2014-56757-P, del que el Dr. Enrique Flores García ha sido investigador principal.

Durante la ejecución de la tesis disfruté de dos estancias, una de cuatro meses y otra de mes y medio, en la Universidad Goethe (Frankfurt, Alemania), formando parte del grupo del profesor Enrico Schleiff y bajo la supervisión del Dr. Rafael Pernil García.

Para comenzar me gustaría dar las gracias a mi director de tesis, el Dr. Enrique Flores, por darme la oportunidad de realizar la tesis en su grupo, su sabiduría a la hora de afrontar los experimentos e interpretación de sus resultados y por su infinita paciencia conmigo. De igual modo quería agradecer a la Dra. Antonia Herrero por su visión crítica y constructiva que, sumada a la de Enrique, ha hecho posible la realización de esta tesis.

También quisiera expresar mi agradecimiento tanto a los miembros actuales del grupo de investigación como a todos aquellos que han formado parte del mismo durante esta etapa, aportando cada uno de ellos su granito de arena a la elaboración del trabajo. Especialmente quisiera agradecer a Gracia su ayuda esencial en los momentos de máximo agobio y saber siempre qué decir para animarme en los malos momentos. A Mireia, que pese a tener la tarea de cargar con el novato del grupo puso todo de su parte para explicarme cómo hacer las construcciones genéticas. A Sergio Camargo por su paciencia a la hora de introducirme en el mundo de la microscopía. A Vicente Mariscal e Ignacio Luque por tener siempre un momento para resolver dudas y no perder la sonrisa pese a lo básico de la pregunta. A Mercedes y Ana por su ayuda y consejos. A José Enrique por toda su ayuda científica (y por ser un gran compañero de debate futbolístico). Y finalmente, quisiera agradecer a Félix toda su ayuda a pesar de tener mil cosas que hacer y por ser el primero en acogerme en Sevilla con los brazos abiertos, siendo mi primer amigo aquí. En general, gracias a todos mis compañeros del grupo y a la gente del IBVF por ser como son y estar siempre dispuestos tanto a ayudarte en el uso de cualquier máquina como a beberse una caña anti-estrés.

I would like to express my gratitude to Prof. Enrico Schleiff for giving me a warm welcome in his group and for all his help during my stays in Frankfurt. I also thank Enrico for his detailed analysis of the results obtained in his laboratory. To Rafa Pernil for all his scientific advises and help with administrative issues (since my first stay in Frankfurt, the term “administrative problems” took on a new meaning for me). To Dr. Roman Ladig for all his help in the analysis of my protein samples. Finally, I would like to thank the whole research group in Frankfurt, both the cyanobacteria line and the other lines, which made my stays there really enjoyable and made the idea that Germans are cold people just a myth. I would especially like to thank Nicklas who made me feel part of the group from the very first moment despite my language problems.

I would also like to thank Drs. Rachel Foster, Martin Ekman and Mercedes Nieves from Stockholm University for providing me with DNA sequences and synthetic genes from symbiotic cyanobacteria.

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e200022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular



To finish acknowledging my colleagues, I would like to thank Dr. Félix Ramos, Prof. Enrico Schleiff, Dr. Ignacio Luque, Dr. Rachel Foster and Dr. Mercedes Nieves for making time in their busy schedules to read parts of this thesis. Your comments and suggestions have been really helpful to complete this work.

Quisiera dar las gracias a toda mi familia. A mi madre por apoyarme en todas mis decisiones y ser para mí un auténtico ejemplo de valor y valentía. A mi tita Manoli, tita Mari, tito Paco y tito José por estar siempre presentes en mi vida y enseñarme lo que realmente significa la palabra familia. A mi prima Ana y mi primo Jesús por ser mucho más que mis primos, siendo como unos hermanos para mí. A mi primo Paco por ser un referente para mí, siendo el hermano mayor que te guía y te enseña que puedes conseguir más de lo que tú mismo creías. Finalmente, no quisiera olvidarme de mi Tía María por haber sido la “abuela” que toda persona necesita. Resumiendo, gracias a todos, sin vosotros no hubiera llegado hasta aquí.

No quiero pasar la oportunidad de agradecer a todos mis amigos, especialmente a Carlos, Champion, Esteban, Jaime y Santi, por hacerme sentir de vuelta en casa cada vez que voy al pueblo, haciendo como si el tiempo no hubiera pasado para nosotros, aunque pasemos meses sin vernos. No hay nada más reconfortante que esas cañas en la barra del bar, en la cual, durante unas horas, todos somos expertos psicólogos, políticos, arquitectos, informáticos... siendo capaces de arreglar todos los problemas del mundo. A mi amiga Mari, que si he llegado hasta aquí en parte ha sido por ella, por nunca darse por vencida conmigo y explicarme mil veces las cosas para que pudiera aprobar los exámenes. A mi amiga Cristina, por ser un soporte para mí en los días malos de tesis (¿cuántos negocios alternativos a la ciencia habremos pensado, por Dios?) y por ser mi alojamiento en mis escapaditas a Granada. Y finalmente me gustaría agradecer a Rubén por transmitirme sus conocimientos tanto en ciencia como en beber cerveza (aunque lo que él beba no se considere cerveza), gran jefe y mejor amigo.

Infine voglio ringraziare Bea, per avermi dato la pazienza e la sanità mentale di cui ho bisogno in quei brutti giorni in cui si vede tutto nero e soprattutto per avermi sopportato ogni giorno, anche dopo un'intera quarantena (questa donna merita un monumento). Senza i tuoi spaghetti alla carbonara e i tuoi gnocchi la stesura di questa tesi sarebbe stata molto più difficile. Ti amo Italianini.

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e200022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular



## INDEX

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**





<b>INDEX</b>	<b>iii</b>
<b>FIGURES INDEX</b>	<b>vii</b>
<b>TABLE INDEX</b>	<b>x</b>
<b>RESUMEN EN ESPAÑOL</b>	<b>xiii</b>
<b>CONCLUSIONES EN ESPAÑOL</b>	<b>xvi</b>
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Multicellularity in bacteria	3
1.2 Cyanobacteria	3
1.2.1 Origin, diversity and classification of cyanobacteria	3
1.2.2 Physiology of cyanobacteria	7
1.2.3 Cell envelope	9
1.2.3.1 Outer membrane	9
1.2.3.2 The peptidoglycan layer	10
1.3 Heterocysts	11
1.3.1 Characteristics of the heterocyst	11
1.3.1.1 Structural characteristics	11
1.3.1.2 Metabolic characteristics	13
1.3.2 Heterocyst differentiation and pattern formation	15
1.3.3 Genes involved in the formation of heterocyst envelopes	16
1.4 Septal junctions and septal proteins	18
1.4.1 SepJ	20
1.4.2 Fra proteins	22
1.5 Intercellular communication	24
1.5.1 Continuous periplasm	24
1.5.2 Cell-cell communication by septal junctions	25
1.5.3 Possible different types of septal junctions	25
1.5.4 Possible regulation of the septal junctions	26
1.6 Pentapeptide repeat proteins	26
1.6.1 The HglK protein	28
1.7 Objectives	30
<b>2. MATERIALS &amp; METHODS</b>	<b>31</b>
2.1 Organisms and Growth conditions	33
2.1.1 <i>Escherichia coli</i>	33
2.1.1.1 Strains	33
2.1.1.2 Culture media and conditions	33
2.1.1.3 Harvest methods	33
2.1.2 Cyanobacteria	34
2.1.2.1 Strain	34
2.1.2.2 Culture media and conditions	35



<b>2.2 Molecular techniques based on DNA manipulation</b>	<b>36</b>
2.2.1 Plasmids	36
2.2.2 DNA isolation	39
2.2.2.1 Isolation of plasmid DNA from <i>E. coli</i>	39
2.2.2.2 Isolation of DNA from PCR reactions or agarose gel	39
2.2.2.3 Isolation of total DNA from <i>Anabaena</i>	39
2.2.3 DNA quantification	39
2.2.4 Polymerase Chain Reaction (PCR)	39
2.2.5 DNA electrophoresis	41
2.2.6 DNA sequencing	41
2.2.7 Enzymatic treatments of DNA	41
2.2.7.1 Restriction	41
2.2.7.2 Ligation	41
2.2.7.3 Dephosphorylation	42
2.2.8 RNA radioactive probes	42
<b>2.3 Genetic methods</b>	<b>42</b>
2.3.1 Transformation of <i>E. coli</i> by heat shock	42
2.3.2 Conjugation to <i>Anabaena</i>	42
<b>2.4 Molecular techniques based on RNA manipulation</b>	<b>45</b>
2.4.1 RNA isolation from cyanobacteria	45
2.4.2 RNA quantification	45
2.4.3 RNA electrophoresis in agarose gels	45
2.4.4 RNA transfer to nylon filters	46
2.4.5 Hybridization with radioactive probes	46
2.4.6 Detection of radioactivity	46
<b>2.5 Protein purification from <i>Anabaena</i></b>	<b>46</b>
2.5.1 Induction and cellular lysis	46
2.5.2 Complex purification using the Strep-tag II	47
2.5.3 Complex purification using the GFP	47
2.5.4 Protein biochemical methods	48
2.5.4.1 SDS-PAGE	48
2.5.4.2 Transference of proteins to PVDF membranes	48
2.5.4.3 Staining PVDF membranes	49
2.5.4.4 Western blotting for Strep-tag II detection	49
2.5.4.5 Western blotting for GFP detection	49
2.5.5 Mass spectrometry analysis	49
<b>2.6 Peptidoglycan isolation</b>	<b>50</b>
<b>2.7 Microscopy</b>	<b>51</b>
2.7.1 Optical microscopy	51
2.7.2 Immunofluorescence localization of proteins	51
2.7.3 Confocal microscopy and FRAP experiments	52
2.7.4 Transmission electron microscopy	52
<b>2.8 Other parameters measured in <i>Anabaena</i></b>	<b>52</b>
2.8.1 Determination of chlorophyll <i>a</i>	52
2.8.2 Growth tests	53
2.8.2.1 Determination of the growth rate constant	53
2.8.2.2 Growth test on solid medium	53
2.8.3 Nitrogenase activity	53
2.8.4 Filament length	54



<b>2.9 <i>In silico</i> analysis</b>	<b>54</b>
2.9.1 Analysis of DNA sequences	54
2.9.2 Analysis of protein sequence and structure	54
<b>3. RESULTS &amp; DISCUSSION</b>	<b>55</b>
<b>Chapter 1</b>	<b>55</b>
<b>Intercellular communication in <i>Anabaena</i></b>	<b>55</b>
<b>3.1 Intercellular communication in <i>Anabaena</i></b>	<b>57</b>
3.1.1 Nanopores	57
3.1.2 Intercellular calcein transfer	61
3.1.3 Nanopores and intercellular calcein transfer	69
3.1.4 Discussion	70
<b>Chapter 2</b>	<b>73</b>
<b>The HglK protein</b>	<b>73</b>
<b>3.2 The HglK protein</b>	<b>75</b>
3.2.1 Subcellular localization of HglK	76
3.2.2 Inactivation of <i>hglK</i> in different genetic backgrounds	83
3.2.3 Intercellular junctions and communication	89
3.2.4 Subcellular localization of SepJ and FraD	91
3.2.5 Discussion	93
<b>Chapter 3</b>	<b>97</b>
<b>HglK and SepJ proteomics</b>	<b>97</b>
<b>3.3 HglK and SepJ proteomics</b>	<b>99</b>
3.3.1 Isolation the HglK and SepJ proteins	99
3.3.1.1 Isolation using the Strep-tag II	99
3.3.1.2 Isolation using SepJ-GFP	103
3.3.2 Mass spectrometry results	104
3.3.3 Discussion	109
<b>Chapter 4</b>	<b>115</b>
<b>The SepJ and Fra proteins from marine symbiotic <i>Richelia</i> and <i>Calothrix</i></b>	<b>115</b>
<b>3.4 The SepJ and Fra proteins from marine symbiotic <i>Richelia</i> and <i>Calothrix</i></b>	<b>117</b>
3.4.1 SepJ and Fra proteins encoded in the RintHH01 and CalSC01 genomes.	117
3.4.2 Expression in <i>Anabaena</i> of symbiont <i>sepJ</i> genes	120
3.4.3 Expression in <i>Anabaena</i> of symbiont <i>fra</i> genes	125
3.4.4 Discussion	132
<b>Chapter 5</b>	<b>137</b>
<b>Subcellular localization of FraE</b>	<b>137</b>
<b>3.5 Subcellular localization of FraE</b>	<b>139</b>
3.5.1 Construction and visualization of FraE-sfGFP	139
3.5.2 Discussion	143



---

<b>4. SUMMARY &amp; CONCLUSIONS</b>	<b>145</b>
<b>4.1 Summary</b>	<b>147</b>
<b>4.2 Conclusions</b>	<b>150</b>
<b>5. REFERENCES</b>	<b>151</b>



## Figure Index

### Introduction

<b>Fig. 1.</b> Example of different taxonomic groups of cyanobacteria defined by Rippka <i>et al.</i> (1979)	5
<b>Fig. 2.</b> Cyanobacterial phylogeny	6
<b>Fig. 3.</b> Heterocyst morphology	12
<b>Fig. 4.</b> Diagram of the main carbon and nitrogen flow between vegetative cells and heterocysts	14
<b>Fig. 5.</b> General structure of the heterocyst envelope in <i>Anabaena</i>	17
<b>Fig. 6.</b> Septal junctions in <i>Anabaena</i>	19
<b>Fig. 7.</b> SepJ: topology and localization	21
<b>Fig. 8.</b> Topology and localization of FraC and FraD	23
<b>Fig. 9.</b> Diagram of HetL	27
<b>Fig. 10.</b> HglK sequence	28
<b>Fig. 11.</b> Phenotype of the <i>hglK</i> mutant	29

### Materials & Methods

<b>Fig. 12.</b> Schematic representation of integration through single cross-over	43
<b>Fig. 13.</b> Schematic representation of gene replacement through double cross-over	44

### Results & Discussion

#### Chapter 1: Intercellular communication in *Anabaena*

<b>Fig. 14.</b> Nanopores in septal peptidoglycan disks of <i>Anabaena</i>	58
<b>Fig. 15.</b> Frequency distribution of nanopore numbers and diameters in septal PG disks	59
<b>Fig. 16.</b> Nanopores of the <i>Anabaena sepI</i> mutant	61
<b>Fig. 17.</b> Intercellular transfer of calcein between vegetative cells of <i>Anabaena</i>	63
<b>Fig. 18.</b> Frequency distribution of calcein transfer values in different <i>Anabaena</i> septal protein mutants	65
<b>Fig. 19.</b> Box-plot representation and statistical analysis of intercellular transfer of calcein between vegetative cells	66
<b>Fig. 20.</b> Analysis of intercellular calcein transfer between vegetative cells in different <i>sepI</i> mutants	67
<b>Fig. 21.</b> Decreased fluorescence recovery rate and presence of non-communicating cells in the $\Delta sepI$ mutant	68
<b>Fig. 22.</b> Correlation between nanopore numbers and intercellular molecular exchange	69

#### Chapter 2: The HglK protein

<b>Fig. 23.</b> Topology of the HglK (All0813) protein	76
<b>Fig. 24.</b> Construction of an <i>Anabaena</i> strain that produces a fusion the superfolder GFP to the C-terminus of HglK	78
<b>Fig. 25.</b> Subcellular localization of HglK-sfGFP	79
<b>Fig. 26.</b> Growth test of <i>Anabaena</i> wild type (PCC 7120) and strains CSSA4 ( <i>hglK-sf-gfp</i> ), CSSA7 ( <i>hglK-6xHis</i> ) and CSSA18 ( <i>hglK-Strep-tag II</i> )	79



<b>Fig. 27.</b> Construction of <i>Anabaena</i> strains that produce an HglK protein with a 6 x His tag or a Strep II tag II at the C-terminus	80
<b>Fig. 28.</b> Subcellular localization of 6xHis- and Strep II-tagged HglK	82
<b>Fig. 29.</b> Construction of <i>Anabaena</i> strains with a $\Delta hglK::C.K1$ mutation	84
<b>Fig. 30.</b> Growth phenotype of <i>Anabaena</i> and <i>hglK</i> mutant strains	85
<b>Fig. 31.</b> Nitrogenase activity of <i>Anabaena</i> and the <i>hglK</i> mutant	86
<b>Fig. 32.</b> Expression of the <i>cox2</i> and <i>nifHDF</i> gene clusters in <i>Anabaena</i> wild type and the <i>hglK</i> mutant.	87
<b>Fig. 33.</b> Filament length in <i>hglK</i> mutant strains with different genetic backgrounds	88
<b>Fig. 34.</b> Nanopores in septal peptidoglycan disks of wild-type <i>Anabaena</i> and the <i>hglK</i> mutant	89
<b>Fig. 35.</b> Effect of the $\Delta hglK::C.K1$ mutation on the intercellular transfer of fluorescent markers in different genetic backgrounds	91
<b>Fig. 36.</b> Subcellular localization of SepJ and FraD in the <i>hglK</i> mutant.	92
<b>Fig. 37.</b> Three-dimensional structure of the pentapeptide-repeat region of HglK . (All0813).	93
<b>Fig. 38.</b> Overview of the PatL (All3305) protein	94
 <b>Chapter 3: HglK and SepJ proteomics</b>	
<b>Fig. 39.</b> Construction of <i>Anabaena</i> strain CSSA17 that produces SepJ-Strep-tag II	100
<b>Fig. 40.</b> Staining with BD71 and <i>western blot</i> analysis from Strep-tag II isolations	102
<b>Fig. 41.</b> Staining with BD71 and <i>western blot</i> analysis of materials from the SepJ-GFP isolation	104
<b>Fig. 42.</b> Analysis of the proteomic results	106
<b>Fig. 43.</b> Properties of putative SepJ-interacting protein Alr2947	111
<b>Fig. 44.</b> Properties of putative SepJ-interacting protein Alr1690	112
<b>Fig. 45.</b> Properties of putative SepJ-interacting protein Alr3277	113
 <b>Chapter 4: The SepJ and Fra proteins from marine symbiotic <i>Richelia</i> and <i>Calothrix</i></b>	
<b>Fig. 46.</b> Micrographs of <i>Hemialus hauckii</i> containing <i>Richelia intracellularis</i> (A) and <i>Chaetoceros compressus</i> with <i>Calothrix rhizosoleniae</i> attached (B)	117
<b>Fig. 47.</b> Comparison of the SepJ amino acid sequences from <i>R. intracellulares</i> (strains HH01 and HM01), <i>C. rhizosoleniae</i> (strain SC01) and <i>Anabaena</i>	119
<b>Fig. 48.</b> Presence of a <i>fraCDE</i> gene cluster in <i>C. rhizosoleniae</i> SC01 (A) and of <i>fraCD</i> gene cluster in <i>R. intracellulares</i> HH01 (B)	120
<b>Fig. 49.</b> Comparison of the FraC, FraD and FraE sequences from symbionts CalSC01, RintHH01, RintRC01 and <i>Anabaena</i>	121
<b>Fig. 50.</b> Constructs for expression of symbiont <i>sepJ</i> genes in <i>Anabaena</i> .	122
<b>Fig. 51.</b> Immunolocalization of SepJ	124
<b>Fig. 52.</b> Growth test on solid medium	124
<b>Fig. 53.</b> Filament length in <i>Anabaena</i> strains bearing <i>sepJ</i> from the symbionts	125
<b>Fig. 54.</b> Constructs for expression of the <i>fraCD</i> genes from RintHH01 in <i>Anabaena</i>	126
<b>Fig. 55.</b> Immunolocalization of FraD	127
<b>Fig. 56.</b> Growth tests on solid medium with different nitrogen sources	127



<b>Fig. 57.</b> Constructs for expression of <i>fraCD</i> genes from different cyanobacteria using the <i>glnA</i> promoter in a replicative plasmid in <i>Anabaena</i>	129
<b>Fig. 58.</b> Immunolocalization of FraD in strains carrying the <i>fra</i> gene constructs in a replicative plasmid	130
<b>Fig. 59.</b> Growth test in solid medium with strains carrying the <i>fra</i> gene constructs in a replicative plasmid	131
<b>Fig. 60.</b> Phylogenetic tree of SepJ	133
<b>Fig. 61.</b> Phylogenetic tree of FraC	134
<b>Fig. 62.</b> Phylogenetic tree of FraD	135
<b>Fig. 63.</b> Phylogenetic tree of FraE	136
 <b>Chapter 5: Subcellular localization of FraE</b>	
<b>Fig. 64.</b> Predicted topology of FraE	139
<b>Fig. 65.</b> Construction and corroboration of <i>Anabaena fraE-sf-gfp</i>	140
<b>Fig. 66.</b> Visualization of FraE-sfGFP	141
<b>Fig. 67.</b> Subcellular distribution of FraE-sfGFP	142
<b>Fig. 68.</b> Top predicted structures of FraE by the Phyre2 program	143



## Table Index

### Materials & Methods

<b>Table 1.</b> <i>E. coli</i> strains used in this work together with their genotypes and references	33
<b>Table 2.</b> Anabaena strains previously available	34
<b>Table 3.</b> Cyanobacterial strains constructed in this work	34
<b>Table 4.</b> General and synthetic plasmids	37
<b>Table 5.</b> Plasmids constructed in this work	38
<b>Table 6.</b> Primers used in this work	40

### Results & Discussion

#### Chapter 1: Intercellular communication in *Anabaena*

<b>Table 7.</b> Nanopores in septal peptidoglycan disks of selected <i>sepJ</i> mutants	60
<b>Table 8.</b> Analysis of non-communicating cells in each replicate	64

#### Chapter 2: The HglK protein

<b>Table 9.</b> HglK, PatL, HetL and the Fra proteins in cyanobacteria	75
<b>Table 10.</b> Expression of the <i>hglK</i> gene in <i>Anabaena</i>	78
<b>Table 11.</b> Intercellular transfer of fluorescent markers in <i>Anabaena</i> with different genetic backgrounds	90

#### Chapter 3: HglK and SepJ proteomics

<b>Table 12.</b> Interaction partners identified with SepJ-GFP as bait -Experiment 1	107
<b>Table 13.</b> Interaction partners identified with SepJ-GFP as bait -Experiment 2	107
<b>Table 14.</b> Interaction partners identified with SepJ-Strep-tag II as bait	108
<b>Table 15.</b> Interaction partners identified with HglK-Strep-tag II as bait - Experiment 1	108
<b>Table 16.</b> Interaction partners identified with HglK-Strep-tag II as bait - Experiment 2	109

#### Chapter 4: The SepJ and Fra proteins from marine symbiotic *Richelia* and *Calothrix*

<b>Table 17.</b> Nitrogenase activity in strain CSSA15 and control strains	128
<b>Table 18.</b> Nitrogenase activity in strains carrying the <i>fra</i> gene constructs in a replicative plasmid	131

#### Chapter 5: Subcellular localization of FraE

<b>Table 19.</b> Relative fluorescence from FraE-sfGFP at different subcellular locations	142
---	-----





## Resumen y Conclusiones

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



## Resumen

La multicelularidad es una forma de organización biológica que ha aparecido de forma independiente en muchos grupos filogenéticos a lo largo de la evolución. En los organismos multicelulares se dan universalmente procesos de adhesión, comunicación y diferenciación celular. La multicelularidad se observa en varios grupos bacterianos, estando las cianobacterias entre los organismos multicelulares más antiguos de la Tierra. Las cianobacterias se caracterizan por llevar a cabo la fotosíntesis oxigénica, mediante la cual se produce la liberación de oxígeno y la producción de “poder asimilatorio” (ATP y ferredoxina reducida/NADPH), el cual se usa en la fijación de CO<sub>2</sub> atmosférico y en la asimilación de algunos otros nutrientes oxidados. Dentro de este filo se pueden encontrar tanto organismos unicelulares como multicelulares. Aquellas cianobacterias que se comportan como organismos multicelulares crecen formando filamentos, y en condiciones de deficiencia de nitrógeno algunas de ellas pueden dar lugar a un proceso de diferenciación celular produciendo un tipo celular llamado heterocisto. Este último lleva a cabo la fijación de nitrógeno atmosférico (N<sub>2</sub>), mientras que el resto de las células del filamento, denominadas células vegetativas, continúan con la fijación de CO<sub>2</sub>. Otras especies del filo tienen la capacidad de formar aquinetos, que son una forma celular de resistencia que aparece en algunas cianobacterias cuando las condiciones ambientales se vuelven adversas, y aún otras forman hormogonios, pequeños filamentos móviles con una función de dispersión.

Este trabajo se ha realizado con la estirpe modelo *Anabaena* sp. PCC 7120, que es una cianobacteria filamentosas formadora de heterocistos. Las cianobacterias poseen una envuelta celular de tipo Gram negativa, y en las cianobacterias filamentosas la membrana plasmática envuelve individualmente a cada célula mientras que la membrana externa, situada por fuera de la capa de peptidoglicano, es continua definiendo por tanto un periplasma continuo. Cuando *Anabaena* crece en ausencia de nitrógeno combinado, se produce la diferenciación de heterocistos, formándose un patrón a lo largo del filamento de un heterocisto cada 10-15 células vegetativas. El crecimiento diazotrófico exige una comunicación intercelular que permita el intercambio de nutrientes y moléculas reguladoras de la diferenciación a lo largo del filamento, llevándose a cabo esta comunicación a través de los nexos septales. Estos son unas estructuras proteicas que comunican de forma directa las células adyacentes. Actualmente se conocen algunas proteínas que forman parte de los nexos septales como son FraC y FraD, y algunas otras que son esenciales para el correcto funcionamiento o ensamblaje de los mismos, como son SepJ, SepI, HgIK (demostrada en esta Tesis) o FraE. Para el ensamblaje de los nexos septales es necesario una perforación del peptidoglicano existente entre las células adyacentes del filamento. Estas perforaciones reciben el nombre de nanoporos y se encuentran en el disco septal de peptidoglicano.

En la primera parte de este trabajo se intentó establecer una relación entre los nanoporos y la comunicación intercelular mediada por los nexos septales, utilizando como herramientas un marcador fluorescente y diferentes mutantes de *Anabaena*. El marcador fluorescente utilizado fue la calceína, que posee la propiedad de difundir desde el exterior hasta el interior de la célula, donde sufre un proceso de hidrólisis que le impide volver a atravesar la membrana plasmática, quedando atrapada en el interior de la célula y emitiendo fluorescencia. Por medio de experimentos de FRAP (“Fluorescence Recovery After Photobleaching”) se pudo calcular un indicador cuantitativo de transferencia entre células adyacentes del filamento. Por otra parte, mediante el aislamiento del peptidoglicano y su visualización por microscopía electrónica de



transmisión se pudo cuantificar el número de nanoporos de las estirpes estudiadas. Estos ensayos se realizaron en filamentos (cultivados en presencia o ausencia de nitrógeno combinado) de la estirpe silvestre y de mutantes tanto de proteínas que intervienen en la formación de los nexos septales como de proteínas que influyen en su ensamblaje y funcionamiento. Los resultados obtenidos permitieron establecer una correlación entre el número de nanoporos y la transferencia intercelular de calceína en filamentos que habían sido incubados en medio sin nitrógeno combinado, pero no en aquellos incubados en presencia de nitrógeno, en los que se observaba una elevada proporción de células no comunicantes.

En la segunda parte de esta Tesis se procedió a la caracterización de la proteína HglK, que pertenece a la familia de las *pentapeptide repeat proteins* (PRP). HglK se compone de cuatro segmentos transmembrana en su mitad N-terminal y de una sección soluble de localización periplásmica que contiene el dominio PRP en su mitad C-terminal. Estudios previos de HglK mostraron que mutantes que carecían de ella eran incapaces de crecer diazotróficamente y tenían un espacio incrementado entre las células del filamento. En esta Tesis se abordó la localización de HglK en estirpes de *Anabaena* que producían HglK fusionada a la *superfolder-GFP* (sfGFP) o con una cola de histidinas o con la etiqueta molecular Strep II. Mediante microscopía confocal y de fluorescencia y ensayos de inmunolocalización, se pudo establecer que HglK poseía una localización preferentemente polar en las células de *Anabaena*. El estudio de mutantes *hglK* determinó que HglK es necesaria para la integridad del filamento y para observar unos niveles normales de intercambio molecular entre las células vegetativas del filamento. Sin embargo, los experimentos realizados con mutantes de HglK y de otras proteínas septales mostraron que ni la fragmentación de los filamentos ni la alteración en el intercambio molecular era mayor cuando se combinaba la inactivación de HglK con la de otras proteínas septales. En la caracterización de HglK también se pudo establecer que el fenotipo de crecimiento diazotrófico negativo del mutante *hglK* estaba determinado por su incapacidad de fijar nitrógeno atmosférico como consecuencia de un defecto en la expresión génica durante la diferenciación de los heterocistos, puesto de manifiesto como una expresión disminuida de los genes del operón *cox2* (citocromo oxidasa del heterocisto) y *nifHDK* (nitrogenasa) en los mutantes *hglK*.

Los avances obtenidos mediante la caracterización de HglK y el conocimiento disponible sobre la proteína SepJ, ponían de manifiesto la implicación de ambas proteínas en los nexos septales. SepJ es una proteína integral de membrana con una localización focalizada en el centro del septo intercelular. SepJ está compuesta por una permeasa en su parte C-terminal y por un *linker* (rico en el amino ácido prolina) y un dominio *coiled-coil* en su parte N-terminal, la cual parece ser periplásmica. En el tercer capítulo de esta Tesis se intentó identificar proteínas con las que pudieran tener algún tipo de interacción tanto HglK como SepJ. El abordaje experimental se hizo mediante el uso de estirpes de *Anabaena* que portaban HglK fusionada a la Strep II o SepJ fusionada a la Strep II o a la GFP. Estas fusiones permitieron la purificación de las proteínas, que se llevó a cabo bajo condiciones no desnaturalizantes, intentando evitar la disgregación de los posibles complejos que pudieran formar con otras proteínas. Las muestras obtenidas tras las purificaciones se analizaron por espectrometría de masas para identificar todas las proteínas presentes en ellas. El análisis de los resultados mostró varias proteínas que podrían interactuar con HglK o SepJ. El primer resultado destacable es la posible interacción entre ambas (SepJ y HglK), aunque la co-purificación de las dos proteínas solamente se observó en los aislados de SepJ. También destaca la presencia de dos proteínas comunes en los aislados de SepJ y HglK. Estas proteínas son Alr2937 y Tic22 (Alr0114), ambas periplásmicas, la segunda de las cuales



con un papel importante en la formación de la membrana externa de *Anabaena*. Finalmente, mientras que no se ha encontrado ninguna proteína que interaccione solamente con HgIK, para SepJ se detectó Alr1690, una proteína anclada a la membrana plasmática que presenta dominios periplásmicos de unión al peptidoglicano.

Continuando con el estudio de las proteínas septales, se quiso profundizar en el conocimiento de las proteínas SepJ, FraC y FraD de otras cianobacterias filamentosas formadoras de heterocistos, concretamente de cianobacterias simbióticas con diatomeas marinas. En el cuarto capítulo de esta Tesis, se estudiaron estas proteínas septales codificadas por los genes de las cianobacterias *Richelia intracellularis* (RintHH01) y *Calothrix rhizosoleniae* (CalSC01). *R. intracellularis* es un endosimbionte obligado que vive en el citoplasma de la diatomea *Hemiaulus hauckii*, mientras que *C. rhizosoleniae* es un exosimbionte facultativo de la diatomea *Chaetoceros compressus*, pudiendo proliferar anclada extracelularmente a la diatomea o como un organismo de vida libre. En *Anabaena*, *fraC* y *fraD* forman parte del operón *fraCDE* que también se encuentra presente en CalSC01 y, sólo *fraCD*, en RintHH01. Respecto a SepJ, tanto RintHH01 como CalSC01 contienen genes que determinan proteínas homólogas a la de *Anabaena*. Los mutantes *sepJ* y *fraC-fraD* de *Anabaena* muestran un fuerte fenotipo de fragmentación e imposibilidad de crecer diazotróficamente, ambos fenotipos de fácil estudio, por lo que se intentó la complementación de dichos mutantes con los genes de las cianobacterias simbióticas. Los genes *sepJ* y *fraCD* de RintHH01 y los genes *sepJ* y *fraCDE* de CalSC01 se obtuvieron por síntesis química, en el caso de *sepJ* de CalSC01 con un uso de codones adaptado a *Anabaena*, para ser introducidos en mutantes de *Anabaena* que portaban las deleciones *sepJ* (estirpe CSVM34) o *fraC-fraD* (estirpe CSVT22), respectivamente. La complementación de CSVM34 con *sepJ* de RintHH01 y de CalSC01 se hizo de forma que la expresión de estos tenía lugar desde el promotor natural del *sepJ* de *Anabaena*. Sin embargo, ni la complementación con *sepJ* de RintHH01 ni la realizada con *sepJ* de CalSC01 consiguieron revertir el fenotipo mutante, ni siquiera parcialmente. En cambio, sí se consiguió revertir al fenotipo silvestre cuando el mutante se complementó con *sepJ* de la propia *Anabaena* utilizado como control positivo. Por otro lado, los genes *fraC-fraD* de RintHH01 se transfirieron a la estirpe CSVT22 con dos estrategias distintas: en una su expresión tenía lugar desde el promotor natural *fraC* de *Anabaena*, y en otra la expresión tenía lugar desde el promotor *glnA* (promotor de fuerte expresión en condiciones de deficiencia de nitrógeno) insertado en un plásmido replicativo de *Anabaena*. Por su parte, la complementación de CSVT22 con los genes *fraC-fraD-fraE* de CalSC01 solamente se intentó con expresión desde el promotor *glnA* en un plásmido replicativo. Nuevamente ninguna de las complementaciones abordadas tuvo éxito. Es más, las expresiones génicas realizadas bajo el promotor *glnA* causaban efectos perjudiciales para *Anabaena*, como pudimos comprobar con controles utilizando los propios genes de *Anabaena*, lo que sugiere que quizás no habíamos expresado los distintos genes en sus niveles óptimos de funcionalidad.

La última parte de esta Tesis se dedicó a la realización de un breve estudio de la proteína FraE, siendo esta la otra proteína codificada en el operón *fraC-fraD-fraE* de *Anabaena*. FraE es una proteína de membrana necesaria para la integridad del filamento. En este capítulo se estudió su localización, para lo que se construyó una estirpe de *Anabaena* que producía FraE-sfGFP. Los ensayos de análisis de fluorescencia de la GFP llevados a cabo con esta estirpe pusieron de manifiesto que FraE se localiza principalmente en los polos de los heterocistos y que sus niveles disminuyen a las 48 horas de inducción en un medio sin nitrógeno combinado.



## Conclusiones

1. En la cianobacteria filamentosa formadora de heterocistos *Anabaena* sp. PCC 7120 existe una fuerte correlación entre el número de nanoporos septales y la transferencia intercelular de calceína cuando los filamentos son incubados en un medio sin nitrógeno combinado. Esta observación apoya la idea de que los nanoporos albergan los nexos septales que median la transferencia molecular intercelular, que es esencial para el crecimiento diazotrófico.
2. HglK es una *pentapeptide repeat protein* necesaria para el crecimiento diazotrófico de *Anabaena*. Los mutantes *hglK* muestran una expresión disminuida de algunos genes que se expresan durante la diferenciación de los heterocistos y están afectados tanto en la integridad del filamento como en la formación de nanoporos y en la transferencia molecular intercelular. La proteína HglK presenta una localización preferente en los polos celulares. Concluimos que HglK contribuye a la formación de estructuras septales funcionales necesarias para la diferenciación de los heterocistos maduros en *Anabaena*.
3. El análisis por espectrometría de masas de preparaciones de las proteínas HglK y SepJ de *Anabaena* marcadas con etiquetas moleculares ha identificado varias proteínas que podrían interactuar con aquéllas. Para SepJ destacamos la propia HglK y la proteína de unión de peptidoglicano Alr1690. Además, se detectaron dos proteínas periplásmicas que podrían interactuar tanto con SepJ como con HglK: Alr2947 y la chaperona Tic22 (Alr0114). Así, SepJ y HglK son dos proteínas septales que pueden interactuar entre sí y con algunas otras proteínas periplásmicas.
4. Las cianobacterias filamentosas formadoras de heterocistos *Richelia intracellularis* y *Calothrix rhizosoleniae*, simbióticas de diatomeas marinas, contienen genes codificantes de SepJ y de FraC, FraD y, en el caso de *C. rhizosoleniae*, FraE. Sin embargo, no hemos tenido éxito en la complementación de mutantes *sepJ* y *fraC-fraD* de *Anabaena* con los genes de estas cianobacterias simbióticas.
5. La proteína FraE presenta una localización preferente en los polos de los heterocistos de *Anabaena*, lo que es consistente con el hecho de que los mutantes *fraE* de esta cianobacteria producen heterocistos que carecen de la estructura polar característica conocida como *heterocyst neck*.



# 1. Introduction

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**





## 1.1 Multicellularity in bacteria

Multicellularity is a form of biological organization in which individual cells aggregate, working together as a single organism and showing greater complexity than when they are separated (Herrero *et al.*, 2016). Multicellularity marks a new form of evolutionary innovation, in which the advantages of aggregate life must be greater than the disadvantages that may appear such as competition for local resources, reduced mobility or the existence of non-cooperating cells that lead to the failure of such cooperation (Bonner, 1998; Szathmary & Smith, 1995; West *et al.*, 2006). This form of organization has appeared several times throughout evolution and is not a characteristic just of eukaryotes (Bonner, 1998; Claessen *et al.*, 2014). All forms of multicellularity have some common characteristics such as some mechanism of cell-cell adhesion, intercellular communication and coordination (Claessen *et al.*, 2014).

There are different forms of bacterial multicellularity including the formation of a biofilm based on the production of an extracellular matrix that surrounds the cells, and where some of them can differentiate to carry out different functions, such as the production of spores in biofilms of *Bacillus subtilis*. Another example of multicellular bacteria is the swarming bacteria such as myxobacteria, which are soil bacteria wrapped in a matrix allowing them to move as a coordinated group through the soil and to produce spores in the absence of nutrients (Muñoz-Dorado *et al.*, 2016). Organisms that grow as chains of cells represent another type of multicellularity. Within filamentous organisms we can find very different structures (Herrero *et al.*, 2016). A concrete example can be the Streptomycetes which is a group of Gram-positive bacteria belonging to the phylum of Actinobacteria. Under adverse conditions Streptomycetes can produce spores that, when those conditions revert, can germinate giving rise to vegetative hyphae divided into compartments by cross-walls (Jakimowicz & van Wesel, 2012). Other different phylogenetic groups such as cyanobacteria, *Desulfobulbaceae* (Pfeffer *et al.*, 2012) and *Lachenospiraceae* (Thompson *et al.*, 2012) contain filamentous forms, while bacteria of the genera *Chloroflexus* (Pierson & Castenholz, 1974), *Beggiatoa* (Strohl & Larkin, 1978) and *Entotheonella* (Wilson *et al.*, 2014) are filamentous. In filamentous cyanobacteria, the separation between daughter cells is incomplete after cell division, so that the outer membrane is continuous along the filament (Wilk *et al.*, 2011).

## 1.2 Cyanobacteria

### 1.2.1 Origin, diversity and classification of cyanobacteria

Cyanobacteria are among the oldest multicellular organisms on Earth. They constitute a monophyletic group of Gram-negative bacteria, are considered the ancestors of plastids (Giovannoni *et al.*, 1988), and are of great ecological importance since they are the only bacteria that can perform oxygenic photosynthesis (Herrero *et al.*, 2016). Cyanobacteria were responsible for the accumulation of oxygen in the atmosphere that led to the Great Oxygenation Event (GEO) 2.4-2.3 billion years ago, which provoked major changes in the life forms on our planet (Lyons *et al.*, 2014). Oxygenic photosynthesis affects the carbon cycle since it efficiently removes atmospheric CO<sub>2</sub> by fixing it to produce organic matter, thus contributing to primary production which is basic in the biosphere's food chain. In addition, cyanobacteria also intervene in the nitrogen cycle since some of them can fix atmospheric nitrogen, which can be assimilated by other organisms (Knoll, 2004, 2008).

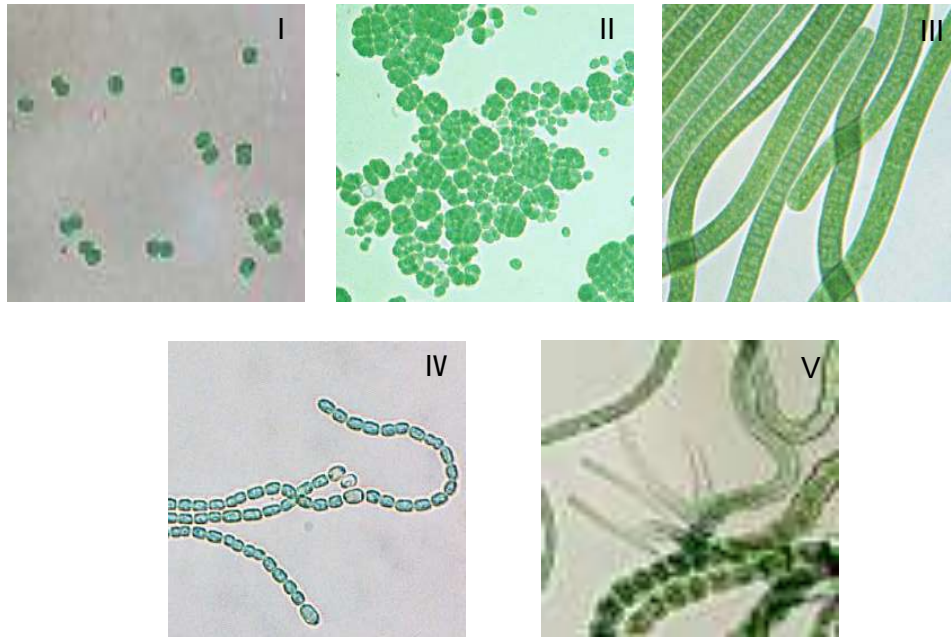


In prokaryotes one of the most diverse phyla is that of cyanobacteria, found in all types of habitats including soils, fresh water, oceans or extreme environments. Moreover, we can even find them in symbiosis with other organisms such as plants or fungi (Meeks & Elhai, 2002). They also show a great cellular diversity presenting unicellular forms and filamentous forms within which different cell types are found (Herrero *et al.*, 2016). Thus, there are different processes of cellular differentiation in filamentous cyanobacteria. Under conditions of nitrogen deficiency, specialized cells called heterocysts are formed that have the capacity to fix atmospheric nitrogen. In these cells, micro-oxic conditions prevail and the oxygen-sensitive nitrogen-fixation machinery can function. Under some particular conditions, other types of differentiated cells are formed including akinetes, which are resistance cells, and hormogonia, which are short filaments formed by cells of small size that have a dispersal function (Flores & Herrero, 2014). Due to an appreciable metabolic versatility and their ability to adapt to different environments, cyanobacteria have a great biotechnological potential as fertilizers, for production of healthy food or bioactive components, and in bioremediation processes (Abed *et al.*, 2009). In spite of this great diversity within cyanobacteria, all of them have a common ancestor (Schirmeister *et al.*, 2011; 2013), with filamentous cyanobacteria being one of the oldest examples of multicellular organisms on Earth (Giovannoni *et al.*, 1988; Tomitani *et al.*, 2006; Shih *et al.*, 2013). Interestingly, some current single-cell species are derived from filamentous forms that seem to have lost the ability to form filaments (Schirmeister *et al.*, 2011).

Based on morphology, cyanobacteria have historically been classified into five taxonomic sections or orders (Rippka *et al.*, 1979; Castenholz, 2001). Section I (order Chroococcales) includes unicellular forms that reproduce by binary fission; section II (Pleurocapsales) is composed of unicellular cyanobacteria that reproduce by multiple fission giving rise to small cells called baeocytes or by both multiple and binary fission; section III (Oscillatoriales) contains filamentous cyanobacteria that divide in a single plane (perpendicular to the major filament axis); finally, sections IV (Nostocales) and V (Stigonematales) have a monophyletic origin and are composed of filamentous cyanobacteria capable of forming heterocysts, but the difference between them is that in section V the angle of the division plane can change producing branched filaments (Fig. 1). New phylogenetic trees based on molecular analysis have generated controversy about the distribution of the different sections, since only sections IV and V are of monophyletic origin consistently with their genetic and morphological coherence (Schirmeister *et al.*, 2015). Phylogenetic analysis has also shown that multicellularity in cyanobacteria appeared very early in evolution, although all extant cyanobacteria evolved from a common unicellular ancestor (Tomitani *et al.*, 2006; Shih *et al.*, 2013) (Fig. 2). Some strains in sections III, IV and V have the capacity to form hormogonia, while the formation of akinetes is framed in sections IV and V, although this characteristic is not found in all strains.

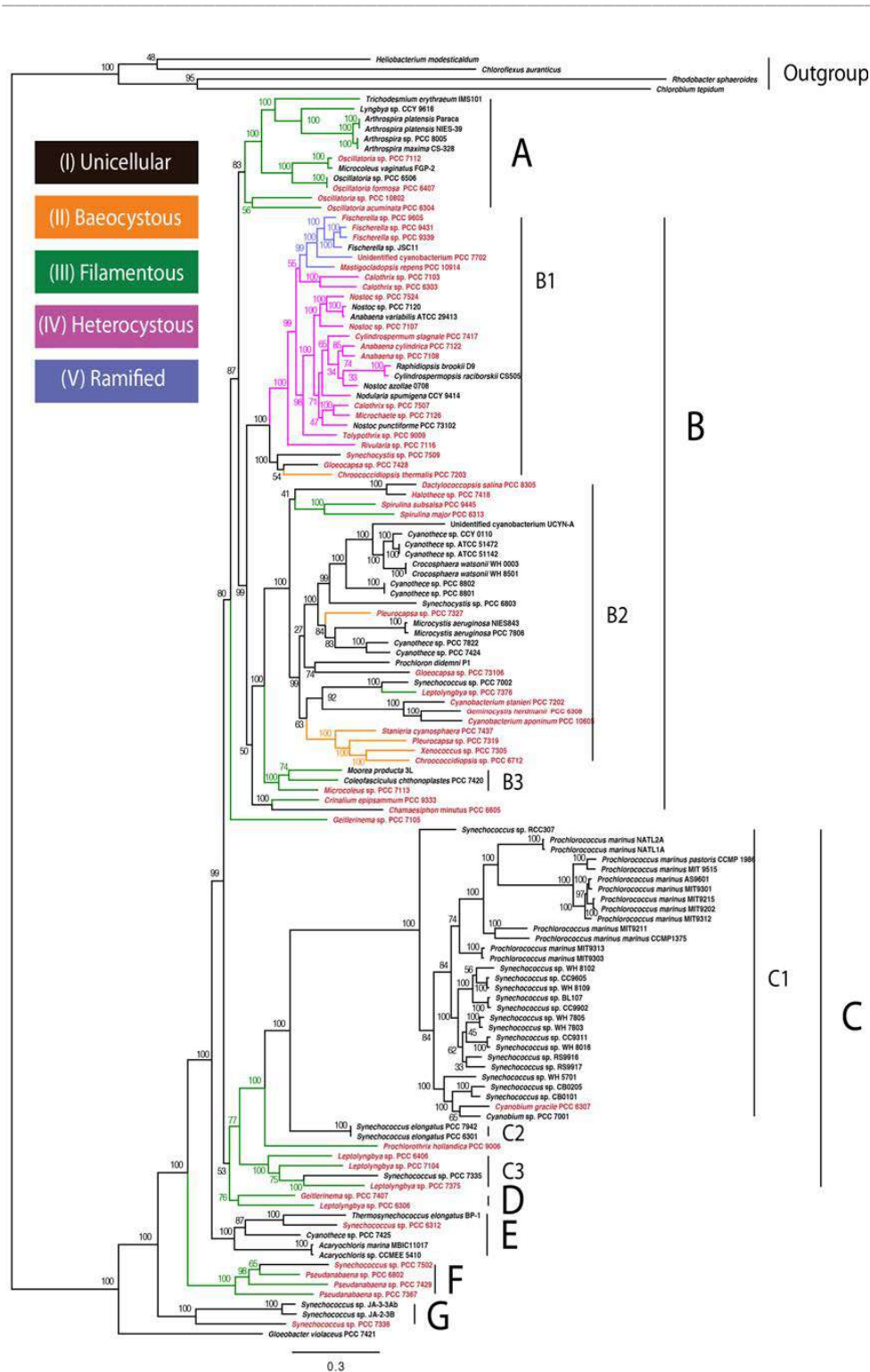
This work has addressed aspects of the biology of heterocyst-forming cyanobacteria, using for experiments the model organism *Anabaena* sp. PCC 7120 (order Nostocales), hereafter referred to as *Anabaena*. Additionally, aspects of symbiotic cyanobacteria have been addressed based on genomic information from strains of *Richelia intracellularis* and *Calothrix rhizosoleniae*, both of which belong also to the Nostocales.





**Fig.1. Examples of the different taxonomic groups of cyanobacteria defined by Rippka *et al.* (1979).**  
 A representative of each section is shown: Section I, *Synechocystis* sp. PCC 6803 (image courtesy of Miguel Roldán, IBVF, CSIC and Universidad de Sevilla; Section II, *Chroococcidiopsis* sp. PCC 7203; Section III, *Oscillatoria* sp. PCC 9325; Section IV, *Anabaena* sp. PCC 7120; Section V, *Fischerella* sp. Images of section II, III and V strains, courtesy of José E. Frías, Servicio de Cultivos Biológicos, CSIC and Universidad de Sevilla.





**Fig. 2 Cyanobacterial phylogeny.** Cyanobacterial species tree from Shih *et al.* (2013) performed with genome sequences from the indicated strains. Branches are colored coded according to the classification of Ripka *et al.* (1979).



### 1.2.2 Physiology of cyanobacteria

The phylum cyanobacteria is defined by a metabolic characteristic that is its ability to perform oxygenic photosynthesis. In this process, cyanobacteria use two photosystems, called PSI and PSII, to generate an electrochemical potential using water as an electron donor and releasing oxygen as a by-product of the reaction, and the electrochemical potential generated is conserved as “assimilatory power” in the form of NADPH and ATP (Hohmann-Marriott & Blankenship, 2011). The cyanobacteria are able to fix atmospheric CO<sub>2</sub> through the reductive pentose phosphate cycle (Calvin-Benson-Bassham cycle) in which ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the enzyme responsible for carrying out the fixation of CO<sub>2</sub> producing two molecules of 3-phosphoglycerate that in successive steps will be reduced rendering the sugars needed to form the carbon skeletons necessary in the processes of biosynthesis. However, Rubisco has a low affinity for CO<sub>2</sub>, so the cells have developed CO<sub>2</sub> concentrating mechanisms (CCMs) to favor the carboxylase reaction over the oxygenase reaction. These CCMs are, basically, a system of membrane transporters for bicarbonate and CO<sub>2</sub>; an enzyme called carbonic anhydrase (CA) that catalyzes the interconversion of bicarbonate to CO<sub>2</sub> and water; and finally, a compartment called carboxysome in which Rubisco and CA are localized, and which prevents CO<sub>2</sub> leakage into the cytosol facilitating increased CO<sub>2</sub> fixation (Cameron *et al.*, 2014).

Although cyanobacteria are generally autotrophic, some strains, especially in the Nostocales, can behave as facultative photoheterotrophs using different sugars such as glucose, fructose and sucrose as carbon source, and some of them can even grow as chemoheterotrophs using the sugars as carbon and energy source (Rippka, 1972, 1979). An important mode of growth seems to be mixotrophy, in which the cells grow photoautotrophically with further support from a sugar (Nieves-Mori3n & Flores, 2018). Known sugar transporters in the Nostocales include a Major Facilitator Superfamily (MFS) transporter for glucose (Ekman *et al.*, 2013) and ATP-Binding Cassette (ABC) transporters for fructose (Ungerer *et al.*, 2008; Ekman *et al.*, 2013) and glucosides (Nieves-Mori3n & Flores, 2018). In *Nostoc punctiforme*, the glucose transporter has been found to be required to establish symbiosis with plant tissues, highlighting the possible importance of sugars as signaling agents in cyanobacteria (Ekman *et al.*, 2013; Picossi *et al.*, 2013).

Another important aspect of the physiology of cyanobacteria is their nitrogen metabolism. These organisms can use different types of combined nitrogen molecules, including ammonium, nitrate, nitrite, urea and some amino acids (using primarily those whose utilization involves less energy expenditure); in addition, some strains are capable of fixing atmospheric dinitrogen in the absence of another source of available nitrogen (Herrero *et al.*, 2001). The incorporation of ammonium and urea by living organisms can take place by diffusion through the plasma membrane, but cyanobacteria express ammonium and urea transporters that increase their efficiency in taking up these substrates from low environmental concentrations. The ammonium transporters belong to the Amt family and have been described in different cyanobacteria such as *Synechocystis* sp. PCC 6803 (Montesinos *et al.*, 1998), *Synechococcus elongatus* (V3zquez-Bermúdez *et al.*, 2002c; Paz-Yepes *et al.*, 2007) and *Anabaena* (Paz-Yepes *et al.*, 2008). Also, an ABC-type urea transporter (Urt transporter) has been described in *Anabaena* which is encoded in an operon, *urtABCDE* (Valladares *et al.*, 2002). Finally, there are also nitrate (and nitrite) transporters of two types, an ABC-type transporter (NrtABCD) (Flores *et al.*, 2005) or an MFS permease (NapA or NrtP) (Wang *et al.*, 2000).





Once the nitrogenous compounds are inside the cell, their metabolization begins. Nitrate is reduced by ferredoxin-dependent nitrate reductase (NarB) to nitrite and nitrite is reduced by ferredoxin-dependent nitrite reductase (NirA) to ammonium. Both enzymes and the nitrate/nitrite transporters (NrtABCD or NrtP) are encoded in the same operon highlighting the relationship between them (Flores *et al.*, 2005). On the other hand, urea is hydrolyzed by means of urease to ammonium, releasing CO<sub>2</sub> in the process (Valladares *et al.*, 2002). Finally, the ammonium obtained by catabolism of other nitrogenous compounds as well as that obtained from the medium is incorporated into organic compounds through the glutamine synthase/glutamate synthase (GS/GOGAT) pathway (Luque & Forchhammer, 2008). Glutamine is first synthesized from ammonium and glutamate via GS in a process that consumes ATP. Then the glutamine-amide N is transferred to 2-oxoglutarate (2-OG) producing two glutamate molecules, a reducing transamination catalyzed by GOGAT. Because 2-OG is an intermediate of the Krebs cycle, it joins the metabolism of carbon with that of nitrogen. It should be noted that 2-OG is an important metabolite for cyanobacteria since it is an indicator of the C/N balance of the cell (Muro-Pastor *et al.*, 2001; Vázquez-Bermúdez *et al.*, 2003).

Recalling what was mentioned a few paragraphs above, in the absence of combined nitrogen some cyanobacteria can fix atmospheric dinitrogen. This is an energy and reductant-dependent process in which dinitrogen is reduced to produce two ammonium molecules, a reaction catalyzed by an enzyme called nitrogenase (Rubio & Ludden, 2008). The nitrogen fixation machinery is very sensitive to oxygen and becomes useless in its presence, which is an especial problem in cyanobacteria since they are organisms that perform oxygenic photosynthesis. To solve this problem, cyanobacteria have developed different mechanisms. Some unicellular and some filamentous cyanobacteria carry out a temporary separation between nitrogen fixation and carbon fixation; that is, they fix carbon during periods of light, since this is directly linked to photosynthesis, while nitrogen fixation is carried out in periods of darkness where photosynthesis is not active and there is not oxygen release (Stal & Zher, 2008). On the other hand, some filamentous cyanobacteria have developed a spatial separation of both processes, producing a type of cell specialized in nitrogen fixation called heterocyst, where carbon fixation does not occur. In this type of cell, which maintains a micro-oxic environment, all the necessary machinery for nitrogen fixation is confined (Flores & Herrero, 2010).

Depending on the initial source of nitrogen, the energy expenditure varies making it necessary a good regulation of nitrogen metabolism. The NtcA protein belongs to the CRP family of transcriptional regulators and is preserved among the different cyanobacteria. NtcA is a very important transcription factor in nitrogen metabolism since it influences the expression of many genes involved in nitrogen assimilation, and it also plays an essential role in the differentiation of heterocysts in the filamentous cyanobacteria that have this option (Herrero *et al.*, 2013). As mentioned above, ammonium is a preferential nitrogen source for cyanobacteria. In conditions that do not limit this compound, alternative mechanisms of nitrogen assimilation are repressed, but when this is not the case, NtcA promotes the expression of all the necessary machineries to assimilate alternative sources of nitrogen. The activity of NtcA is modulated by 2-oxoglutarate, by the PII signal transduction protein, and by the PII-binding protein PipX. The action of NtcA is stimulated by binding of 2-OG whose concentration increases when nitrogen levels are low (Vázquez-Bermúdez *et al.*, 2002b; Valladares *et al.*, 2008). Additionally, PipX can bind to NtcA stabilizing its active confirmation (Espinosa *et al.*, 2006), but PipX can also bind to PII preventing binding (of PipX) to NtcA. The mechanism behind these relations is that 2-OG also binds to PII



producing a conformational change that prevents PipX from binding, thus leaving more PipX available for stabilizing NtcA (Espinosa *et al.*, 2014). This demonstrates the importance of 2-OG in NtcA activation and highlights that N control indeed responds to the availability of C and N rather than simply to the levels of ammonium (Vázquez-Bermúdez *et al.*, 2003).

The structure of NtcA is that of a homodimer in which each subunit contains a helix-turn-helix domain; its form of action is to bind to DNA, where it recognizes the consensus sequence GTAN<sub>8</sub>TAC (Luque *et al.*, 1994), acting as a repressor or an activator of transcription depending on the localization of its binding site in the gene promoter (Picossi *et al.*, 2014). The NtcA-activated promoters frequently follow the pattern of bacterial Class II activator-dependent promoters (Busby & Ebright, 1999), since the NtcA-binding sequence is located about 41.5 nucleotides upstream of the transcription start point for the gene, and the promoter also includes a -10 box which has as consensus sequence TAN<sub>3</sub>T (Luque *et al.*, 1994). However, when NtcA binds the DNA overlapping the -10 box, it acts as a repressor (Vázquez-Bermúdez *et al.*, 2002a). In spite of this consensus, some promoters activated by NtcA have been found in which the binding site is located upstream of position 41.5, others in which the sequence for binding differs from the consensus, and even some genes regulated by NtcA have been found without any recognizable binding site (Herrero *et al.*, 2004).

### 1.2.3 Cell envelope

The structure of the cyanobacterial cell envelope is that described in all Gram-negative bacteria, which is composed of a cytoplasmic membrane (CM), a peptidoglycan layer (PG) located in an area known as the periplasm, and finally an outer membrane (OM) above this layer (Flores *et al.*, 2006; Wilk *et al.*, 2011). In filamentous cyanobacteria, there is the unusual characteristic that the OM is continuous along the entire filament, making the periplasm common among the different cells, not producing an absolute septal separation between cells, and helping to maintain the integrity of the filament (Mariscal *et al.*, 2007; Burnat *et al.*, 2014). In addition, some cyanobacteria can produce an extra layer with functions of protection, cell integrity, adhesion and swimming mobility (in those cyanobacteria that can do so) (Liberton & Pakrasi, 2008). This paracrystalline layer, called the S-layer, is of non-membranous nature and is localized outside of the outer membrane (Hahn & Schleiff, 2014).

#### 1.2.3.1 Outer membrane

The outer membrane is composed of an asymmetric bilayer, where the inner layer is composed of phospholipids and the outer layer is rich in lipopolysaccharide (LPS) (Bos *et al.*, 2007). In addition, it contains a series of integral membrane proteins that provide a selective permeability for different molecules; these proteins are called porins and consist of antiparallel amphipathic  $\beta$ -chains that form a  $\beta$ -barrel cylindrical structure (Koebnik *et al.*, 2000; Nikaido, 2003). Four of the nine porins encoded in the *Anabaena* genome are readily detected in outer membrane preparations (Nicolaisen *et al.*, 2009a; Moslavac *et al.*, 2005). The structure of lipopolysaccharide in cyanobacteria may not be identical to that best known of the enterobacteria (Hahn & Schleiff, 2014; Simkovsky *et al.*, 2016). Nonetheless, three parts can be differentiated: (i) the anchorage zone of lipid A to the membrane; (ii) the core region that serves as a link between lipid A and the outer polysaccharide; and (iii) the O-antigen, which is a chain of polysaccharides whose



composition in sugars and length is variable between bacteria but specific to bacterial strains (Sperandeo *et al.*, 2019; Xu *et al.*, 1997).

### 1.2.3.2 The peptidoglycan layer

The peptidoglycan layer is found in the periplasm, the space between the outer and cytoplasmic membranes. This periplasmic zone provides the oxidative environment necessary for some enzymes present there to catalyze the formation of disulfide bonds (Nakamoto & Bardwell, 2004). The periplasm is a zone in which there is no source of energy in the form of ATP but, in spite of this, there are many factors that modulate the position of the polymers of the envelope. In addition, some proteins that transport small molecules can be found in the periplasm (Ruiz *et al.*, 2006).

The peptidoglycan is basically composed by a repetition of a disaccharide made of  $\beta$ -1-4-linked N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) that is linked by peptide side chains (Ruiz *et al.*, 2016), acting all of this structure as a huge single molecule (the murein sacculus) (see Fig. 5 below). This sacculus determines the cellular form, provides osmotic resistance, and serves as an anchor for many components of the cellular envelope, as for example occur in Gram-positive bacteria in which it serves as an anchor to teichoic acids (Vollmer *et al.*, 2008). Although the basic composition of PG is as indicated above, there are many chemical, structural and architectural differences between different bacterial strains (Turner *et al.*, 2014). Hence, it is interesting to indicate that in cyanobacteria the degree of density and cross-linking between the different chains of peptidoglycan is more similar to that of Gram-positive than to that of Gram-negative bacteria (Hoiczky & Hansel, 2000). Specifically, in the case of *Anabaena*, the PG has two or three layers and is found wrapping each cell individually, being also present between the cells at the intercellular septa (Flores *et al.*, 2006; Wilk *et al.*, 2011).

The murein sacculi of some filamentous cyanobacteria, such as *Anabaena* or *Nostoc punctiforme*, can be isolated as a single structure encompassing several cell units, allowing their study by transmission electron microscopy (TEM). Their study has shown the presence of a rigid disk-shaped septal structure composed of the PG from the adjacent cells. These disks bear perforations in their central part that are 15-20 nm in diameter and have been called "nanopores" (Lehner *et al.*, 2013; Nürnberg *et al.*, 2015). It has been proposed that nanopores are openings in the peptidoglycan disk that hold protein structures that connect adjacent cells (Mariscal *et al.*, 2014). The size of nanopores appears to be regulated, since a protein called SjcF1 that is conserved in filamentous cyanobacteria and could be involved in nanopore size determination has been described in *Anabaena* (Rudolf *et al.*, 2015). SjcF1 is a membrane-anchored, mostly periplasmic protein that contains three functional domains, two peptidoglycan-binding domains and a protein-protein interaction SH3 domain. *Anabaena* mutants of this protein present nanopores wider and of more variable diameter than the wild-type strain (Rudolf *et al.*, 2015).

PG is a dynamic structure since it has to be hydrolyzed to allow the incorporation of new peptidoglycan subunits and, subsequently, covalent bonds are re-established, so that the structure of PG is not altered during cell growth. The proteins that perform this hydrolysis are called autolysins, which are found in the periplasm and not only play an essential role in cell growth but are also important to determine the shape, size and density of the PG, as well as for the separation of the daughter cells after cell division (Typas *et al.*, 2012). Although genes encoding autolysins have been studied mainly in model organisms such as *E. coli*, their homologues have also been





found in cyanobacteria (Hahn & Schleiff, 2014). Autolysins are generally PG amidases (Vollmer *et al.*, 2008; Typas *et al.*, 2012). In *E. coli*, there are three known amidases: AmiA, AmiB and AmiC. Amidase AmiC has two variants in heterocystous cyanobacteria, AmiC1 and AmiC2 (Zhu *et al.*, 2001). However, the role of AmiC amidases in filamentous cyanobacteria is complex because there is no hydrolysis of the peptidoglycan and invagination of the outer membrane at cell division; instead, it is known that AmiC1 in *Anabaena* and AmiC2 in *Nostoc punctiforme* are necessary for the differentiation of heterocysts and formation of nanopores (Lehner *et al.*, 2011; Berendt *et al.*, 2012). In heterocystous cyanobacteria, *amiC1* and *amiC2* are found within a gene cluster with *murI*, which codes for a glutamate racemase (*amiC1-amiC2-murI*). In some strains, a gene named *amiC3* has also been found, but it is not located within that cluster (Bornikoel *et al.*, 2017; Zheng *et al.*, 2017).

### 1.3 Heterocysts

Heterocysts are a differentiated type of cell in which the conditions necessary for the fixation of atmospheric dinitrogen occur. Heterocysts are present in strains of filamentous cyanobacteria that belong to taxonomic sections IV and V (Rippka *et al.*, 1979). They are easily recognized along the filament, because they have a shape and size different from the rest of the cells. Regarding to their localization in the filament, this can be intercalary (between vegetative cells) or terminal (at the end of the filament).

#### 1.3.1 Characteristics of the heterocyst

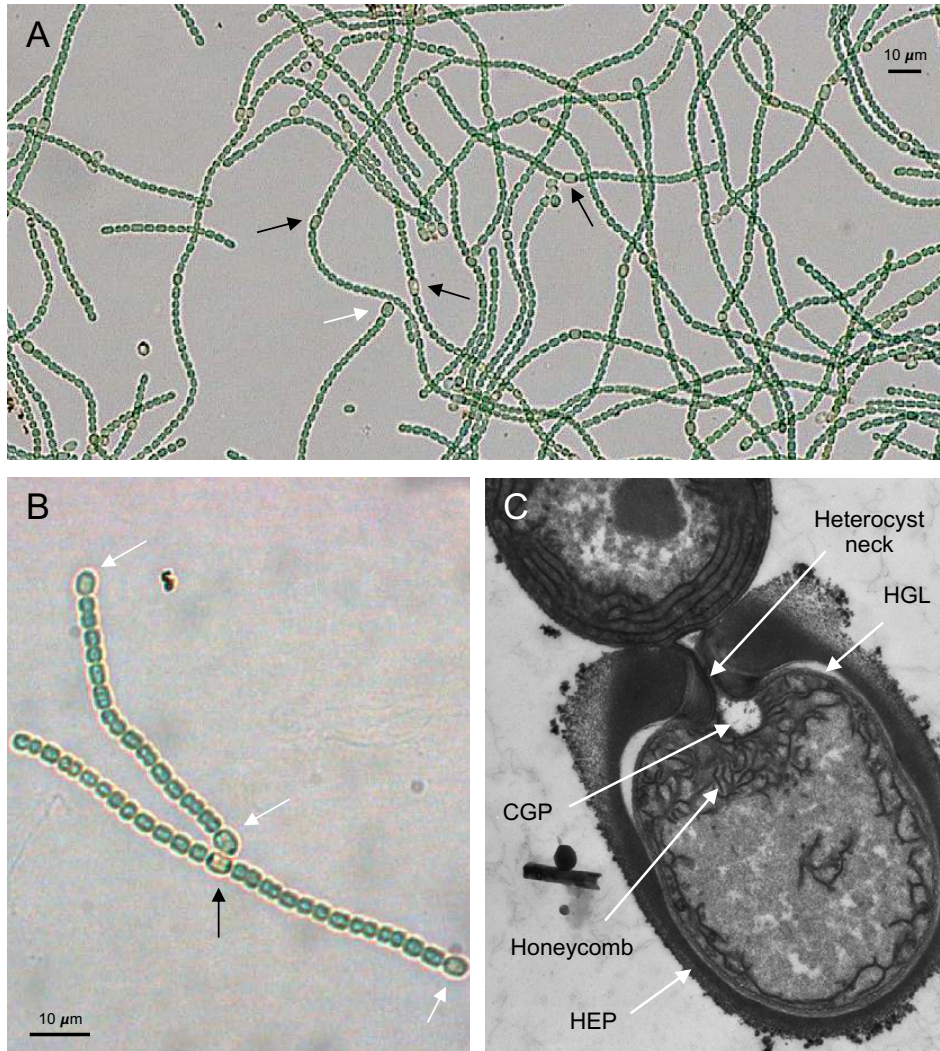
In order to facilitate the description of the different characteristics of the heterocyst, this section will be divided into two sub-sections, one of which will refer to structural characteristics and the other one to metabolic aspects.

##### 1.3.1.1 Structural characteristics

Beyond size and shape, there are certain structures (or processes that take place during heterocyst differentiation) that have a role in providing adequate conditions for nitrogen fixation in the heterocysts and that clearly differ from those found in the vegetative cells. (i) A reorganization of intracellular membranes places the thylakoid membranes arranged near the heterocyst cell poles in the so-called "honeycomb" structure (Lang & Fay, 1971); a high respiratory rate takes place in the *honeycomb* membranes that allows the elimination of oxygen thus contributing to avoid the inactivation of nitrogenase (Valladares *et al.*, 2007). (ii) A loss from the cytoplasm of carboxysomes and glycogen granules can be seen, as well as a narrowing of the cytoplasm in the heterocyst poles giving rise to a structure known as the "heterocyst neck", which makes the septum between heterocysts and vegetative cells narrower than between vegetative cells; this could help to control the entry of gases in the heterocyst (Fig. 3) (Walsby, 2007). (iii) Granules of cyanophycin also appear at the heterocyst poles; these granules are composed of a polymer of arginine and aspartic acid of non-ribosomal synthesis that is a nitrogen reservoir (Fig. 3) (Lang *et al.*, 1972). Finally, (iv) the heterocysts have a thick envelope that provides protection against gas diffusion and consists of heterocyst-specific glycolipids (HGL) and polysaccharides (HEP) deposited outside of the outer membrane (see Fig. 3) (Cardemil & Wolk, 1979; Flores & Herrero, 2010). The layer closest to the outer membrane is the glycolipid layer that functions as a barrier against gases (Fay, 1992), whereas the polysaccharide layer protects the glycolipid layer from physical damage (Xu *et al.*, 2008). These extra wrappers are important because the lack or harm



of any of them impair the diazotrophic growth of the filament (Fan et al., 2005; Huang et al., 2005).



**Fig. 3. Heterocyst morphology.** (A and B) Filaments of *Anabaena* sp. PCC 7120 grown without combined nitrogen photographed with the optical microscope with the 20x and 40x objectives respectively. The white arrows show terminal heterocysts and the black arrows show intercalary heterocysts. (C) Transmission electron micrograph of a fragment of an *Anabaena* filament showing a vegetative cell (top cell) and a terminal heterocyst (bottom cell). The image shows the place that the cyanophycin granule normally occupies (CGP), the neck of the heterocyst, the *honeycomb* membranes, the heterocyst glycolipid layer (HGL) and the heterocyst polysaccharide layer (HEP). Image C by Victoria Merino-Puerto (CSIC, Seville) and Iris Maldener (Universität Tübingen, Germany).



### 1.3.1.2 Metabolic characteristics

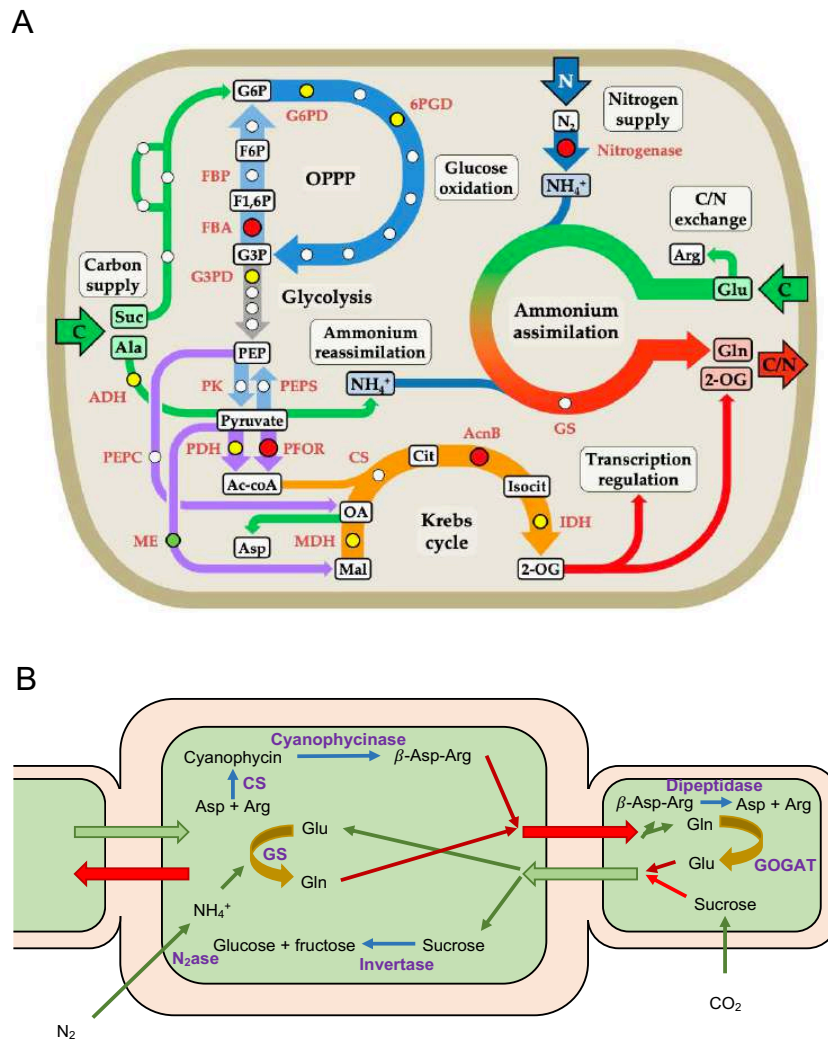
From a metabolic point of view, heterocysts also have some characteristics that differentiate them from the rest of the cells of the filament. Among these characteristics, we can highlight: (i) the inability to carry out oxygenic photosynthesis due to the loss of PSII activity; (ii) the inability to fix atmospheric CO<sub>2</sub> as a consequence of repression of Rubisco; (iii) the increase in respiratory rate to reduce oxygen levels; and (iv) the expression of all the machinery necessary to fix atmospheric dinitrogen (Golden & Yoon, 2003; Herrero *et al.*, 2004). Although not a metabolic characteristic itself, it should be noted that the heterocysts are terminally differentiated cells, so they are unable to divide once they have completed cell differentiation and cannot revert to vegetative cells (Wang & Xu, 2005; Kuhn *et al.*, 2000).

Due to the inability of heterocysts to perform certain functions (photosynthesis) and their capacity to perform others (N<sub>2</sub> fixation), there must be an exchange of molecules between the heterocysts and the vegetative cells. For example, the heterocysts transfer glutamine and a dipeptide ( $\beta$ -aspartyl-arginine) to the vegetative cells and receive sucrose and some amino acids (glutamate, alanine) from the vegetative cells. In order to explain more clearly the importance of this exchange of metabolites, its involvement in carbon and nitrogen metabolism will be presented below.

As mentioned above, the nitrogen-fixing machinery is very sensitive to oxygen, and the heterocysts have several ways of maintaining a micro-oxic environment. The oxygen that enters the heterocyst is eliminated by the high respiratory rate in the *honeycomb* membranes and the action of flavodiiron proteins that absorb O<sub>2</sub> (Valladares *et al.*, 2003, 2007; Ermakova *et al.*, 2014). The synthesis of ATP necessary for nitrogen fixation depends on respiration as well as on PSI-dependent photophosphorylation (Herrero *et al.*, 2013). Because the heterocysts are unable to fix atmospheric CO<sub>2</sub>, they need a contribution of reduced carbon from the vegetative cells. The vegetative cells transfer reduced carbon to the heterocysts mainly in the form of sucrose, which is split into glucose and fructose by heterocyst-specific invertases (López-Igual *et al.*, 2010; Vargas *et al.*, 2011).

In relation to nitrogen metabolism, the oxidation of sugars received from vegetative cells produces the necessary reducing power for nitrogenase. Nitrogenase is formed by a tetramer of NifD and NifK that contains the active site of the enzyme, and a dimer of NifH acts as an electron donor to reduce N<sub>2</sub> to ammonium (Rubio & Ludden, 2008). The resulting ammonium is incorporated into glutamate forming glutamine, which is one of the amino acids transferred to the vegetative cells (Wolk *et al.*, 1976; Meeks *et al.*, 1977). The glutamate needed to form glutamine is received from the vegetative cells because the heterocysts possess GS but not GOGAT, making them unable to produce enough glutamate for the incorporation of ammonium (Thomas *et al.*, 1977; Martín-Figueroa *et al.*, 2000). Finally, the heterocysts also transfer to the vegetative cells  $\beta$ -aspartyl-arginine dipeptide, which results from the cyanophycinase-catalyzed degradation of cyanophycin (previously synthesized by cyanophycin synthetase). Once the dipeptide is in the vegetative cells, it is hydrolyzed by an aspartyl-arginine dipeptidase releasing L-arginine and L-aspartic acid, which are used as sources of nitrogen (Burnat *et al.*, 2014) (Fig. 4).





**Fig. 4. Diagram of the main carbon and nitrogen flows between vegetative cells and heterocysts.** (A) Heterocyst metabolism. The OPPP (oxidative pentose phosphate pathway) and the Krebs cycle are the major routes used by heterocysts to generate reducing equivalents, while glycolysis has a smaller contribution. Metalloenzymes (red circles) and enzymes catalyzing reactions that generate (yellow circles) or require (green circles) NAD(P)H are indicated. *Abbreviations for metabolites:* Suc, sucrose; G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; F1,6P, fructose 1,6-bisphosphate; G3P, glyceraldehyde 3-phosphate; PEP, phosphoenolpyruvate; Ac-coA, acetyl-coA; Mal, malate; OA, oxaloacetate; Cit, citrate; Isocit, isocitrate; 2-OG, 2-oxoglutarate. *Abbreviations for enzymes:* G6PD, glucose 6-phosphate dehydrogenase; 6PGD, 6-phosphogluconate dehydrogenase; FBA, class-II fructose biphosphate aldolase; FBP, fructose biphosphatase; G3PD, glyceraldehyde 3-phosphate dehydrogenase; PK, pyruvate kinase; PEPS, PEP synthase; PDH, pyruvate dehydrogenase; PFOR, pyruvate:ferredoxin oxidoreductase; MDH, malate dehydrogenase; CS, citrate synthase; AcnB, aconitase; IDH, isocitrate dehydrogenase; GS, glutamine synthetase. Picture and legend from Pernil & Schleiff (2019). (B) Simplified diagram of the exchange of compounds rich in nitrogen and compounds rich in carbon between the heterocyst and the vegetative cell. It includes the synthesis and degradation of cyanophycin (polymer of Asp + Arg that serves as a nitrogen reservoir in the heterocyst). Asp, aspartic acid; Arg, arginine; CS, cyanophycin synthetase; Gln, glutamine; Glu, glutamic acid; GOGAT, glutamine-oxoglutarate amido transferase (glutamate synthase); GS, glutamine synthetase; N<sub>2</sub>ase, nitrogenase.





### 1.3.2 Heterocyst differentiation and pattern formation

When some filamentous strains of cyanobacteria are under nitrogen deprivation, a process of cell differentiation begins in which about 495 genes are induced and about 196 genes are inhibited producing structural and metabolic changes that result in heterocyst differentiation (Ehira & Ohmori, 2006a; data re-evaluated in Xu *et al.*, 2008). The appearance of multiple non-coding RNA transcripts suggests that this type of nucleic acid may also play an important role in the differentiation process (Mitschke *et al.*, 2011; Flaherty *et al.*, 2011).

The response to nitrogen deprivation is mainly regulated by two transcriptional regulators. On one hand, the *ntcA* gene is early expressed when 2-OG levels increase, and the NtcA protein is in charge of governing the expression of genes that respond to the cellular need for nitrogen (Picossi *et al.*, 2014; Li *et al.*, 2003). On the other hand, *hetR* is also early induced in differentiation, and HetR acts by binding to DNA as a homodimer (Kim *et al.*, 2011) or a tetramer (Valladares *et al.*, 2016) and regulates the expression of some genes related to the development, function and pattern of the heterocysts (Black *et al.*, 1993; Buikema & Haselkorn, 1991). To maintain high levels of NtcA and HetR during the first steps of differentiation, both genes are self-induced and also mutually induced (Frías *et al.*, 1993; Muro-Pastor *et al.*, 2002). Another regulator called NrrA has been identified whose involvement in cell differentiation is not entirely clear (Flores & Herrero, 2010), but it does appear to act as a regulatory link between NtcA and HetR (Ehira & Ohmori, 2006b). After 6-12 hours of nitrogen deprivation (intermediate differentiation phase), the induction of genes involved in the synthesis of the heterocyst envelope begins (Maldener *et al.*, 2014). Finally, genes encoding different oxidases and nitrogen assimilation proteins are induced in a late differentiation period (Herrero *et al.*, 2013; Flores *et al.*, 2019).

Together with the previously mentioned genes, during cell differentiation the expression of *patS* and *hetN* is also induced. These genes encode a peptide (PatS) or a protein (HetN) that, after being processed (at least in the case of PatS, see below), play a fundamental role in the establishment of the pattern of heterocysts along the filament, which in *Anabaena* consists of one heterocyst for 10-15 vegetative cells (reviewed in Herrero *et al.*, 2016). If the conditions of deprivation of combined nitrogen continue, the filament will begin to grow diazotrophically as a consequence of the division of the vegetative cells, making new heterocysts to appear in the center of the increasing interval, maintaining in this way the described pattern (Herrero *et al.*, 2016).

In *Anabaena*, the *patS* gene is constitutively expressed in the presence of combined nitrogen and induced in small cell clusters in the initial phases of cell differentiation in response to nitrogen deficiency. Mutants in which *patS* has been eliminated have a phenotype in which multiple contiguous heterocysts (Mch) are observed (Yoon & Golden 1998, 2001; Corrales-Guerrero *et al.*, 2013; 2015). The expression of *patS* produces a peptide of 17 amino acids that is processed to produce a peptide that corresponds to the C-terminal part of the original peptide, has the sequence ERGSGR, and behaves as a morphogen that is transferred to neighboring cells and inhibits their differentiation (Corrales-Guerrero *et al.*, 2013; Hu *et al.* 2015; Yoon & Golden 1998, 2001; Zhang *et al.*, 2017). HetN also has the sequence ERGSGR, but in this case internal to the protein; HetN or a derivative of it has a negative effect on heterocyst formation and, as it happens with PatS, its inactivation produces multiple contiguous heterocysts, although this phenotype is delayed compared to the one of inactivation of *patS* (Callahan & Buikema, 2001; Corrales-Guerrero *et al.*, 2013, 2014a; Higa *et al.*, 2012).

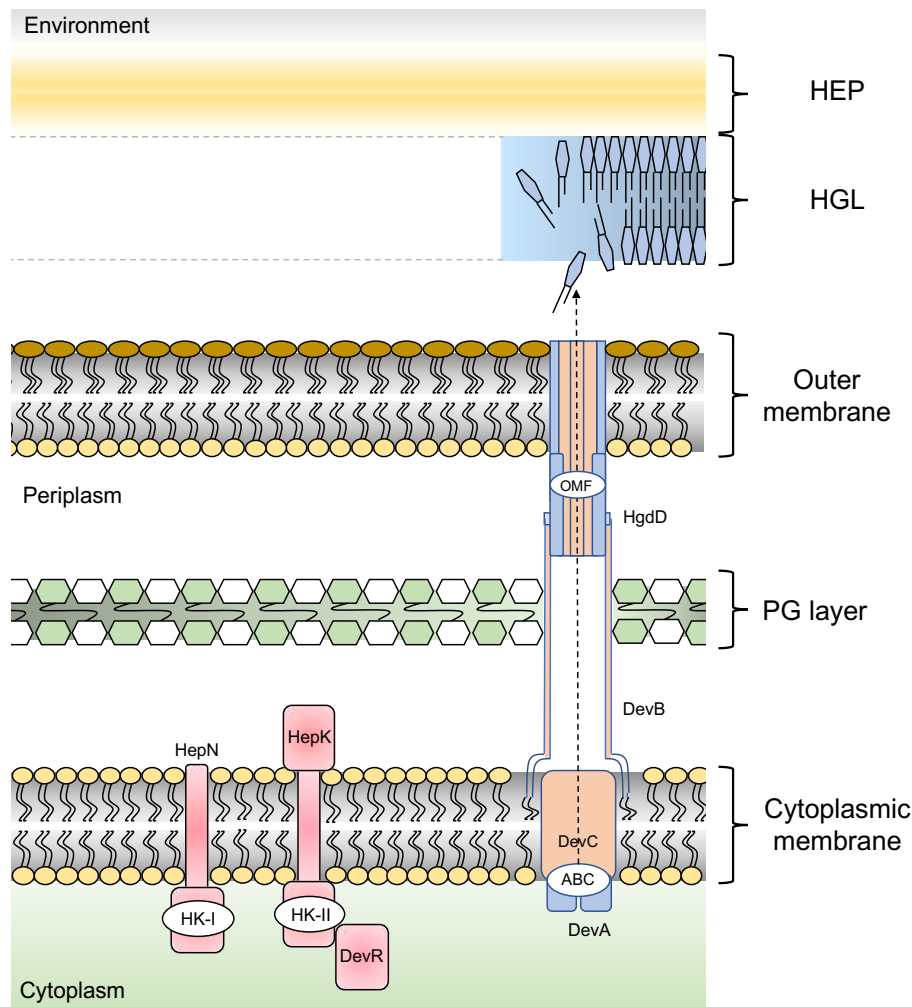


Finally, HetR, in addition to be regulated at the level of gene expression, can be regulated post-translationally by phosphorylation, preventing the accumulation of HetR tetramers, which are the active form of the protein (Valladares *et al.*, 2016). HetR binds, and is inhibited by, the PatS morphogen and HetN (or a HetN-derived factor), and has been suggested to be regulated by the HetF protease and by PatA, which has similarity with CheY (Risser & Callahan, 2008). Away from the heterocyst, the concentration of HetR increases making it possible to stimulate further heterocyst differentiation (Risser & Callahan, 2009).

### 1.3.3 Genes involved in the formation of heterocyst envelopes

The formation of the heterocyst envelope requires enzymes that carry out the synthesis and export of the materials that make up HEP and HGL, and is also regulated by various proteins. In *Anabaena*, HEP formation is one of the first morphological changes that occur after nitrogen deprivation, and many of the genes responsible for this process are found in a cluster called "HEP island" (*alr2825-alr2841*) (Huang *et al.*, 2005). Although some genes encoding proteins involved in this process have been described, such as *hepA* (similar to an ABC-type exporter of lipid A of lipopolysaccharide; Huang *et al.*, 2005), *hepB*, *alr3699* (both glycosyl transferases), *alr4388* (polysaccharide export protein) and *hepP* (predicted membrane protein), the detailed process of production and deposition of the HEP layer remains unknown (Maldener *et al.*, 2003; Wang *et al.*, 2007; Lechno-Yossef *et al.*, 2011; López-Igual *et al.*, 2012). As with *hep* genes, several genes encoding proteins for HGL layer formation are found in a cluster (*all5343-alr5347*) within the *Anabaena* genome (Fan *et al.*, 2005). The HGL layer is formed by fatty alcohols glycosidically linked to sugar residues, and within the HGL cluster there are several genes encoding fatty acid synthases and polyketide synthases, examples of which are *hglB*, *hglC*, *hglD* and *hglE* (Fan *et al.*, 2005; Maldener *et al.*, 2014). Some genes have also been identified that are necessary for the correct deposition of the HGL layer, some of which constitute the *devBCA* operon that encodes an ABC-type exporter (Fiedler *et al.*, 1998) and *hgdD* that encodes an outer membrane TolC-like protein (see Fig. 5) (Fan *et al.*, 2005; Moslavac *et al.*, 2007). Finally, an example of a protein that regulates the process of synthesis and deposition of the HGL layer is the DNA-binding protein DevH (Ramírez *et al.*, 2005), however, a very complex regulation appears to be operative that is not yet fully understood (reviewed in Flores *et al.*, 2019).





**Fig. 5. General structure of the heterocyst envelope in *Anabaena*.** The figure shows the cytoplasmic membrane formed by a bilayer of phospholipids. As an example of cell wall biosynthesis regulation, there are two histidine kinase complexes in the cytoplasmic membrane that are necessary for the activation of the synthesis of heterocyst envelope polysaccharide (HEP), although they do not intervene in the transport of polysaccharides. In the histidine kinase complex known as HK-I, HepN acts as a phosphorylation regulator, whereas HK-II has HepK as a phosphorylation regulator and phosphorylates DevR, which must be another intermediate regulatory factor since it does not have a transcription activation domain (for further details see Flores *et al.*, 2019). The cytoplasmic membrane also assembles part of the DevBCA-HgdD secretion complex that crosses the peptidoglycan layer (PG layer) and the outer membrane and is essential for the export of the glycolipids responsible for forming the heterocyst-specific glycolipid (HGL) layer, although the mechanism by which they are assembled is unknown. The DevBCA complex is an ABC-transporter while HgdD forms a tunnel that crosses the outer membrane. Between the cytoplasmic membrane and the outer membrane, there is an aqueous space called periplasm with the peptidoglycan layer (represented by hexagons, where one color represents N-acetyl glucosamine and the other represents N-acetyl muramic acid). The outer membrane is an asymmetric lipid bilayer composed of phospholipids in its inner layer and lipopolysaccharides in its outer layer. Outside the outer membrane is the HGL layer (composed of glycolipids) and the HEP layer (formed by polysaccharides). Both layers together protect the heterocyst from the entry of ambient oxygen. Although they have not been drawn to simplify the scheme, the cytoplasmic membrane and the outer membrane also have many other proteins including lipoproteins and, in the outer membrane, porins. The elements represented in the scheme are not to scale. The figure is adapted from Fig. 2 of Nicolaisen *et al.* (2009a).



#### 1.4 Septal junctions and septal proteins

As it has been commented above, filamentous cyanobacteria grow as chains of cells in which differentiated cells are observed if the environmental conditions are appropriate. This type of growth made microscopists focus their attention on the septal areas between two adjacent cells, being Wildon & Mercer (1963) among the first to show them. Thin structures perpendicular to the cytoplasmic membrane were observed that received the name "microplasmodesmata" (Lang & Fay, 1971). Later, by freeze-fracture electron microscopy, Giddings & Steahelin (1978) observed these microplasmodesmata as a pit and a protrusion on the exoplasmic and protoplasmic fracture faces, respectively, of each adjacent cell membrane. They also observed that the microplasmodesmata had an outer diameter of less than 20 nm and occupied a circular area in the center of the septum (Giddings & Steahelin, 1978). In the last decade, the septum between two adjacent vegetative cells of *Anabaena* has been studied by electron tomography (Wilk *et al.*, 2011). For this purpose, the samples were chemically fixed and stained, either with positive staining (potassium permanganate) showing protein structures communicating the two cells and having a diameter of 5.5 nm and a length of 18 nm, or with negative staining (osmium tetroxide) showing holes in the septal PG that had a diameter of 14 nm and a length of 18 nm. The structures connecting the two adjacent cells were called "septosomes" (Wilk *et al.*, 2011).

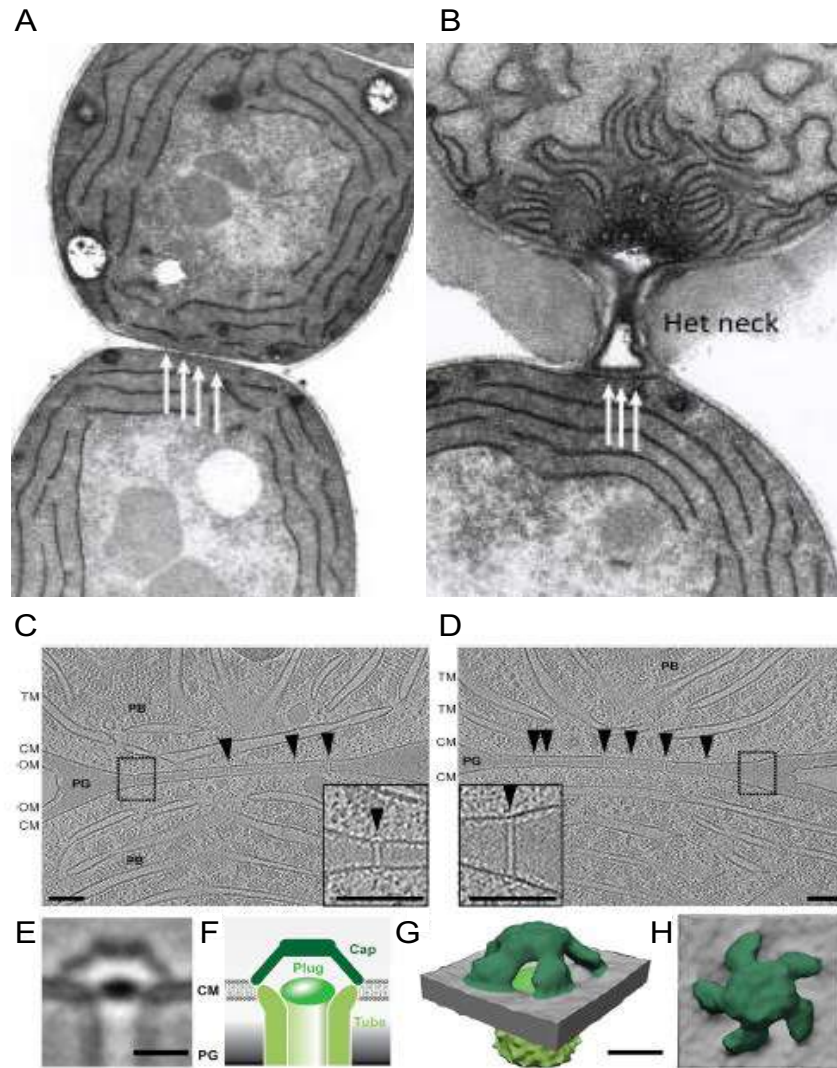
Studies on the septa between vegetative cells and on those between vegetative cells and heterocysts in filaments of *Anabaena* prepared by cryo-fixation, subjected to negative staining (osmium tetroxide) and visualized by electron tomography, showed the presence of structures, termed "channels", that cross the septal PG (Omairi-Nasser *et al.*, 2014). The size of the channels varies from 12 nm long and 14 nm in diameter in the septum between vegetative cells to 21 nm long and 14 nm in diameter between vegetative cells and heterocysts. This suggests a remodeling of the septa during the development of the heterocysts, and that the septal channels are larger between heterocysts and vegetative cells than between vegetative cells (Omairi-Nasser *et al.*, 2014, 2015). These septal structures, now termed "septal junctions" (Mariscal, 2014; Flores *et al.*, 2016), have been recently visualized by cryo-electron tomography (Weiss *et al.*, 2019). In the cytoplasmic membrane, in the part that is in contact with the cytoplasm, the septal junctions have a cap-like structure and just below it a plug-like density (see Fig. 6). These cytoplasmic membrane structures from adjacent cells are connected by a 11-nm wide tube (lumen, 7 nm).

The channels present in the septal disks (Wilk *et al.*, 2011; Omairi-Nasser *et al.*, 2015) would correspond to the perforations described as nanopores (Lehner *et al.*, 2013) and could be the places through which the septal junction traverse the septal PG (Flores *et al.*, 2016). The number and diameter of nanopores per septal disk is not the same for all strains of cyanobacteria, since for example in septa between vegetative cells the number of nanopores in *Nostoc punctiforme* is about 155 with a diameter of about 20 nm (Lehner *et al.*, 2013), while in *Anabaena* the number has been described to be about 75 with a diameter of about 15 nm (Nümborg *et al.*, 2015).

As mentioned earlier, in the septum between adjacent cells there are proteinaceous structures that appear to join the cells in the filament. Some proteins have been identified whose localization is septal and their absence results in filament fragmentation, which led to think that they could have a role in the septal junctions (Bauer *et al.*, 1995; Flores *et al.*, 2007; Nayar *et al.*, 2007; Merino-Puerto *et al.*, 2010).







**Fig. 6. Septal junctions in *Anabaena*.** (A) Electron micrograph of a portion of a filament of *Anabaena* where the white arrows indicate structures perpendicular to the cytoplasmic membrane located in the septum between two vegetative cells, known as septal junctions. (B) Electron micrograph of the septal zone between a vegetative cell (bottom) and a heterocyst (top) where the septal junctions are visible (white arrows). Het neck, the narrowed region located at the pole of the heterocyst. (C and D) Cryo-electron-tomograms of an *Anabaena* filament and magnified areas in each box (same size bars). The multiple septal junctions present in the septum between two vegetative cells are marked with arrowheads. CM, cytoplasmic membrane; OM, outer membrane; PG, septal peptidoglycan; TM, thylakoid membranes; PB, pycobilisomes. (E, F) Subtomogram of the terminal modules of the septal junctions and their schematic representation showing tube, cap and plug. (G) Oblique view of the septal junction. (H) View from above of the septal junction scheme. A and B from Flores *et al.* (2016); C-H from Weiss *et al.* (2019).



### 1.4.1 SepJ

ORF *alr2338* of the genome of *Anabaena* (Kaneko *et al.*, 2001) is the *sepJ* gene, whose inactivation or deletion results in lack of diazotrophic growth (Ernst *et al.*, 1992; Buikema & Haselkorn, 1991; Mariscal *et al.*, 2011). Heterocyst differentiation in these mutants aborts at an early stage, after production of HEP but before synthesis of HGL (Flores *et al.*, 2007; Mariscal *et al.*, 2011). A *sepJ* mutant grown in the presence of combined nitrogen makes short filaments (about seven cells per filament), and filament fragmentation increases after transfer to a medium without combined nitrogen (filament length means, about two cells per filament) (Flores *et al.*, 2007; Nayar *et al.*, 2007; Mariscal *et al.*, 2011). The expression of *sepJ* in *Anabaena* takes place regardless of the nitrogen source in the medium, although it is moderately increased in medium without combined nitrogen (Flores *et al.*, 2007), with this increase in expression concentrated in the heterocysts (Nayar *et al.*, 2007). A fusion of GFP (Green Fluorescent Protein) to the C-terminal part of SepJ shows localization in the center of the septa between vegetative cells (Flores *et al.*, 2007; Mariscal *et al.*, 2011), as well as two foci of GFP fluorescence in the septa between vegetative cells and a heterocysts (Flores *et al.*, 2007; Flores & Herrero, 2010).

The *Anabaena sepJ* gene encodes a 751-amino acid protein that is highly conserved in heterocystous cyanobacteria and in which four different conserved domains have been identified (Herrero *et al.*, 2016). (i) N-terminal region, composed of the first 27 amino acids whose sequence is strongly conserved. (ii) Coiled-coil domain, ranging from amino acid 28 to 207, in which two coiled-coil motifs are strongly predicted. (iii) Linker domain, from residue 208 to 411, which is rich in proline (15-25% in the protein from different cyanobacteria), serine (10-20%), threonine (6-15%) and in some cases glutamic acid (about 10%). Unlike the previous ones, this domain presents a great variety with respect to its length (ranging from 67 to 317 amino acids) and amino acid sequence, but its composition of amino acids is similar in different cyanobacteria. (iv) Integral membrane domain (permease domain) composed by the amino acids ranging from 412 to 751. Analyzing its sequence, two subdomains can be distinguished, IM1 and IM2 (IM, from Integral Membrane). IM1 has two or three transmembrane segments (TMSs) that are quite variable; on the other hand, IM2 has eight TMSs and is topologically conserved in different cyanobacteria (Ramos-León *et al.*, 2018).

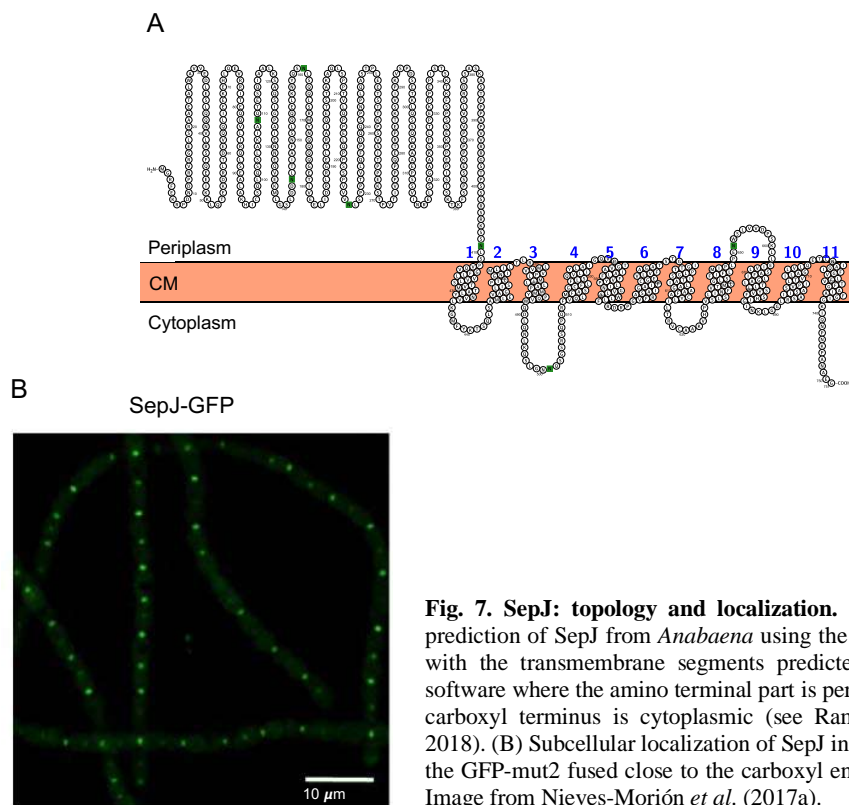
SepJ is limited to the cyanobacterial clade and, as just discussed, its general structure is highly conserved in cyanobacteria that form heterocysts. Such is the degree of conservation, that antibodies specifically generated against the coiled-coil domain of *Anabaena* are functional against *Mastigocladus laminosus* SepJ (Nürnberg *et al.*, 2014), which is framed in section V and is phylogenetically distant from *Anabaena* (Shih *et al.*, 2013). However, in non-heterocyst forming filamentous cyanobacteria, SepJ can be found but in most cases with a short linker domain, as for example occurs in *Arthrospira platensis* NIES-39 whose linker domain has only 40 amino acids but maintains abundance in proline and serine (Herrero *et al.*, 2016). Finally, homologues to the permease domain of SepJ can be found in unicellular cyanobacteria (Escudero *et al.*, 2015).

The number of TMSs of SepJ is still not clear, since SepJ from different heterocystous cyanobacteria is predicted to contain 9 to 11 TMS by different topology prediction programs (prediction for *Anabaena* SepJ in Fig. 7a) (Rudolf *et al.*, 2015; Ramos-León *et al.*, 2018). Based on this information and on evidence that the C-terminus is cytoplasmic (Flores *et al.*, 2007;



Ramos-León *et al.*, 2015), if SepJ has an odd number of TMSs the coiled-coil and linker domains would be periplasmic, while if the number of TMSs is even, these domains would be cytoplasmic (Flores *et al.*, 2016). Some evidence supports that the coiled-coil and linker domains are periplasmic: (i) protein-protein interaction has been observed between the linker-domain of SepJ and the periplasmic domain of FstQ, which is a divisome protein (Ramos-León *et al.*, 2015); (ii), an interaction between SepJ and a periplasmic domain (SH3) of the peptidoglycan-binding protein SjcF1 has been demonstrated (Rudolf *et al.*, 2015); (iii) immunogold TEM analysis performed on the coiled-coil domain of an *Anabaena* SepJ overexpression strain showed localization in the septa between vegetative cells (Omairi-Nasser *et al.*, 2015).

Coiled-coil domains are frequently involved in protein-protein interactions, which together with the observed formation of SepJ multimers and the fact that SepJ self-interacts in bacterial adenylate cyclase two-hybrid (BACTH) analysis (Ramos-León *et al.*, 2015; 2017), suggests that SepJ forms large protein complexes. Additionally, the polar localization of SepJ (Fig. 7b) (Flores *et al.*, 2007) makes us wonder whether the coiled-coil domains of SepJ from adjacent cells could interact mediating the formation of cell-cell joining complexes (Flores *et al.*, 2007). The decrease in the number of nanopores observed in the murein sacculi from a  $\Delta sepJ$  mutant is consistent with a role of SepJ in the septal junctions (Nürnberg *et al.*, 2015).



**Fig. 7. SepJ: topology and localization.** (A) Topological prediction of SepJ from *Anabaena* using the Protter software with the transmembrane segments predicted by TMHMM software where the amino terminal part is periplasmic and the carboxyl terminus is cytoplasmic (see Ramos-León *et al.*, 2018). (B) Subcellular localization of SepJ in *Anabaena* using the GFP-mut2 fused close to the carboxyl end of the protein. Image from Nieves-Mori3n *et al.* (2017a).

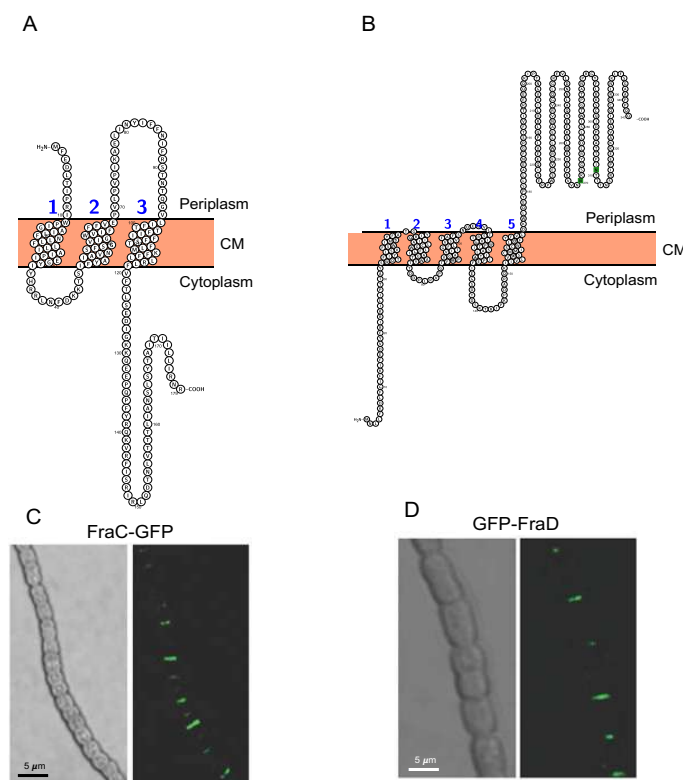


### 1.4.2 Fra proteins

The set of Fra proteins received this name because the inactivation of some of them leads to the fragmentation of the *Anabaena* filament under certain conditions. These proteins are encoded in a gene cluster comprising *fraC*, *fraD* and *fraE*, which are expressed as an operon (Merino-Puerto *et al.*, 2010). Another gene in the cluster is *fraF*, which is found downstream of the operon and is transcribed in the opposite orientation (*fraC-fraD-fraE*  $\leftarrow$  *fraF*) (Merino-Puerto *et al.*, 2013). The deletion of any gene in the operon generates a phenotype of inability to grow diazotrophically and filament fragmentation mainly under nitrogen deprivation, whereas inactivation of *fraF* results in impaired diazotrophic growth only in solid medium and an increased frequency of long filaments. The expression of the *fra* operon is constitutive at low levels, whereas on the other hand the expression of *fraF* is increased under conditions of nitrogen deprivation producing an antisense RNA that overlaps *fraE* and, partially, *fraD* (Merino-Puerto *et al.*, 2013; Ehira & Ohmori, 2014). Finally, it should be mentioned that all filamentous cyanobacteria have retained a *fra* operon, except for a group in section III that includes *Pseudanabaena*.

Bioinformatics analysis predicts that FraC, FraD and FraE are membrane integral proteins, and fusions with the GFP-mut2 protein have shown that FraC-GFP and GFP-FraD are localized at the cell poles in the intercellular septa, although less focused than SepJ (Merino-Puerto *et al.*, 2010). In *Anabaena*, FraC is a 179-amino acid protein with three (or less likely four) predicted TMSs and the C-terminus in the cytoplasm (Fig. 8a) (Herrero *et al.*, 2016; Rudolf *et al.*, 2015). This protein could be important in the correct location and assembly of the septal junctions (Merino-Puerto *et al.*, 2011b; Weiss *et al.*, 2019). FraD is a 343-amino acid protein with five TMSs and a periplasmic coiled-coil domain that could be involved in protein-protein interactions (Fig. 8b) (Herrero *et al.*, 2016). FraD has been shown by immunogold TEM to be localized at the intercellular septa (Merino-Puerto *et al.*, 2011b) and is a component of the septal junctions that have been visualized by cryoET (Weiss *et al.*, 2019). FraE is a 267-amino acid protein in which six TMSs were predicted, and it shows homology to the permease protein of some ABC transporters, specifically with one of *Myxococcus xanthus* which is essential for the assembly of the pilus and export of the pilus subunit (Herrero *et al.*, 2016).





**Fig. 8 Topology and localization of FraC and FraD.** (A and B) Topology of FraC and FraD, respectively, using the Protter software and the transmembrane segments predicted in Rudolf *et al.* (2015). FraC is presented with 3 TMSs where the amino terminus is periplasmic and the carboxy terminus is cytoplasmic, while FraD has 5 TMSs, the amino terminus cytoplasmic and the carboxy terminus periplasmic. (C) The image shows the subcellular localization of FraC in *Anabaena* by attaching the GFP-mut2 at the carboxy terminus of FraC. (D) Subcellular location of FraD in *Anabaena* by fusion of the GFP-mut2 at the amino terminus of the protein. C and D from Herrero *et al.* (2016). The picture on the left is a bright-field image and the picture on the right the GFP fluorescence.

From the physiological point of view, the single  $\Delta fraC$  and  $\Delta fraD$  mutants show nitrogenase levels lower than the wild type, but the single  $\Delta fraE$  mutant shows negligible nitrogenase values; this is consistent with the inability of all of them to grow diazotrophically (Merino-Puerto *et al.*, 2010). From the structural point of view, it is remarkable that the  $\Delta fraE$  mutant produces heterocysts that lack the heterocyst neck, whereas the  $\Delta fraC$  and  $\Delta fraD$  mutants produce heterocysts that show a heterocyst neck but have the septal cytoplasmic membrane retracted into the heterocyst (Merino-Puerto *et al.*, 2011b). The observation that the single  $\Delta fraC$  and  $\Delta fraD$  mutants and the double  $\Delta fraC \Delta fraD$  mutant have similar phenotypes (Merino-Puerto *et al.*, 2010, 2011b) and the conservation of the *fra* operon in filamentous cyanobacteria suggest that FraC and FraD indeed work together. Consistently, as observed by cryoET, FraD has been described to be a component of the septal junctions and FraC to be involved in their assembly (Weiss *et al.*, 2019).





FraF belongs to the pentapeptide-repeat protein family (which will be described later). An *Anabaena* strain carrying GFP attached to the C-terminus of FraF showed that FraF-GFP is localized in the cytoplasm (Merino-Puerto *et al.*, 2013). FraF is known to influence the expression of the *fra* operon, but the mechanism by which FraF restricts filament length is not yet known (Herrero *et al.*, 2016).

## 1.5 Intercellular communication

As mentioned above, in filamentous heterocystous cyanobacteria, communication occurs between adjacent cells for the transport of both nutrients and regulators. In these organisms, atmospheric nitrogen and carbon fixation takes place in different cells, making the exchange of nutrients between cells essential. Glutamine and glutamate (Wolk *et al.*, 1976; Thomas *et al.*, 1977), products of cyanophycin catabolism (Gupta & Carr, 1981; Picossi *et al.*, 2004; Burnat *et al.*, 2014), sucrose (López-Igual *et al.*, 2010; Vargas *et al.*, 2011) and alanine (Jüttner, 1983; Pernil *et al.*, 2010) have been described as intercellularly transferred compounds. On the other hand, the intercellular transfer of regulators such as the PatS morphogen and a HetN-dependent regulator are essential for the correct formation of the pattern of heterocysts (Corrales-Guerrero *et al.*, 2013; Rivers *et al.*, 2014).

Because of the importance of communication between vegetative cells or between vegetative cells and heterocysts, many studies have been carried out to shed light on the structures and mechanisms involved in these intercellular communication processes. A continuous periplasm and cell-cell joining structures have been proposed as two possible routes of communication (Mariscal *et al.*, 2007; Mullineaux *et al.*, 2008).

### 1.5.1 Continuous periplasm

The continuous periplasm, which is consequence of the fact that the outer membrane is not subjected to a process of constriction after cell division, has been suggested as a route of nutrient exchange in *Anabaena*. A study, in which GFP (27 kDa) was specifically expressed in heterocysts and exported to the periplasm, showed that it diffused to the periplasm of neighboring vegetative cells (Mariscal *et al.*, 2007). This observation suggested the possible movement of molecules along the periplasm, although diffusion of large proteins could be prevented by the peptidoglycan (Mariscal *et al.*, 2007; Zhang *et al.*, 2008, 2013).

For this communication mechanism, cells should have transporters in their cytoplasmic membranes to export substrates to the periplasm in one type of cells and to import the transferred substrates in the other type of cells. In addition, the outer membrane should have a low permeability to these substrates, otherwise they would leak out to the environment. In *Anabaena*, some ABC transporters that are essential for diazotrophic growth are involved in the uptake of amino acids and glucosides (Picossi *et al.*, 2005; Pernil *et al.*, 2008; Nieves-Mori3n *et al.*, 2017a). Additionally, it has been shown that the outer membrane of *Anabaena* has low permeability for intercellularly exchanged metabolites that are necessary for diazotrophic growth such as glutamate, glutamine or sucrose (Nicolaisen *et al.*; 2009b). These observations support a possible role of the periplasm as a communication conduit. In contrast, no amino acid or sugar exporters have been characterized.



### 1.5.2 Cell-cell communication by septal junctions

The transfer of molecules through the septal joining cell-cell structures is another mechanism that has been proposed for intercellular communication. To test this hypothesis, small fluorescent molecules have been used to observe and quantify intercellular exchange in real time (Mullineaux *et al.*, 2008). Three fluorescent markers have been used: (i) calcein (622.5 Da), which has four ionizable groups resulting in a mixed molecular population with mainly three negative charges at pH 7; (ii) 5-carboxyfluorescein (5-CF; 376.3 Da), which is smaller than calcein and has mainly one negative charge at pH 7; and (iii) esculin (340.3 Da) that is a sucrose analog mainly neutral at pH 7 (Nürnberg *et al.*, 2015). Both calcein and 5-CF pass into the cell because they are initially hydrophobic, but once in the cytoplasm they are hydrolyzed and become highly hydrophilic and are retained inside the cell (Mullineaux *et al.*, 2008, Mariscal *et al.*, 2011). Unlike calcein or 5-CF, esculin is taken up from the medium by glucoside transporters (Nürnberg *et al.*, 2015; Nieves-Mori3n *et al.*, 2017a). The studies carried out with these fluorescent molecules are based on the fact that they are visible when excited with a monochromatic light, but if a high light density is applied they are destroyed and the fluorescence is lost. This allows to bleach a cell and to follow the time it takes to recover fluorescence, a process known as FRAP (Fluorescence Recovery After Photobleaching) that permits to calculate a recovery rate constant (Merino-Puerto *et al.*, 2011b, Nieves-Mori3n *et al.*, 2017a) or, in a more complex way, to determine the molecular exchange coefficient (Mullineaux *et al.*, 2008). The fluorescence that reaches the bleached cell comes from the adjacent cells showing movement of the marker along the filament. It has been shown that this movement takes place by simple diffusion, since the movement of the markers is always down the concentration gradient (from higher to lower concentration) (Mullineaux *et al.*, 2008) and is directly proportional to the absolute temperature (Nieves-Mori3n *et al.*, 2017b). Because diffusion can hardly take place through active transporters as those discussed to be involved in movement through the periplasm, structures directly joining the cytoplasm of adjacent cells have been considered to mediate diffusion (Mullineaux *et al.*, 2008).

### 1.5.3 Possible different types of septal junctions

Results from intercellular transfer experiments with fluorescent markers have shown that deletion mutants of the genes encoding septal proteins SepJ, FraC and FraD are impaired in rapid intercellular molecular exchange (Mullineaux *et al.*, 2008; Merino-Puerto *et al.*, 2010, 2011; Nürnberg *et al.*, 2015). However, whereas  $\Delta sepJ$  mutants were more altered in calcein than in 5-CF transfer, in  $\Delta fraC$  and  $\Delta fraD$  mutants (as well as in the double  $\Delta fraC \Delta fraD$  mutant) calcein and 5-CF transfer were similarly altered (Nürnberg *et al.*, 2015, Merino-Puerto *et al.*, 2011b). On the other hand, the  $\Delta sepJ$  and double  $\Delta fraC \Delta fraD$  mutants showed a decrease in the number of nanopores compared to the wild type, indicating the involvement of these proteins in the formation of nanopores (Nürnberg *et al.*, 2015). Based on these observations, it has been suggested that there can be at least two different types of septal junction (Merino-Puerto *et al.*, 2011b; Nürnberg *et al.*, 2015), one related to SepJ that could allow the exchange of larger molecules (its mutation affected mainly calcein transfer) and another related to FraC/FraD that could be involved in the intercellular transfer of sucrose (its mutation affected mainly 5-CF and esculin transfer). Nonetheless, the presence of other types of structures in the intercellular septa that could carry out intercellular communication has not been ruled out, since a  $\Delta sepJ$ ,  $\Delta fraC$  and  $\Delta fraD$  triple mutant still presents over 7% of nanopores with respect to the wild type and still shows some exchange of the fluorescent markers (Nürnberg *et al.*, 2015). The recent cryoET



study of *Anabaena* has only identified, however, one type of septal junctions, which include FraD and need FraC for full assembly (Weiss *et al.*, 2019).

#### 1.5.4 Possible regulation of the septal junctions

A study of esculin transfer in diazotrophically-grown *Anabaena* filaments showed that, in contrast to heterocysts that had the same level of labeling as their neighboring vegetative cells, some heterocysts showed a lower level of labeling, defining the existence of non-communicating heterocysts (Nürnberg *et al.*, 2015). These results suggested that the septal junctions in a heterocyst could be regulated in some way, so that they were open or closed allowing or not, respectively, the passage of esculin (Flores *et al.*, 2018). The same conclusion, that the septal junctions can be regulated in a cell, was reached after re-evaluation of data in which the fluorescent marker was calcein or 5-CF (Flores *et al.*, 2018). These data showed that the recovery rate constants in the wild type conformed to a normal distribution of the data, but when mutants of the *glsC* and *glsD* genes (encoding components of an ABC transporter for glucosides) were analyzed, there was a clear increase in the number of cells with a low recovery rate constant ( $< 0.01 \text{ s}^{-1}$ ), defining them as non-communicating cells (Flores *et al.*, 2018). It was hypothesized that GlsC and GlsD could intervene in the opening or closing of the septal junctions, perhaps transmitting information on the availability of sucrose to regulate the septal junctions (Flores *et al.*, 2018).

Finally, in a mutant of gene *hepP* (which encodes an MFS glucoside transporter), it was observed that the transfer of esculin between vegetative cells is decreased, but transfer from vegetative cells to heterocysts is increased (Nieves-Mori3n *et al.*, 2017a). The hypothesis has been considered that HepP could inhibit specifically the heterocyst septal junctions under conditions of sucrose scarcity (Flores *et al.*, 2018). In addition, FRAP experiments performed under different stresses showed the presence of a high number of non-communicating cells, but if the filaments were transferred and incubated without stress conditions the number of non-communicating cells decreased (Weiss *et al.*, 2019). These data support the hypothesis that the septal junctions in a cell can be opened or closed to allow or not, respectively, the passage of molecules.

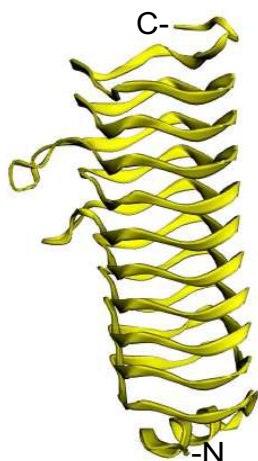
#### 1.6 Pentapeptide repeat proteins

Proteins of the pentapeptide repeat protein (PRP) family are defined as proteins that have at least eight tandem repeats of a five-amino acid sequence, which could be framed within the consensus (STAV)(D/N)(L/F)(S/T/R)(X) (Bateman *et al.*, 1998; Vetting *et al.*, 2006; Ni *et al.*, 2009). More than 3,500 PRPs have been identified through computer platforms (Pfam database) (Bateman *et al.*, 2000; El-Gebali *et al.*, 2019). Approximately 10% of PRPs are found in eukaryotes and 90% in prokaryotes, both archaea and bacteria, and they are also present in some bacteriophages (Vetting *et al.*, 2006; Zhang *et al.*, 2019). Specifically, in *Anabaena*, 32 PRPs have been identified (Ni *et al.*, 2009). The PRPs present a structure that can be described as a collection of type II and IV beta turns, because the conformation that they usually adopt is that of four-sided, right-handed beta helical structure (Buchko *et al.*, 2006, 2008; Vetting *et al.*, 2006, 2007). However, beyond the fact that their general structure is known, their biochemical function is still unknown in most cases (Ni *et al.*, 2009). Recently, a new PRP type with a type I beta turn structure has been discovered in *Anabaena* (termed *Nostoc* sp. PCC 7120 by the authors) (Zhang *et al.*, 2019), making it clear that other types of structures could appear as the study of PRPs deepens.





In cyanobacteria, several PRPs have been described, but only in a few cases a function could be envisaged. For example, RfrA isolated from *Synechocystis* sp. PCC 6803 is involved in the absorption of manganese (Chandler *et al.*, 2003), although its structure and precise biochemical function are unknown (Ni *et al.*, 2009). HetL described in *Anabaena* is fundamental for the establishment of the pattern of heterocysts along the filament, its structure has been determined (Fig. 9) but its mechanism of action is unknown (Ni *et al.*, 2009). Other PRPs have also been described in *Anabaena*, such as PatL, which is also related to the establishment of the heterocyst pattern (Liu & Wolk, 2011), or HglK, whose inactivation prevents diazotrophic growth apparently affecting Hgl deposition (see below; Black *et al.*, 1995). Apart from cyanobacteria, other PRPs have also been described such as MfpA from *Mycobacterium tuberculosis* that has been determined to perform a DNA gyrase inhibitor function, likely because the DNA gyrase could bind to the DNA-mimicking PRP (Hegde *et al.*, 2005; Ni *et al.*, 2009). Finally, to give an example in eukaryotes, a PRP has been described that is related to the tetramerization of potassium channels in the membrane (Vetting *et al.*, 2006).



**Fig 9. Diagram of HetL.** The image shows the  $\alpha$ -helix at the N-terminal part and the  $\beta$ -sheets that constitute most of HetL (Ni *et al.*, 2009). Structure PDB 3DU1.

Although the pentapeptide repeat sequence described above is the major pentapeptide found in PRPs, a second pentapeptide repeat has been described in the Pfam database as "peptapetide\_2 repeat", which responds to the consensus sequence (N)(T,L,I)(G)(S,N)(G), with 61 examples in prokaryotes and two examples in eukaryotes (Vetting *et al.*, 2006). To date, not much is known about the structure of the pentapeptide\_2 repeat, but it is thought that its structure should not be the same as that of the original pentapeptide (Vetting *et al.*, 2006). Both types of PRPs are associated with other types of domains, including, for example, transmembrane segments, WD40 repeats, Ser/Thr protein kinase and PPE (Pro-Pro-Glu) domains (Vetting *et al.*, 2006). However, there is no protein where both pentapeptide repeat types occur at the same time, nor is there a domain that has been associated with one type of PRP and is present in the other type of PRP. Therefore, a certain type of domain appears to be associated only with one PRP or the other (Vetting *et al.*, 2006).



### 1.6.1 The HglK protein

More than half of the sequences found in the Pfam database belonging to the PRP family are within the phylum of cyanobacteria (Zhang *et al.*, 2019), so it is not surprising that it was in a cyanobacterium that this type of proteins was first discovered and characterized. The first PRP described was HglK (Black *et al.*, 1995), ORF *all0813* from the *Anabaena* genome. HglK is composed of several parts since it has an N-terminal domain with four TMSs and a C-terminal domain in which the pantapeptide repeat with 36 repetitions of consensus sequence ADLSG is found (Black *et al.*, 1995) (Fig. 10).

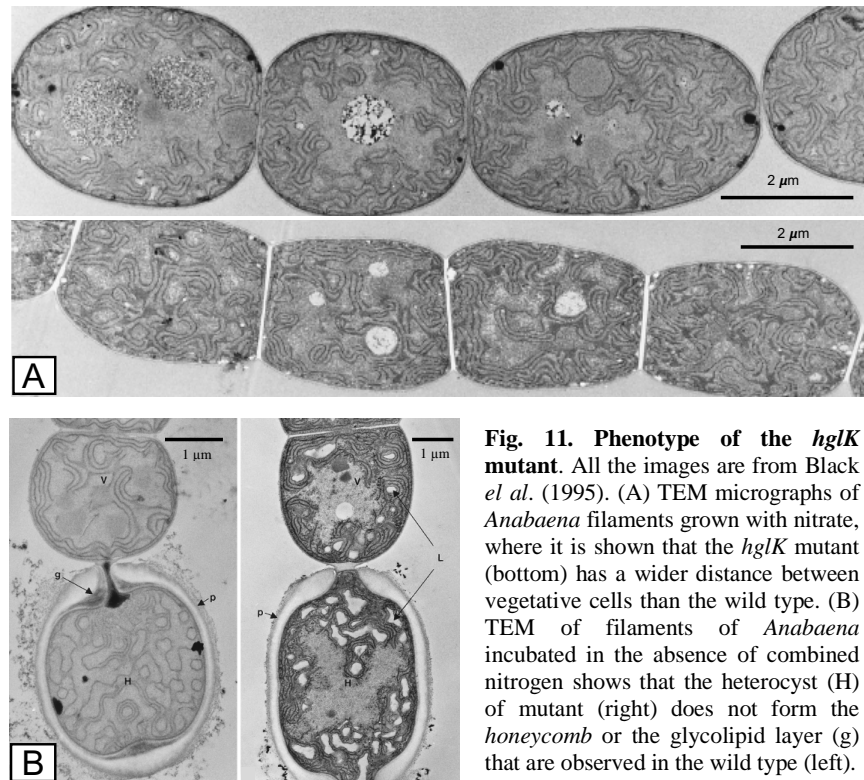
```

MTTPIVKKSNNAPSKNLVKPNSLPLATRRLAAWATEITLLATTGLVFPGLGVYINSRSDIN
REPLNPALVVTERAIARPLALPADYGIRNVAWPTNYLWMLALLAPTALSWWQLYLLAKTGS
TLPKRWFQVGLVNEEGTPPLATVVVREGIGRWTVPMSVAYILWRYSFAFPNLGLFTSLAV
LMVIGEALALPARRGRKALHDWLAGTYVVDANRPVPSPDVALNGRGLSGVSPQPEEGNATL
ATTAMAMSYPPQGEVITTDNSLSLSLWRRMQQNPSLTLFGVALTSMTAVLATLIGTQVYIQT
QQGNRESQKINSQQFLEFVKQLSPEGASIEDRQRTILALGSLKDFQSIQFLTDMMVKETN
PILIDTIQQALTSVGTAAIPELQNKQFLATELDSVGSASPEREARQKRLQINQQTINKIL
NVYSGKTLGLDLRSRTQLGQSGTVGGSFFNLILDNIDLSGIKFKSANLNQASFKGSRFRSVG
DDGRWDTYDDAIADLSQAQMKQANFTDANLSRVLMTRSDLSRATLNRANLSNARLIGANLS
SAQLVGADLRGTVLENASLTGADLGDALQEANLYGARLSRVIAIGAQLSFANLTKTDWQS
SDLSGADLERANLSNADLSATRMGTAILRSAQLENANLRNADLSLVDLRGANVAGADFKDT
ILTPSRQDPADQFVQTPELGSVSAVVKGVDVFSQAKNLDGKQLAYICTQGGVHPRCP
    
```

**Fig. 10. HglK sequence.** Amino acid sequence with indication of transmembrane segments (magenta) and pentapeptide repeats (alternatively shown in blue and red color), as described by Black *et al.* (1995).

The *hglK* mutant was unable of diazotrophic growth. This mutant also presented some structural alterations in the heterocysts as observed by TEM. Among these alterations, the following stand out: the lack of the HGL layer, the lack of reorganization of the intracellular membranes and the lack of a polar cyanophycin granule (Fig. 11b) (Black *et al.*, 1995). The lack of the glycolipid layer is not due to the inability of the heterocysts to synthesize Hgl, since the images show lipid vesicles inside the cells; these observations suggested that HglK could be involved in glycolipid transport (Black *et al.*, 1995). Additionally, the microscopy images also showed that the intercellular septum between vegetative cells (Fig. 11a) or between vegetative cells and heterocysts (Fig. 11b) was much wider than in the wild type (Black *et al.*, 1995).





**Fig. 11. Phenotype of the *hglK* mutant.** All the images are from Black *et al.* (1995). (A) TEM micrographs of *Anabaena* filaments grown with nitrate, where it is shown that the *hglK* mutant (bottom) has a wider distance between vegetative cells than the wild type. (B) TEM of filaments of *Anabaena* incubated in the absence of combined nitrogen shows that the heterocyst (H) of mutant (right) does not form the *honeycomb* or the glycolipid layer (g) that are observed in the wild type (left).



## 1.7 Objectives

In this work, we addressed some aspects of the intercellular septa in filamentous cyanobacteria, mainly the *Anabaena* sp. PCC 7120 model strain, by carrying out further studies of mutants of the septal proteins SepJ, FraC and FraD and a characterization of the cytoplasmic-membrane, pentapeptide-repeat protein HglK, previously thought to be involved in the export of glycolipids. The main objectives of this Thesis were:

1. To study the possible relationship between intercellular molecular transfer through septal junctions and nanopores under different culture conditions.
2. To characterize the HglK protein from *Anabaena* by the study of the phenotype of *hglK* mutants and the study of HglK protein fusions.
3. To search for possible partners of HglK and SepJ in protein complexes.
4. To study the SepJ and Fra proteins from the marine symbiotic cyanobacteria *Richelia intracellularis* and *Calothrix rhizosoleniae*.
5. To study the subcellular localization of the FraE protein in *Anabaena*.

Part of the results of this thesis have been included in the following articles:

- Arévalo S & Flores E (2020) Pentapeptide-repeat, cytoplasmic-membrane protein HglK influences the septal junctions in the heterocystous cyanobacterium *Anabaena*. *Mol Microbiol* 113: 794-806.
- Ramos-León F, Arévalo S, Mariscal V & Flores E (2018) Specific mutations in the permease domain of septal protein SepJ differentially affect functions related to multicellularity in the filamentous cyanobacterium *Anabaena*. *Microb Cell* 5: 555-65.
- Springstein BL, Arévalo S, Helbig AO, Herrero A, Stucken K, Flores E & Dagan T (2020) A novel septal protein of multicellular heterocystous cyanobacterial is associated with the divisome. *Mol Microbiol* (doi: 10.1111/mmi/14483).



## 2. Materials & Methods

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

## 2.1 Organisms and Growth conditions

The living organisms used in this work were two types of gram-negative bacteria, *Escherichia coli* and *Anabaena* sp., whose growth conditions are detailed below.

### 2.1.1 *Escherichia coli*

#### 2.1.1.1 Strains

The *Escherichia coli* strains used in this work, together with their genotypic characteristics, are described in Table 1.

**Table 1. *E. coli* strains used in this work together with their genotypes and references.** The DH5 $\alpha$  strains were obtained from the Servicio de Cultivos Biológicos del Centro de Investigaciones Científicas Isla de la Cartuja, whereas the ED8654 and HB101 strains are stored in the strains bank of our research group.

Strain	Genotype	Uses	Reference
DH5 $\alpha$	F- supE44 hsdR17(r <sub>k</sub> -m <sub>k</sub> +) recA1 girA96 (NaI) endA1 thi-1 relA1 $\Delta$ (lacaya-argF) ( $\emptyset$ 80lacZAM15) U59	Gene cloning and plasmid construction	Hanahan, 1983
ED8654	Lac-3 o LacY1 supE44 supF58 hsdR514 (r <sub>k</sub> -m <sub>k</sub> -) recA56mcrA1 metB1 lacY galk2 galT22 trpR55	<i>Anabaena</i> conjugation	Murray <i>et al.</i> , 1977
HB101	F- hdsS20 (r <sub>B</sub> -m <sub>B</sub> -) leu supE44 ara14 galK2 lacY1 proA2 rpsL20 xyl-5 mtl-1 recA13 mcrB	<i>Anabaena</i> conjugation	Boyer & Roulland- Dussoix, 1969

#### 2.1.1.2 Culture media and conditions

*E. coli* strains were grown axenically in rich Luria-Bertani (LB) medium consisting of 10 g/L NaCl, 10 g/L tryptone and 5 g/L yeast extract (Sambrook & Russell, 2001). For preparing a solid medium, it was supplemented with 1.5% (w/v) agar (Panreac) before autoclaving. Both types of medium (solid or liquid) were sterilized in autoclave.

The liquid media were incubated in orbital shakers at 200 rpm and 37°C, while the solid media were grown in Petri dishes in a 37°C incubator. If necessary, the media were supplemented with antibiotics that were sterilized by filtration and used at a final concentration of: ampicillin (Ap) 50  $\mu$ g/ml; chloramphenicol (Cm) 30  $\mu$ g/ml; streptomycin sulfate (Sm) 25  $\mu$ g/ml; spectinomycin dihydrochloride pentahydrate (Sp) 100  $\mu$ g/ml; and kanamycin sulfate (Km) 50  $\mu$ g/ml. The antibiotic stocks were prepared in ultrapure water (milliQ) with the exception of chloramphenicol, which was dissolved in absolute ethanol.

#### 2.1.1.3 Harvest methods

For volumes of less than 3 mL, cells were collected by centrifugation using 1.5-mL eppendorf tubes in an Eppendorf microcentrifuge model miniSpind plus at 13,000 rpm for 1 minute. For volumes between 5 and 50 mL, a Haraeus Megafuge 1.0 centrifuge was used at 4000 rpm for 5 minutes.



## 2.1.2 Cyanobacteria

### 2.1.2.1 Strain

The prokaryotic organism on which this work has been focused is the filamentous heterocyst-forming cyanobacterium *Anabaena* sp. (also known as *Nostoc* sp.) strain PCC 7120 (ATCC 27893), which is in section IV of the taxonomic classification of Rippka *et al.* (1979). All *Anabaena* strains used during the development of this Thesis are described in the following tables.

**Table 2. *Anabaena* strains previously available.**

Strain	Genotype	Resistance	Reference
PCC 7120	Wild type	None	Rippka <i>et al.</i> , 1979
CSAM137	<i>sepJ-gfp<sub>mut2</sub>::pCSV3</i>	Sm Sp	Flores <i>et al.</i> , 2007
CSVM34	$\Delta sepJ$	None	Mariscal <i>et al.</i> , 2011
CSVM141	$\Delta sepJ \Delta fraC \Delta fraD$	None	Nürnberg <i>et al.</i> , 2015
CSVT22	$\Delta fraC \Delta fraD$	None	Merino-Puerto <i>et al.</i> , 2011b

**Table 3. Cyanobacterial strains constructed in this work.**

Strain	Genotype	Description	Resistance
CSSA1	$\Delta hlgK::C.K1$	Strain obtained by double homologous recombination (DHR) in which <i>hlgK</i> is inactivated by deletion of an internal fragment and insertion of the C.K1 cassette	Nm
CSSA2	$\Delta fraC \Delta fraD \Delta hlgK::C.K1$	Strain obtained by DHR in which <i>hlgK</i> is inactivated by deletion/insertion of C.K1 in strain CSVT22 ( $\Delta fraC \Delta fraD$ )	Nm
CSSA3	$\Delta sepJ \Delta hlgK::C.K1$	Strain obtained by DHR in which <i>hlgK</i> is inactivated by deletion/insertion of C.K1 in strain CSVM34 ( $\Delta sepJ$ )	Nm
CSSA4	<i>hlgK::sf-gfp</i> (+pCSV3)	Strain obtained by single homologous recombination (SHR) in which <i>hlgK</i> lacks the stop codon and bears <i>sf-gfp</i> fused at the 3' end (pCSV3 vector also integrated)	Sm Sp
CSSA7	<i>hlgK::6xHis</i> (+pCSV3)	Strain obtained by SHR in which <i>hlgK</i> lacks the stop codon and bears a 3' sequence encoding a 6xHis tag (+ pCSV3)	Sm Sp





<b>CSSA12</b>	$\Delta sepJ \Delta fraC$ $\Delta fraD \Delta hglK::C.K1$	Strain obtained by DHR in which <i>hglK</i> is inactivated by deletion/insertion of C.K1 strain CSVM141 ( $\Delta sepJ \Delta fraC \Delta fraD$ )	Nm
<b>CSSA14</b>	$\Delta sepJ::sepJ_{RintHH01}$ (+pCSV3)	Strain obtained by SHR to complement <i>Anabaena</i> $\Delta sepJ$ with <i>sepJ</i> from <i>Richelia intracellularis</i> HH01 (+ pCSV3)	Sm Sp
<b>CSSA15</b>	$\Delta fraC-\Delta fraD::fraC-fraD_{RintHH01}$	Strain obtained by SHR to complement <i>Anabaena</i> $\Delta fraC-\Delta fraD$ with <i>fraC fraD</i> from <i>Richelia intracellularis</i> HH01	Sm Sp
<b>CSSA17</b>	<i>sepJ::Strep-tag II</i> (+pCSV3)	Strain obtained by SHR in which <i>sepJ</i> lacks the stop codon and bears a 3' sequence encoding the Strep-tag II is his C-terminal (+ pCSV3)	Sm Sp
<b>CSSA18</b>	<i>hglK::Strep-tag II</i> (+pCSV3)	Strain obtained by SHR in which <i>hglK</i> lacks the stop codon and bears a 3' sequence encoding Strep-tag II (+ pCSV3)	Sm Sp
<b>CSSA28</b>	$\Delta sepJ::sepJ_{CalSC01}$ (+pCSV3)	Strain obtained by SHR to complement <i>Anabaena</i> $\Delta sepJ$ with <i>sepJ</i> from <i>Calothrix rhizosoleniae</i> SC01 (pCSV3 vector integrated as well)	Sm Sp
<b>CSSA29</b>	$\Delta sepJ::sepJ_{PCC7120}$ (+pCSV3)	Strain obtained by SHR to complement <i>Anabaena</i> $\Delta sepJ$ strain with <i>sepJ</i> from <i>Anabaena</i> , used as a positive control (pCSV3 vector integrated as well)	Sm Sp
<b>CSSA30</b>	$\Delta fraC-\Delta fraD$ [pCSSA48]	Strain with replicative plasmid pCSSA48 to complement <i>Anabaena</i> $\Delta fraC-\Delta fraD$ with <i>fraC fraD</i> from <i>R. intracellularis</i> HH01 cloned in pRL3845	Em
<b>CSSA31</b>	$\Delta fraC-\Delta fraD$ [pCSSA49]	Strain with replicative plasmid pCSSA49 to complement <i>Anabaena</i> $\Delta fraC \Delta fraD$ with <i>fraC fraD</i> from <i>Anabaena</i> cloned in pRL3845 (positive control)	Em
<b>CSSA31</b>	$\Delta fraC-\Delta fraD$ [pCSSA50]	Strain with replicative plasmid pCSSA50 to complement <i>Anabaena</i> $\Delta fraC \Delta fraD$ with <i>fraC fraD fraE</i> from <i>C. rhizosoleniae</i> SC01 cloned in pRL3845	Em

DHR, double homologous recombination; SHR, single homologous recombination.

### 2.1.2.2 Culture media and conditions

The different strains of *Anabaena* were grown in axenic culture under photoautotrophic conditions in a modified BG11 medium (Rippka *et al.*, 1979), which consists of: 0.2 mM Na<sub>2</sub>CO<sub>3</sub>,



0.3 mM MgSO<sub>4</sub>, 0.24 mM CaCl<sub>2</sub>, 0.2 mM K<sub>2</sub>HPO<sub>4</sub>, 28.5 μM citric acid, 6 mg/L iron (III) citrate hydrate (19% Fe), 2.4 μM H<sub>3</sub>BO<sub>3</sub>, 9.1 μM MnCl<sub>2</sub>, 1.6 μM Na<sub>2</sub>MoO<sub>4</sub>, 0.8 μM ZnSO<sub>4</sub>, 0.2 μM CoCl<sub>2</sub> and 17.2 mM NaNO<sub>3</sub> as nitrogen source. The culture medium was prepared from a 100x concentrate containing all the components except K<sub>2</sub>HPO<sub>4</sub>, which was added before sterilization in the autoclave. Sometimes a culture medium without any source of combined nitrogen was used (BG11<sub>0</sub> medium), whose composition was as mentioned above except for NaNO<sub>3</sub> that was not added. A culture medium was also used that had ammonium (NH<sub>4</sub><sup>+</sup>) as a source of combined nitrogen. For its preparation, the BG11<sub>0</sub> medium was supplemented with 4 mM of NH<sub>4</sub>Cl buffered with 8 mM TES·NaOH (pH 7.5).

In liquid media, the strains were grown in sterile flasks of 50- or 100-mL capacity depending on whether 25-mL or 50-mL cultures were used respectively; in both cases, they were placed in orbital shakers at 100 rpm in a 30°C chamber with continuous light of 30 μE/(m<sup>2</sup>·s). Some experiments were performed with liquid cultures bubbled with a sterile gaseous mixture containing 1% CO<sub>2</sub> (v/v). These cultures are referred to as bubbled cultures and were supplemented with 10 mM NaHCO<sub>3</sub> regardless of whether they contained BG11 or BG11<sub>0</sub> medium (called BG11C or BG11<sub>0</sub>C respectively). The chamber where the bubbled cultures were grown provides lateral illumination of about 40 μE/(m<sup>2</sup>·s) and has a temperature of 30°C.

Solid culture media were prepared by adding 1% (w/v) agar (Bacto-Agar, Difco) to BG11 or BG11<sub>0</sub> medium after sterilizing agar and nutrients apart. The solid cultures were grown in Petri dishes and incubated in methacrylate boxes illuminated from above at 15-30 μE/(m<sup>2</sup>·s) in 30°C chambers.

When needed, antibiotics were added to the culture medium at a final concentration of: streptomycin sulfate and spectinomycin dihydrochloride pentahydrate (Sm Sp), 2 μg/mL each for liquid cultures and 5 μg/mL each for solid cultures; neomycin (Nm), 15 μg/mL for liquid and 30 μg/mL for solid cultures; and erythromycin (Em), 5 μg/mL for liquid or solid cultures. All antibiotic stocks were sterilized by filtration and dissolved in ultrapure water (milliQ) with the exception of erythromycin which was prepared in absolute ethanol.

### 2.1.2.3 Harvest methods

Unless specifically indicated otherwise, cultures were collected by centrifugation at 4,000 rpm for 5 minutes, washed with BG11<sub>0</sub> medium and resuspended in fresh medium (BG11 or BG11<sub>0</sub>) at a concentration of 1 μg chlorophyll *a* (Chl)/mL. In some cases, the cells were collected by filtration through 0.45-μm pore diameter filters with a vacuum system.

## 2.2 Molecular techniques based on DNA manipulation

### 2.2.1 Plasmids

In this work, several plasmids were used, which are detailed in the following tables. The first table describes the plasmids of general use or specific plasmids I was provided with, while the second table describes the plasmids generated in this work.



**Table 4. General and synthetic plasmids.** All of these plasmids were available for use in our plasmid bank or synthesized by a company. Plasmid pCSAL39 provides resistance to Km and Nm; Km used in *E. coli* and Nm normally used in *Anabaena*.

Name	Resistance	Description	Reference
CalSC01_fraCDE	Ap	500 bp upstream of the <i>fraC</i> gene from <i>Anabaena</i> linked to the <i>fraC</i> <i>fraD</i> <i>fraE</i> genes from <i>Calothrix rhizosoleniae</i> SC01, inserted in EcoRI site of pUCIDT	Synthesized by Integrated DNA Technologies
pCSAL39	Km/Nm	Derived from pMBL-T; contains gene <i>sf-gfp</i> , a sequence encoding a 4-Gly linker and a BsaI site in its 5' end.	Corrales-Gerrero <i>et al.</i> , 2014a
pCSRO	Sm Sp	Contains the polylinker and resistance cassette from pCSV3 and the <i>sacB</i> gene from pRL278	R. López-Igual
pCSV3	Sm Sp	Plasmid derived from pRL500, used for conjugation to <i>Anabaena</i>	Valladares <i>et al.</i> , 2011
pMBL-T	Ap	Commercial vector for cloning purposes	Dominion MBL
pRL3845	Em	Plasmid used to overexpress genes under the <i>Anabaena glnA</i> promoter, also has the pDU1 replicon	C.P. Wolk
pRL443	Ap Tc	Conjugative plasmid, mobilizes ColE1-derived plasmids	Elhai & Wolk, 1988
pRL161	Nm	Contain Nm-resistance cassette C.K1	Elhai & Wolk, 1988
pRL623	Cm	Derived from ColK; helper plasmid for conjugation; contains genes encoding methylases corresponding to <i>Anabaena</i> restriction systems	Elhai <i>et al.</i> , 1997
pSpark	Ap	Commercial vector for cloning purposes	Canvax
pEX-A258-sepJ SC01	Ap	500 bp upstream of the <i>sepJ</i> gene from <i>Anabaena</i> linked to the <i>sepJ</i> gene from <i>C. rhizosoleniae</i> SC01 synthesized with <i>Anabaena</i> codon usage. Inserted in SacI site of pUCIDT	Synthesized by Eurofins Genomics
RINTHH13270_FraCD	Ap	500 bp upstream of the <i>fraC</i> gene from <i>Anabaena</i> linked to the <i>fraC</i> <i>fraD</i> genes from <i>R. intracellularis</i> HH01. Inserted in EcoRI site of pUCIDT	Synthesized by Integrated DNA Technologies
RINTHH17990_sepJ	Ap	500 bp upstream of the <i>sepJ</i> gene from <i>Anabaena</i> , linked to the <i>sepJ</i> gene from <i>R. intracellularis</i> HH01. Inserted in EcoRI site of pUCIDT	Synthesized by Integrated DNA Technologies



Table 5. Plasmids constructed in this work.

Name	Resistance	Description
pCSSA4	Ap	Plasmid containing the fusion in an XbaI site of a fragment near the 5' end of <i>all0813</i> (obtained by PCR with the primers all0813-5 and all0813-6) and a fragment near the 3' end of <i>all0813</i> (obtained by PCR with the primers all0813-7 and all0813-8). This fragment is cloned in SacI site of pSpark
pCSSA5	Ap	Insertion of the C.K1 cassette in the XbaI site between the two PCR products cloned in pCSSA4
pCSSA6	Sm Sp Km	Fragment composed of the 5' and 3' PCR products of <i>all0813</i> , with C.K1 in between, introduced into pCSRO by digestion of both parts with SacI
pCSSA11	Sm Sp	Plasmid generated by cloning of the 5' end of <i>all0813</i> with the primers all0813-12 and all0813-13, which removes the stop codon and adds a sequence encoding 6xHis at the 3' end. This fragment was inserted into pCSV3 by digestion of both parts with EcoRI
pCSSA12	Ap	PCR product from the 3' end of <i>all0813</i> (generated with primers all0813-14 and all0811-15 that remove the stop codon) cloned in pCSAL39 by digestion of both parts with HindIII and BsaI
pCSSA13	Sm Sp	Fragment composed of the 3' end of <i>all0813</i> (without stop codon), a sequence for a 4-Gly linker and the <i>sf-gfp</i> cloned into pCSV3 by the digestion of both parts with KpnI
pCSSA21	Sm Sp	Insert from RINTHH17990_sepJ cloned in pCSV3 after digesting both plasmids with EcoRI
pCSSA22	Sm Sp	Insert from RINTHH13270_FraCD cloned in pCSV3 after digesting both plasmids with EcoRI
pCSSA25	Ap	500 bp of the 3' end of <i>alr2338</i> amplified by PCR with primers alr2338-57 and alr2338-58, which removes the stop codon and inserts the Strep II tag at the 3' end of <i>alr2338</i> ( <i>sepJ</i> ). Fragment introduced into pSpark by EcoRI digestions
pCSSA26	Sm Sp	Fragment corresponding to the 3' end of <i>alr2338</i> attached to the Strep II tag and EcoRI-inserted into pCSV3
pCSSA29	Ap	600 bp of the 3' end of <i>all0813</i> PCR-amplified with primers all0813-24 and all0813-25, which removes the stop codon and adds the Strep II tag at the 3' end of <i>all0813</i> , cloned into pSpark by EcoRI digestion
pCSSA30	Sm Sp	Fragment corresponding to the 3' end of <i>all0813</i> attached to the Strep II tag and EcoRI-inserted into pCSV3
pCSSA46	Sm Sp	Fragment from pEX-A258-sepJ SC01 inserted in pCSV3 with SacI
pCSSA47	Sm Sp	PCR product using <i>Anabaena</i> DNA and primers alr2336-62/alr2336-63, which includes the <i>sepJ</i> ( <i>alr2338</i> ) gene and 600 upstream bp, cloned into pCSV3 with SacI (used as a positive control in complementation experiments)
pCSSA48	Em	The <i>fraC fraD</i> genes (RINTHH13270_FraCD) from <i>R. intracellularis</i> HH01 cloned in pRL3845 by digestion with SmaI and BamHI.
pCSSA49	Em	The <i>fraC fraD</i> genes from <i>Anabaena</i> amplified with primers alr2392-93-1 and alr2392-93-2 cloned in pRL3845 by digestion with SmaI and BamHI (used as a positive control in complementation experiments)
pCSSA50	Em	The <i>fraC fraD fraE</i> genes (CalSC01_fraCDE) from <i>C. rhizosoleniae</i> SC01 cloned in pRL3845 by digestion with SmaI and BamHI



## 2.2.2 DNA isolation

### 2.2.2.1 Isolation of plasmid DNA from *E. coli*

Plasmidic DNA was obtained from 3 mL of liquid *E. coli* cultures grown overnight at 37°C with shaking. The plasmids were isolated using the Nucleospin Plasmid kit (Machery Nagel), following the manufacturer's instructions.

### 2.2.2.2 Isolation of DNA from PCR reactions or agarose gel

The DNA fragments obtained from a PCR or after an electrophoresis were isolated by using the Isolate II PCR & Gel kit (Bioline) following the manufacturer's instructions.

### 2.2.2.3 Isolation of total DNA from *Anabaena*

Cells from 50-mL liquid cultures were harvested by filtration and resuspended in a final volume of 500  $\mu$ L T<sub>1/10</sub> (10 mL Tris-HCl, 0.1 EDTA-Na<sub>2</sub>, pH 8) in an Eppendorf tube. Then, 150  $\mu$ L of sterile glass beads (250-300 nm diameter, sterilized at 180°C), 20  $\mu$ L of 10% sodium dodecyl sulfate and 450  $\mu$ L phenol-chloroform (1:1, v/v) were added. This mix was shaken in a vortex for 1 minute and incubated in ice (~ 0°C) for 1 minute. This step was repeated 3-4 times (depending on whether the original culture was overgrown). After that, the lysate was centrifuged at 16,000 x g, 15 minutes, at 4°C and the resulting clear supernatant was transferred to a new Eppendorf tube for successive extractions with 1 volume of each 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. Tubes were incubated for 1 hour (at least) at -20°C to allow DNA precipitation. After centrifugation at 16,000 x g, 15 minutes at 4°C, the pellet containing genomic DNA was washed with 70% ethanol (v/v), dried at room temperature and resuspended with 50  $\mu$ L of ultrapure water (milliQ).

## 2.2.3 DNA quantification

The isolated DNA was subjected to agarose gel electrophoresis to ensure the presence of nucleic acids. The concentration of DNA in the samples subjected to electrophoresis could be estimated by comparison with samples of known concentration. For a more precise quantification, NanoDrop *Spectrophotometer ND-1000* was used, performing the quantification of the samples at an absorbance of 260 nm (extinction coefficient  $\epsilon = 0.020 \text{ mL}/(\text{mg} \cdot \text{cm})$ ).

## 2.2.4 Polymerase Chain Reaction (PCR)

PCR for routine analysis (such as checking genotypes or verifying inserts) was performed with Biotaq DNA Polymerase, while for PCR for cloning, the iProof high-fidelity polymerase was used. Both polymerases are included in kits that include the necessary buffers for reaction and were always used following the manufacturer's instructions. Regardless of the polymerase used, the deoxyribonucleotide triphosphates (dNTPs) and specific primers for each reaction were added and the PCRs were carried out in a T100 thermal cycler (Bio-Rad).



As stated in the previous paragraph, to make a PCR it is necessary to use primers specific to the sequences to be amplified. Table 6 describes all the primers used in this work.

**Table 6. Primers used in this work.** The sequences in red color indicate targets for restriction enzymes, and the underlined letters indicate the sequence of the tags that were inserted by means of the primers.

Primer	Sequence (5'→ 3')
All0813-5	GCGAGCTCTTTGGCATTACC
All0813-6	TTTCTAGATTCCCGATTCCCTTG
All0813-7	CTTCTAGACGTCCTCATGACTCG
All0813-8	ACCCCTTTAACTACCGCAGATACT
All0813-9	CCAAATTCTCTACCCCTAGCCAC
All0813-10	GGGAGGCTCGACAAAACGC
All0813-12	GAATCCGCCAGCTTGACAGGGG
All0813-13	GAATTCCTAGTGATGGTGATGGTGATGCGGACAACGTG
All0813-14	GATAAGCTTGGTACCCTAGAGTCAAGCCCAGATG
All0813-15	GGTCTCACGCCGGACAACGTGGATG
All0813-16	CTGACCGCAAGATTTTCGCC
All0813-17	TAGTTGCTTGCCATCTAGATTTTTAG
All0813-18	GTGATGGTGATGGTGAT
All0813-24	GAATTCCTACTTTTTCGAACTGCGGGTGGCTCCACGGACAACGTGGATGAAC
All0813-25	GCCAATTATCCAATGCACGCTTAAT
Alr2336-62	ATGAGCTCCTGGGATGCGTAAGTTC
Alr2336-63	AAGAGCTCTTAACCTTCTGCATTGGC
Alr2336-64	CTACTGAGCCAGAAGTCCAGAG
Alr2338-57	TGTGCGGCTAAAGTCCATCCAG
Alr2338-58	GAATTCCTACTTTTTCGAACTGCGGGTGGCTCCACCTTCTGCA TTGGCAGG
Alr2338-61	CAGCACTGGTATAGGAAC
Alr2392-93-1	CCCGGGATGTTTGAAGATTGAC
Alr2392-93-2	GGATCCTCACTGCTGCGGTGGCG
Alr2394-5	TATAAGCTTCGCTGTCTTTGTGGGT
Alr2394-6	CAAGGTCTCACGCCAAACTCCC GGCGTG
Alr2394-10	GGTACGCGATCGCATCC
Alr2394-11	CAAGGCACATTAGTAGAAGCATC
FraC Calothrix	CTTCACTGAGTCTTCATTTCCCTTAC
FraCD_Rich-2	GCGATTTGATGCTATTGGC
FraCD_Rich-3	GGTCAAAAGTTGTGATTTCTTCTG
FraCD_Rich-4	CCCGGGATGCAAACATTCCTCC
FraCD_Rich-5	GGATCCTTATCTTTGAGATGGTGG
FraE Calothrix	GGTGAAATCAGGTAATGCTTG





<b>GFP-8</b>	TTATTTGTATAGTTCATCCATGCC
<b>SepJ Calothrix_1</b>	CAACGTGGTCGACAAATCAAC
<b>SepJ Calothrix_5</b>	GTTCTGAGCGCTGCCATCTTTAG
<b>SepJ Calothrix_6</b>	GATGCTCCCAATTTGCGAATAC
<b>SepJ_Rich-2</b>	GTGGGATACACCTAACTTCCG
<b>SepJ_Rich-3</b>	CCTACTTTATCCGGTCTCAACAG
<b>Strep-tag reverse</b>	CTTTTCGAACTGCGGGTGGCTCCA

### 2.2.5 DNA electrophoresis

DNA fragments were separated by agarose gel electrophoresis (Sambrook & Russell, 2001). Gels were made with 0.8% or 2% agarose (w/v; depending on whether larger or smaller fragments were to be separated) in TBE buffer (90 mM Tris-Borate, 2 mM EDTA- $\text{Na}_2$ , pH 8), supplemented with 0.03% (v/v) *GelRed Nucleic Acid Gel Stain* (Biotium). Alternatively, to the *GelRed*, to visualize the DNA in the gel ethidium bromide at a concentration of 1 mg/mL was added. Before loading samples into the gel, they were mixed with 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)). The size of the fragments was estimated by using the commercial size marker *1 kb leader* (Biotools) for comparison. The electrophoresis was carried out using a *Mini-Sub Cell GT* or *Wide Mini-sub Cell GT* (Bio-Rad) apparatus, and finally 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.

### 2.2.6 DNA sequencing

Isolated PCR products or isolated plasmids were sequenced by the Secugen S.L. Company service (address Ramiro de Maeztu 9, Madrid, Spain). To perform the sequencing, this company used primers from their own service for commercial vectors or our own primers for other plasmids.

### 2.2.7 Enzymatic treatments of DNA

#### 2.2.7.1 Restriction

Both DNA PCR products and plasmidic DNA were digested with restriction enzymes from New England Biolabs or Thermo Scientific (FastDigest), and each enzyme was accompanied by its own necessary buffers. The incubation was from 15 minutes to 2 hours, according to the manufacturer's recommendations, at a temperature of 37°C. After treatment, endonucleases were inactivated and the buffer salts were removed using the *Isolate II PCR & Gel kit* (Bioline).

#### 2.2.7.2 Ligation

To perform the ligation of different DNA fragments, a ligation mixture composed of vector, insert, ligation buffer and T4 bacteriophage DNA ligase (T4 ligase) in a final volume of 10  $\mu\text{L}$  was used. The vector:insert concentration ratio was approximately 1:3, except when using commercial plasmids such as pSpark, in which case the ratio was 1:5. The T4 ligase used was of





two types, T4 ligase from Dominion MBL or T4 ligase that was part of the pSpark DNA Cloning System kit. Both ligases were provided with the necessary buffers. Finally, the ligation mixture was incubated for 1 to 24 hours at a temperature of 21°C.

### 2.2.7.3 Dephosphorylation

To avoid recirculation, the 5' protruding extremes of linearized vectors were dephosphorylated. DNA from the vector was incubated with rAPid alkaline phosphatase (Roche) following the manufacturer's instructions.

### 2.2.8 RNA radioactive probes

DNA fragments that had been previously amplified by PCR were used as a template for radioactively labeling the DNA. 40-50 ng of two-stranded DNA were mixed with 10 pmol of oligonucleotides in a final volume of 34.5 µL. The mixture was denatured for 3 min by heating to 100°C. A mixture containing 0.03 mM dATP, 0.03 mM dGTP, 0.03 mM dTT, 1 U Klenow enzyme (Fermentas), 5 µl Klenow buffer 10X (Fermentas) and 5 µCi of ( $\alpha$ -<sup>32</sup>P)-dCTP (10 µCi/µL)-dCTP was then added. After incubation at 37°C for 30 min, the reaction was terminated by addition of 2 µL of 1 mM dCTP and incubation for 5 min at 37°C. Finally, the unincorporated ( $\alpha$ -<sup>32</sup>P)-dCTP was removed by gel filtration. After quantifying the incorporated radioactivity by scintillation counting (in a model LS 6000 Beckman counter), the labeled DNA fragment was used as a radioactive probe in *northern* blot experiments.

## 2.3 Genetic methods

### 2.3.1 Transformation of *E. coli* by heat shock

Exogenous DNA in the form of a plasmid was introduced into *E. coli* DH5 $\alpha$  and HB101 strains by means of thermal shock transformation. To perform the transformation in an Eppendorf tube, 100 µL of competent cells were mixed with either 10 µL of a ligation product or 2 µL of purified plasmid. These mixtures were incubated for 10 minutes on ice, then subjected to a thermal shock at 42°C for 90 seconds and finally incubated on ice for 5 minutes. After this last incubation, 1 mL of liquid LB medium was added, and the mix was incubated in a shaker for one hour at 37°C to allow expression of resistance genes. The cells from the transformation were collected by centrifugation and the pellet (resuspended in approximately 20 µL of LB medium) was spread on a Petri dish with LB medium supplemented with the corresponding antibiotic. For pSpark-derived constructs, the plates were supplemented with the antibiotic and 40 µg/mL of X-gal, which facilitates the selection of positive colonies ( $\alpha$  complementation assay) (Sambrook & Russell, 2001).

### 2.3.2 Conjugation to *Anabaena*

The conjugation protocol is described in Elhai & Wolk (1988b). Two strains of *E. coli* are needed to perform the conjugation. Strain ED8654 that bears plasmid pRL443 which provides conjugation functions, and strain HB101 that carries the *cargo* plasmid and the helper plasmid pRL623. The latter bears genes that encode methylases for AvaI, AvaII and AvaIII restriction



sites (Elhai *et al.*, 1997), which protects the *cargo* plasmid from digestion by the *Anabaena* restriction endonucleases, as well as the *mob* gene. In a first step, pRL433 is transferred from ED8654 to HB101, and in a second step, the cargo plasmid is transferred from HB101 to *Anabaena*.

Once HB101 [pRL623] is transformed with the construct to be sent to *Anabaena*, the conjugation is performed following the next steps. Two independent pre-inocula each in 3 mL of liquid LB medium supplemented with the antibiotics corresponding to the strains of *E. coli*, HB101 [pRL623 (30 µgCm/mL) and *cargo* plasmid (antibiotic depending on particular plasmid)] and ED8654 [pRL443 (50 µg Ap/mL)], are incubated at 37°C with shaking. When the cultures reach their stationary phase, 250 µL of the ED8654 strain and 350 µL of the HB101 strain are taken, 10 mL of fresh medium with the appropriate antibiotics are added, and the cultures are incubated at 37°C for 2.5 hours in a shaker. Then, the cells are collected by centrifugation, washed 3 times with fresh LB to remove the antibiotics and resuspended together to approximately 0.5 mL. The cell suspension of the *E. coli* strains is incubated at room temperature for at least 2 hours, and an amount of *Anabaena* (wild type or a mutant used as parental strain) containing 10 µg of Chl (5 µg Chl if the resistance cassette is C.K1) is added, everything is gently homogenized, and finally spread on a nitrocellulose filter (Nucleopore) on a Petri dish with BG11 medium + 5% LB (v/v) without antibiotics. The Petri dish is incubated at 30°C under low white light intensity (10 µE/(s·m<sup>2</sup>)) for at least 3 hours and then incubated at standard light conditions (30 µE/(s·m<sup>2</sup>)). Twenty-four hours after inoculation, the nitrocellulose filter is transferred to a new BG11 plate where it is incubated for a further 24 hours before transferring the filter to a plate of BG11 medium supplemented with the corresponding antibiotic(s). The filter is transferred to a new plate with fresh medium and antibiotic(s) every 48 hours until antibiotic-resistant colonies appear.

When the transferred plasmid has not a replication origin for *Anabaena*, two types of homologous recombination can take place depending on the type of construct used:

- Single cross-over: in this case, the plasmid only bears a fragment that is homologous with a region of the chromosome. Here, the final result is the integration of the plasmid in a specific place of the *Anabaena* chromosome through homologous recombination. To ensure that the conjugation was performed correctly, isolated colonies are checked by PCR (Fig. 12).

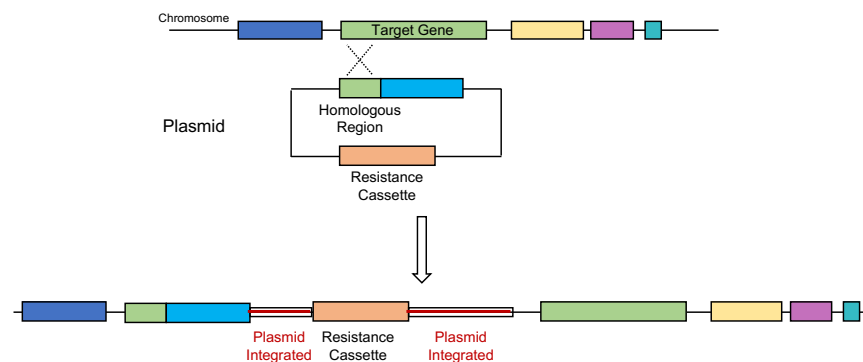
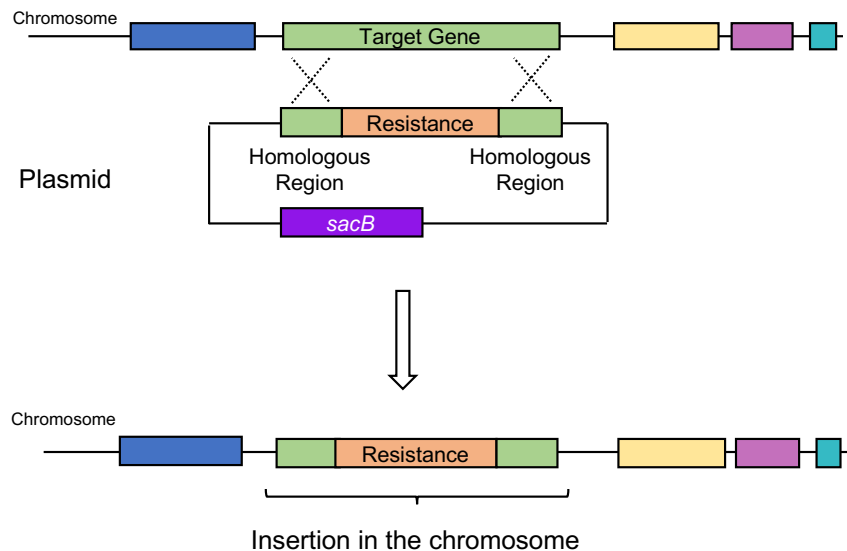


Fig. 12. Schematic representation of integration through single cross-over.



- Double cross-over: the plasmid has two fragments homologous to sequences in a region of the chromosome. The final result will be the integration of a DNA fragment in a specific region of the chromosome in a process of gene replacement (in this case the whole plasmid is not inserted). The process takes place in two steps and includes a step of positive selection through sucrose sensitivity provided by the *sacB* gene (encoding the *Bacillus subtilis* levansucrase; this enzyme converts sucrose into levan, which accumulates in the periplasm and is toxic). After the first cross-over, the whole plasmid will be inserted, which is checked by PCR. In case of a positive result (ideally, presence only of chromosomes with the insertion of the whole plasmid), the strain is transferred to a plate of BG11<sub>0</sub> medium supplemented with ammonium (NH<sub>4</sub><sup>+</sup>) (as nitrogen source) + 5% sucrose. Exconjugants that can grow in sucrose-supplemented medium should have lost the plasmid vector by a second cross-over event. In this work, when the double recombination was tried, a neomycin resistance gene cassette (C.K1) was introduced, so exconjugants that can grow in sucrose can also grow in the presence of neomycin (Fig. 13). Finally, to verify that the segregation was complete and that the mutation is correct, the selected clones are tested by PCR. In some occasions, clones with the correct mutation but still not segregated (i.e., containing some wild-type chromosomes) are obtained. To facilitate the segregation, some clones that can grow in the presence of sucrose are incubated in liquid BG11 medium with antibiotic (Nm); once grown, they are subjected to sonication in a cleaning bath to fragment the filaments to a size of between 3-5 cells per filament. After sonication, serial dilutions are plated in BG11 medium with antibiotic (Nm). The exconjugants obtained should originate from a short filament, increasing the probability that all the cells in the filament are genetically homogeneous.



**Fig. 13. Schematic representation of gene replacement through double cross-over.**



## 2.4 Molecular techniques based on RNA manipulation

### 2.4.1 RNA isolation from cyanobacteria

Isolation of total RNA from cyanobacteria was performed following a modified protocol based on the hot phenol method of Mohamed & Jansson (1989).

Cells from 50 mL of culture were collected by centrifugation (see 2.1.2.3), washed with 5 mL T<sub>50</sub>E<sub>100</sub> buffer (50 mM Tris-HCl, 100 mM EDTA-Na<sub>2</sub>, pH 8) and resuspended in 2 mL of the same buffer. This cell suspension was transferred to a 2-mL Eppendorf tube where it was centrifuged. After centrifugation, the supernatant was removed and the pellet was frozen with liquid nitrogen and stored at -20°C until use.

The cell pellet was resuspended in 300 µL of resuspension buffer (300 mM sucrose, 10 mM sodium acetate pH 4.5) supplemented with 100 µL of 250 mM EDTA-Na<sub>2</sub> and 400 µL of lysis buffer (2% SDS, 10 mM sodium acetate pH 4.5). When the pellet was well homogenized in this mixture, 1 mL of acid phenol (pH 4.5) preheated to 65°C was added, and the mixture was shaken in a vortex for 30 s and incubated at 65°C for 2 minutes. The process of shaking and incubation was repeated 3 times. Then the mixture was centrifuged for 5 minutes at 4°C and 15,000 x g. The aqueous phase was transferred to a clean Eppendorf tube and treated successively with 1 mL of acid phenol preheated to 65°C, then with a mixture of acid phenol:chloroform (1:1 (v/v)), and finally with 1 mL of chloroform (Merck). Before adding the solvents previously indicated, the mixture was vortexed for 30 seconds, incubated at 65°C for 2 minutes and centrifuged for 5 minutes (15,000 x g, 4°C), always collecting the aqueous phase. After treatment with chloroform, the aqueous phase was again transferred to a clean Eppendorf tube, 1 mL of isopropanol was added and the mixture was incubated at -20°C for at least 1 hour to precipitate the nucleic acids. Then, the sample was centrifuged for 30 minutes, the supernatant was discarded, and the pellet was washed with 70% ethanol (v/v) and dried at room temperature. Finally, the nucleic acid precipitate was resuspended in 90 µL of water (previously treated with diethylpyrocarbonate [DEPC]) and treated with 10 units of I-turbo DNase (Ambion) at 37°C for 30 minutes in the buffer supplied by the manufacturer. After 30 minutes, the DNase was removed using the DNase inactivating reagent that is supplied with the DNase.

### 2.4.2 RNA quantification

The concentration of RNA was quantified in the same way as the concentration of DNA, using the Nanodrop Spectrophotometer ND-1000. The measurement was performed at an absorbance of 260 nm (extinction coefficient  $\epsilon = 0.020 \text{ mL}/(\text{mg}\cdot\text{cm})$ ).

### 2.4.3 RNA electrophoresis in agarose gels

RNA electrophoresis was performed in 1% (w/v) agarose gels in MOPS 1x buffer (40 mM MOPS, 10 mM sodium acetate, 1 mM EDTA pH 7.2) to which formaldehyde 2% (v/v) was added. The samples were prepared to load a final volume of 20 µL, which was composed of 10 µg total RNA and 15 µL loading buffer consisting of 60% (v/v) deionized formamide, 6% (v/v) formaldehyde, MOPS 1x buffer, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 0.1% (v/v) ethidium bromide, and 10% (v/v) glycerol. Before loading the samples, they were incubated for



3 minutes at 80°C and immediately cooled in ice until they were loaded into the gel. Once all the samples were loaded, a commercial Riboruler High Range RNA Ladder (Thermo Scientific) was also loaded, and finally the gel was run for 1-2 hours at 100 V.

#### 2.4.4 RNA transfer to nylon filters

To perform this step, after the electrophoresis the gel is washed three times for 5 minutes each time with abundant deionized water and then balanced with 10x SSC buffer (1.5 M NaCl and 150 mM sodium citrate, pH 7) for 10 minutes. The RNA was then transferred to a nylon filter GeneScreen Plus (PerkinElmer) by capillarity for 16 hours. When the transfer was completed, the filter was briefly washed with 2x SSC buffer and the RNA was fixed to the nylon filter by irradiating it with UV light for 2 minutes and allowing it to dry for 1-2 hours in an oven at 80°C.

#### 2.4.5 Hybridization with radioactive probes

Before hybridizing the RNA-carrying nylon membrane with radioactive DNA probes, the filter was incubated with a pre-hybridization solution (0.3 M phosphate buffer [pH 6.9], 7% SDS and 1 mM EDTA [pH 8]) for 1-2 hours in a hybridization oven at 65°C. After pre-hybridization of the filter, the radioactively marked DNA probe (section 2.2.8) was denatured for 3 minutes at 95°C and added to the filter, which was incubated for 12-16 hours in the hybridization oven at 65°C. After the incubation, the filter was washed with a first washing solution (0.3 N NaCl, 0.3 M trisodium citrate dihydrate and 0.1% SDS) for 10 minutes at 65°C and then with a second solution (0.15 M NaCl, 0.015 M trisodium citrate dihydrate and 0.1% SDS) at 65°C for 10 minutes again. Although the agarose gel was loaded with the same amount of RNA in each well, the filters were hybridized with a probe of *mnpB* gene (Vioque, 1997) as a control for the amount of RNA loading and transfer.

#### 2.4.6 Detection of radioactivity

Both for detection of hybridization signals in the GeneScreen Plus nylon membranes (PerkinElmer) and for analyzing those signals, electronic auto-radiography with the Cyclone Storage Phosphor System (PerkinElmer) and image analysis with the program *Optiquant* (Packard) were carried out.

### 2.5 Protein purification from *Anabaena*

#### 2.5.1 Induction and cellular lysis

The strains from which the purification of proteins was performed were grown in 1.5 L of BG11C medium (supplemented with antibiotics if necessary) with bubbling (see section 2.1.2.2). Once grown, they were harvested by centrifugation at 6,000 x g for 10 minutes, the pellet was resuspended in 1.5 L of liquid BG11C without antibiotics, and an incubation was performed for 48 hours with bubbling (induction under nitrogen deprivation). After induction, the cells were collected again by centrifugation until a single pellet was obtained in a Falcon tube. The pellet was resuspended in 10 mL of lysis buffer (20 mM Tris/HCl pH 8, 150 mM  $\epsilon$ -amino-n-caproic acid, 5 mM  $\beta$ -mercaptoethanol and one pill of protease inhibitor cocktail [Roche]), which was



poured dropwise with the aid of a pipette into a mortar containing liquid nitrogen. Once poured into the mortar, the sample was grinded until all the liquid nitrogen was evaporated to give a cyanobacterial powder. The powder was collected with a spoon and poured into a clean Falcon tube to which a volume of 2 mL of lysis buffer was added; when the sample was thawed, it was supplemented with 1 mg/mL lysozyme and 1% (w/v) n-dodecyl- $\beta$ -D-maltoside (DDM), and this mixture was incubated on a rotary shaker at 4°C for 1 hour. The sample was then centrifuged at low speed (1,000 x g) for 5 minutes at 4°C to remove any unbroken cells; the supernatant was passed into a clean Falcon tube and placed on ice to begin the purification of the sample.

### 2.5.2 Complex purification using the Strep-tag II

The supernatant obtained in the above-described process (approximately 12 mL) was supplemented with 2  $\mu$ g of avidin and incubated with 120  $\mu$ L (v/v) MagStrep "type3" XT Beads (IBA) (previously equilibrated with Wash Buffer 1 –described below– following the manufacturer's instructions) for 1.5 h at 4°C on a rotary shaker. After incubation, the sample was transferred to an Eppendorf tube in a magnetic rack that allowed us to separate the magnetic beads from the rest of the sample. After collecting all the magnetic beads, they were resuspended in 1.2 mL of Wash Buffer 1 (20 mM Tris/HCl [pH 8], 100 mM  $\epsilon$ -amino-n-caproic acid, 5 mM  $\beta$ -mercaptoethanol and 1 mM EDTA), separating again the supernatant from the magnetic beads using the magnetic rack and freezing the supernatant in liquid nitrogen before being stored at -80°C. This step was repeated using Wash Buffer 2 (20 mM Tris/HCl [pH 8], 150 mM  $\epsilon$ -amino-n-caproic acid, 5 mM  $\beta$ -mercaptoethanol and 1 mM EDTA) and Wash Buffer 3 (20 mM Tris/HCl [pH 8], 150 mM  $\epsilon$ -amino-n-caproic acid, 5 mM  $\beta$ -mercaptoethanol, 1 mM EDTA and 0.5% NP-40). After using the latter buffer, the magnetic beads pellet was resupplied again in 120  $\mu$ L of Wash Buffer 3, where a small aliquot (20  $\mu$ L) of the magnetic beads were incubated with biotin elution buffer (20 mM Tris/HCl [pH 8], 150 mM  $\epsilon$ -amino-n-caproic acid, 5 mM  $\beta$ -mercaptoethanol, 1 mM EDTA and 50 mM D-biotin) at room temperature for 10 min with occasional vortexing to bring beads into suspension. To ensure that no proteins remained attached to the magnetic beads, they were incubated for 5 minutes at 95°C with sample buffer. Finally, the rest of magnetic beads (~100  $\mu$ L) that were resuspended in Wash Buffer 3 were frozen in liquid nitrogen and stored at -80°C until further analysis in the mass spectrometer. All aliquots (from each of the washes and elution buffer) were checked by Western blot analysis to see if the purification had come out correctly.

### 2.5.3 Complex purification using the GFP

The supernatant obtained in section 2.5.1 was mixed with 75  $\mu$ L (v/v) of GFP-Trap<sup>®</sup>\_MA bead (Chromotek) previously equilibrated in dilution buffer as indicated in the manufacturer's instructions. This mixture of supernatant and magnetic beads was incubated for 1 hour at 4°C and shaking. After incubation, the magnetic beads were separated from the supernatant by passing the supernatant to a 1.5-mL Eppendorf tube on a magnetic rack. The beads were washed twice with 500  $\mu$ L of washing buffer (10 mM Tris/Cl [pH 7.5], 150 mM NaCl and 0.5 mM EDTA); after the last washing the beads were resuspended in 75  $\mu$ L of washing buffer, frozen with liquid nitrogen and stored frozen at -80°C until processing. Before freezing the magnetic beads (75  $\mu$ L suspension), 25  $\mu$ L were taken for elution and analysis by Western Blot to check if the isolation had gone correctly. To this aliquot, 50  $\mu$ L of glycine-elution buffer (200 mM glycine, pH 2.5) was added and incubated for 30 seconds in constant agitation. After incubation, the supernatant





was transferred to a clean Eppendorf tube containing 5  $\mu$ L of 1 M Tris (pH 10.4). This step was repeated again to improve protein elution. Finally, to elute proteins attached to the magnetic beads, these were resuspended in 100  $\mu$ l of sample buffer and incubated for 10 min at 95°C. Aliquots of each of the washes and the starting supernatant were kept to be checked by Western blot.

## 2.5.4 Protein biochemical methods

### 2.5.4.1 SDS-PAGE

For protein separation by polyacrylamide gel electrophoresis, the protocol of Laemmli (1970), which is described in Sambrook & Russell (2001), was followed. The polyacrylamide gels are composed of two gels, a running gel that is prepared with acrylamide:bisacrylamide (29:1) at 12% in 375 mM Tris-HCl (pH 8.8) and the stacking gel (where the wells are located) prepared with acrylamide:bisacrylamide (29:1) at 12% in 125 mM Tris-HCl (pH 6.8). Both gels are supplemented with 0.1% (w/v) SDS, which facilitates the denaturation of the proteins allowing them to migrate better through the gel. To make the polymeric gels, it is necessary to add 0.05% APS (ammonium persulfate) and 0.1% TEMED just before pouring the gels between the crystals. Before loading into the stacking gel wells, the samples are supplemented with 2x protein sample buffer consisting of 125 mM Tris/HCl (pH 6.8), 20% (v/v) glycerol, 4% SDS (v/v), 10% (v/v) 2-mercaptoethanol and 0.0025% (w/v) bromophenol blue. Finally, to ensure that all proteins are denatured, samples are heated for 5 minutes at 95°C just before loading into the gel.

To carry out the electrophoresis, the gels were immersed in an electrophoresis buffer consisting of 25 mM Trizma base, 192 mM glycine and 0.1% SDS (pH 8.3). In addition to the above mentioned, a protein size marker (gTPbio) was loaded into the gels.

### 2.5.4.2 Transference of proteins to PVDF membranes

In order to perform protein immunodetection, proteins separated by size in a polyacrylamide gel are transferred to a polyvinylidene difluoride (PVDF) Hybond-P 0.45- $\mu$ m pore size membrane (GE Healthcare) using a GE Healthcare Amersham semi-Dry Transfer system. Prior to the transfer, the PVDF membranes were activated by soaking them in methanol for a few seconds, then washed in ultrapure water for about 10 minutes (changing the water at least twice) and finally soaked in transference buffer (478 mM Trizma base pH 8, 0.03% (w/v) SDS, 39 mM glycine and 5% (v/v) methanol) for 10 minutes. Regardless of the solution where the membrane was immersed, it was always shaken in a rotatory shaker to prevent the membrane surface from drying out. The gels were also immersed for 10 minutes in transference buffer. Once impregnated by the transference buffer, the transfer components were assembled as a "sandwich" consisting of (from bottom to up): three Whatman papers previously wetted in transference buffer, the PVDF membrane, the acrylamide gel and three other Whatman papers also wetted in transference buffer. The electrical current applied for transfer was calculated using the following formula:  $I = S \cdot t$ ; where I is the current in mA, S the surface of the gel in  $\text{cm}^2$  and t is the thickness of the gel in mm. The transfer was carried out for 70 min at a continuous current intensity.





### 2.5.4.3 Staining PVDF membranes

The membranes were incubated with DB71 solution (0.008% (v/v) (DB71 in 40% (v/v) ethanol and 10% (v/v) acetic acid) for 5 minutes with shaking. After staining, the membrane was rinsed with double distilled water and the protein size marker was marked with a pen. Then, the membrane was destained with DB71 destaining solution (50% (v/v) ethanol and 150 mM NaHCO<sub>3</sub>) for 5 minutes at room temperature on a shaker. Finally, the membrane was incubated once with PBS and twice with PBS-T (each time for 5 minutes at room temperature and shaking). (This procedure was dispensable when a prestained protein size marker was used.)

### 2.5.4.4 Western blotting for Strep-tag II detection

The membranes were incubated with blocking buffer containing PBS buffer (5.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>HPO<sub>4</sub>, 137 mM NaCl and 2.7 mM KCl [pH 7.4]), 3% BSA (w/v) and 0.5% v/v Tween 20, for 1 hour at room temperature with gentle shaking or overnight at 4°C (with shaking too). Then they were washed 3 times with 20 mL of PBS-Tween buffer (PBS buffer with 0.1% v/v Tween 20) per membrane for 5 min with shaking. After the washings, the membrane was immersed in 10 mL of PBS-Tween buffer and 10 µL Biotin Blocking Buffer was added, and then it was incubated for 20 min at room temperature and gentle shaking. For the detection of the Strep-tag II, the membrane in 10 mL of PBS-Tween buffer was supplemented with 10 µL of Strep-Tactin horseradish peroxidase conjugate diluted 1:100 in enzyme dilution buffer (PBS with 0.2 % w/v BSA and 0.1% v/v Tween 20) and incubated during 60 min at room temperature and gentle shaking. After the incubation, the membrane was washed twice with PBS-Tween buffer for 5 min and then washed twice with PBS-buffer for 5 min (all washes were done at room temperature and with gentle shaking). Finally, *WesternBright ECL* (Advansta) reagents were used for Strep-Tactin detection and exposure was made with *Hyperfilm* (GE Healthcare).

### 2.5.4.5 Western blotting for GFP detection

Western Blot analysis was carried out on membranes with transferred proteins. The membrane was incubated with a blocking buffer consisting of PBS-T (5.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>HPO<sub>4</sub>, 137 mM NaCl and 2.7 mM KCl [pH 7.4] and 1% (v/v) Tween-20) with 5% (w/v) milk powder for 1 hour at room temperature (or overnight at 4°C) and shaking. After the blocking step, the membrane was washed three times with PBS-T for 5 minutes and then incubated with the primary antibody dissolved in 5% milk PBS-T (GFP monoclonal antibody, a 1:1,000 dilution was used) for 1 hour at room temperature and shaking. Then the membrane was washed three times with PBS-T for 10 minutes with shaking, and after the last washing the secondary antibody (anti-rabbit antibody) was added to a 1:10,000 dilution, and the membrane was incubated overnight at 4°C and with shaking. The next step was to wash the membrane twice with PBS-T and then 3 times with PBS for 5 min with shaking and room temperature. Finally, *WesternBright ECL* (Advansta) chemiluminescent substrate was added and radiographic film (GE Healthcare) was used for the detection of the secondary antibody.

### 2.5.5 Mass spectrometry analysis

The identification of peptides from the samples isolated using the magnetic beads against Strep-tag II and GFP, respectively, was done by mass spectrometry performed by Dr. Roman Ladig



from Prof. Enrico Schleiff's group at Goethe University of Frankfurt (Germany). The samples were processed as described in Brouwer *et al.* (2019).

The label-free quantification (LFQ) values were calculated with MaxQuant quantitative proteomics software for all experiments. The LFQ values for the experiments for SepJ-GFP (three with baits and three control experiments with same affinity column material), for SepJ-Strep-tag II (three with baits and three control experiments with same affinity column material) and for HglK-Strep-tag II (five with baits and three control experiments with same affinity column material) were extracted from the list. All proteins for which no LFQ value was calculated for any of the experiments with bait were removed. Subsequently, all zero LFQ values were set to the obtained minimal LFQ value in each of the three sets. This simulates a high false negative discovery rate. Subsequently, the p-value for the distribution of the ratio was calculated by t-test. Two different cut offs were used, namely  $p < 0.05$  (95% confidence) and  $p < 0.1$  (90% confidence), the latter considering the low abundance of the bait.

## 2.6 Peptidoglycan isolation

The isolation of the murein sacculi (peptidoglycan, PG) was carried out following the protocol of Lehner *et al.* (2011) with some modifications. The strains were grown in 50 mL of liquid BG11 medium to a concentration of 3-4  $\mu\text{g Chl/mL}$  (supplemented with antibiotics if necessary) and collected by centrifugation. To isolate peptidoglycan from induced cultures, the BG11-grown cells were washed, resuspended in 50 mL of liquid BG11<sub>0</sub> medium (without antibiotics) and incubated for 24 or 48 hours (section 2.1.2.2). After harvesting, the pellet was resuspended in 1 mL of PBS (5.5 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{HPO}_4$ , 137 mM NaCl and 2.7 mM KCl [pH 7.4]) and the filaments were fragmented by sonication in a cleaning bath to obtain individualized cells or filaments with a length of between 2-3 cells per filament. The cell suspension was poured dropwise into 10 ml of a boiling 6% (w/v) SDS solution, and the mixture was boiled for at least 2 hours with strong stirring. After this, the samples were incubated overnight at 37°C with gentle stirring. Then, the samples were centrifuged at 320,000  $\times g$  at 25°C for 35 minutes in a Beckman Coulter Optima XPN-100 ultracentrifuge using rotor 90Ti. After centrifugation, the supernatant was removed and the pellet was resuspended in 3 ml of 3% SDS (w/v) and boiled for 2.5 hours. The samples were then sedimented by centrifugation (320,000  $\times g$  at 25°C for 35 minutes). Again, the supernatant was removed and the pellet was resuspended in 2 mL of 0.05% (w/v) SDS and boiled for 2.5 hours (always maintaining stirring). After boiling, the samples were sedimented as indicated above, the supernatant was removed and the pellet was resuspended in 1.5 mL of 50 mM sodium phosphate buffer (pH 6.8) supplemented with 50  $\mu\text{g}$  (corresponding to 2U) of  $\alpha$ -chymotrypsin (from bovine pancreas, Sigma). The mixture of sample and enzyme was incubated at 37°C overnight with gentle stirring. After treatment with the enzyme, 0.5 mL of double distilled water and 750  $\mu\text{L}$  of 6% (w/v) SDS were added to the samples, which were boiled for 2.5 hours and centrifuged at 320,000  $\times g$  at 25°C for 35 minutes. The supernatant was removed and the pellet was washed three times with double distilled water to remove the remaining SDS. In each wash step the pellet was resuspended in a volume of 2 mL and centrifuged at 320,000  $\times g$  at 25°C for 20 minutes. Finally, after the last centrifugation the pellet was resuspended in 100  $\mu\text{L}$  of double distilled water.



## 2.7 Microscopy

### 2.7.1 Optical microscopy

In order to visualize the *Anabaena* filaments, bright-field microscopy was performed using an Olympus BX60 microscope. The following objectives were used: UplanFI 20x/0.5, UplanFI 40x/0.75Ph2 and UplanFI 100x/1.30 oil. In addition, the microscope had attached a Leica DFC 300FX camera that allowed the samples to be photographed using the PhotoED (Microsoft) software.

### 2.7.2 Immunofluorescence localization of proteins

The strains on which immunolocalization was to be performed were grown in liquid BG11 cultures (supplemented with antibiotics if necessary). When grown, the cells were collected and resuspended in liquid BG11 or BG11<sub>0</sub> medium (without antibiotics), at a concentration of 1 µg Chl/ml, and incubated for 24 or 48 hours. Samples of 1.5 mL of the cultures were centrifuged in an Eppendorf tube. After centrifugation, a small volume of the sample was placed atop a poly-L-lysine pre-coated microscope slide (Sigma) and covered with a 45-µm pore-size Millipore filter. After a few minutes, the filter was removed and the slide was allowed to dry at room temperature, then it was placed in cold 70% ethanol and left at -20°C for 45 minutes. After this time, the slide was removed from the ethanol and allowed to dry at room temperature for 20 minutes. The area of the holder containing the sample was delimited with a marker (liquid blocker) creating a border that allowed us to add various solutions. The sample was washed twice by covering the slide with PBS-T (5.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>HPO<sub>4</sub>, 137 mM NaCl, 2.7 mM KCl [pH 7.4]) and 0.1% (v/v) Tween-20 for 2 minutes at room temperature, and the liquid was removed with a vacuum pump. Afterwards, the slide was treated with blocking buffer (5% milk power in PBS-T) for 30 minutes at room temperature. After removing the blocking buffer, the primary antibody was added.

For His- or Strep II-tagged proteins, the primary antibody was a fluorophore-conjugated commercial antibody against His-tag (Penta-His™ Alexa Fluor® 488 conjugate from Qiagen GmbH, Germany) or against Strep-tag II (StrepMAB-Clasic, specific for Strep-tag II, conjugated to Cromeo™ 488 from IBA GmbH, Germany). The primary antibody was diluted 1:200 and 1:250 for His-tag II and Strep-tag II, respectively, in blocking buffer and incubated in the dark for 45 min at room temperature. After incubation, the slide was washed 3 times with 2 ml of PBS-T for 2 min at room temperature. Finally, 2 drops of FluorSave (Calbiochem) were placed on the sample, a cover was placed on top, and the samples were visualized with a Leica DM6000B fluorescence microscope, using the FITC L5 filter (excitation, band-pass (BP) 480/40 filter; emission, BP 527/30 filter) to detect the fluorescence. The microscope had attached an ORCA-ER camera (Hamamatsu) that allowed the samples to be photographed. Images were analyzed using the ImageJ software (<http://imagej.nih.gov/ij>).

To detect SepJ or FraD, anti-SepJ-CC or anti-FraD antibody was used as the primary antibody at a 1:200 dilution in blocking buffer and incubated for 105 minutes at room temperature. After incubation, the samples were washed 3 times with PBS-T for 2 minutes at room temperature to remove all primary antibody. Secondary anti-rabbit antibody conjugated to fluorescein isothiocyanate (FITC) (Sigma) was then added at a 1:500 dilution in PBS-T buffer and incubated



in the dark for 45 min at room temperature. After this step, the samples were washed twice with PBS-T for 2 min at room temperature, FluorSave was added and the samples were visualized as indicated in the previous paragraph.

### 2.7.3 Confocal microscopy and FRAP experiments

The Fluorescence Recovery After Photobleaching (FRAP) experiments were performed with two different fluorescent markers, calcein and 5-carboxyfluorescein (5-CF). In both cases, the filaments were grown in BG11 medium supplemented with antibiotics if necessary, collected by centrifugation, washed and re-inoculated in BG11 or BG11<sub>0</sub> medium (without antibiotics) at a concentration of 1 µg Chl/mL. The cell suspensions were then incubated under culture conditions for 24 or 48 hours. To mark the filaments, 1 mL of culture was supplemented with 20 µL of 1 mg/mL calcein-AM solution or 5 µL of 1 mg/mL 5-CF solution. In the case of calcein the filaments were incubated for 90 min and in the case of 5-CF for 60 min; both incubations were carried out in darkness at 30°C with shaking. After incubation, the filaments were washed twice with BG11 or BG11<sub>0</sub> medium (depending on the origin of the filaments), and a few drops of the filament suspension were placed on a plate of solid BG11 or BG11<sub>0</sub> medium, respectively. When the drops had dried, the piece of agar was cut out and placed with on cover slip in an especial holder in which the temperature of the sample was controlled at 30°C. Images were taken on a Leica TSC SP2 confocal microscope using the immersion objective HCX PLAM-APO 63x 1.4 NA. An argon laser was used to excite the samples at 488 nm. Fluorescent emission was monitored by collection across a window of 500-520 nm using a 150-µm pinhole. Once we had correctly focused a filament that had a homogeneous fluorescence, a cell (chosen randomly) was bleached by increasing the power of the laser by a factor of 10 using the microscope in Y-scanning mode. After bleaching, the laser intensity was reduced again, the microscope was put back into XY-imaging mode and a sequence of 10 photos was taken at 1 to 2 s intervals.

With all the photographs obtained in the FRAP analysis, a stack of images was created to quantify the fluorescence of the cell that had been bleached using the ImageJ program. The recovery rate constant  $R$  was calculated from the data using the formula  $C_B = C_0 + C_R (1 - e^{-2Rt})$ , in which  $C_B$  is the fluorescence of the bleached cell,  $C_0$  is the fluorescence immediately after bleach and tending towards  $(C_0 + C_R)$  after fluorescence recovery,  $t$  is time, and  $R$  is the recovery rate constant due to transfer of the marker from one neighbor cell (Merino-Puerto et al., 2011b).

### 2.7.4 Transmission electron microscopy

The murein sacculi whose isolation is described in section 2.6 were deposited on a grid that had previously been covered with a formvar / carbon film and stained with 1% (w/v) uranyl acetate. After staining, they were visualized on a ZEISS LIBRA 120 PLUS electron microscope at 120 kV.

## 2.8 Other parameters measured in *Anabaena*

### 2.8.1 Determination of chlorophyll *a*

The concentration of chlorophyll *a* (Chl) was determined by a protocol derived from that described by Mackinney (1941). A volume of 100 µL was taken from the culture and mixed with



900  $\mu$ L of methanol in an Eppendorf tube, the mix was stirred in a vortex for 1 min and centrifuged for 1 min. The supernatant was transferred to a plastic spectrophotometer cuvette (1-cm light pass), and its absorbance was measured at 665 nm. The Chl concentration was calculated using the formula  $[\text{Chl}] = \text{Abs} \cdot \text{Fd} \cdot \epsilon$ ; where Abs is the absorbance of the sample, Fd is the dilution factor, which in this case is 10, and  $\epsilon$  is the extinction coefficient, 74.46 mL/(mg·cm).

## 2.8.2 Growth tests

### 2.8.2.1 Determination of the growth rate constant

To calculate the growth rate constant, the cultures were grown in 25 mL BG11 liquid medium (supplemented with antibiotics if necessary). Once grown, the cultures were collected by centrifugation, washed with liquid BG11<sub>0</sub> medium and their Chl concentration was determined (see 2.8.1). Flasks containing 50 mL of liquid BG11 or BG11<sub>0</sub> medium (without antibiotics) were re-inoculated at a concentration of 0.1  $\mu$ g Chl/mL and incubated at 30°C with shaking. Samples of 1 mL were taken every 12 hours, approximately, always ensuring that the sample was well homogenized with the pipette; the first sample was taken immediately after re-inoculation. The OD of these samples was measured in a spectrometer at a wavelength of 750 nm. The OD values were plotted, allowing us to calculate the growth rate constant,  $\mu = (\ln 2)/t_d$ , where  $t_d$  is the doubling time.

### 2.8.2.2 Growth test on solid medium

The strains were grown in liquid BG11 medium (with antibiotics if necessary), and then collected by centrifugation, washed with BG11<sub>0</sub> medium and their Chl concentration determined. A suspension was prepared in BG11<sub>0</sub> medium at 1  $\mu$ g Chl/mL, from which four serial dilutions were prepared (0.5, 0.25, 0.125 and 0.0625  $\mu$ g Chl/mL, respectively). From each dilution, 8- $\mu$ L drops were placed on Petri dishes containing solid BG11 or BG11<sub>0</sub> medium without antibiotics (so each drop had an initial amount of 8, 4, 2, 1 and 0.5 ng Chl), and the plates were incubated for 10 to 14 days under standard growth conditions (section 2.1.2.2).

### 2.8.3 Nitrogenase activity

The determination of nitrogenase activity was carried out by a method based on that of Stewart *et al.* (1967), which involves reduction of acetylene by this enzyme. The strains were grown in 25 mL of liquid BG11 medium, collected by centrifugation, washed with BG11<sub>0</sub> medium, re-inoculated in 25 mL of liquid BG11<sub>0</sub> medium at 1  $\mu$ g Chl/mL, and incubated for 48 hours under standard growth conditions with shaking. After incubation, the cells were collected by centrifugation and a cell suspension with 10  $\mu$ g Chl/mL was prepared in a final volume of 2 mL and placed in small flasks. To determine the nitrogenase activity under oxic conditions, the flasks with the 2-mL suspension were sealed with a rubber stopper and incubated for 30 min at 30°C with shaking. Then, 2 mL of acetylene was injected, the gas in the flask was mixed, and 1-mL gas samples were taken at different times up to 90 minutes. During the assay, the flasks were kept under standard conditions of growth at 30°C with light and shaking. Gas samples were analyzed by gas chromatography to determine the production of ethylene from acetylene. The concentration of Chl of each sample was determined again at the end of the assay, and together



with ethylene production, the nitrogenase activity was calculated and expressed as nmol ethylene/ $\mu\text{g}$  Chl.

To carry out the determination of nitrogenase activity under anoxic conditions, the protocol described above was modified as follows. The 2-mL cell suspension (10  $\mu\text{g}$  Chl/mL) was supplemented with DMCU (3-(3,4-dichlorophenyl)-1,1-dimethylurea) to a final concentration of 4  $\mu\text{g}$ /mL. The flask was then sealed with the rubber stopper, bubbled with argon for 4 minutes, and incubated for one hour under standard growth conditions. Then, acetylene (2 mL) was injected and the protocol continued as described in the previous paragraph.

#### 2.8.4 Filament length

In order to quantify the number of cells that make up each filament, the strains were grown in liquid BG11 medium (with antibiotics if necessary), collected by centrifugation, washed with BG11<sub>0</sub> medium and re-inoculated in the different liquid media of study (in this case BG11 and BG11<sub>0</sub>, without antibiotics) at a concentration of 1  $\mu\text{g}$  of Chl/mL. After 24, 48 and 72 hours of incubation under standard growth conditions, 500  $\mu\text{L}$  of sample was taken with great care to prevent fragmentation of the filaments, and several drops were deposited on a Petri dish with solid BG11 or BG11<sub>0</sub> medium, respectively. Once the drops dried, they were cut from the agar and deposited on a slide for visualization and photography under an optical microscope (see 2.7.1). The cells that formed each filament were visualized in the micrographs and counted using the ImageJ software.

### 2.9 *In silico* analysis

#### 2.9.1 Analysis of DNA sequences

The search for sequences necessary to prepare constructs with the different genes of *Anabaena*, *Richelia intracellularis* or *Calothrix rhizosoleniae*, as well as for the study of their genomic context, was carried out using the Integrated Microbial Genomes (IMG) page of the Joint Genome Institute (<https://img.jgi.doe.gov>). On the other hand, the design of all plasmids and the study of the sequencing results were carried out using the *SerialCloner* software.

#### 2.9.2 Analysis of protein sequence and structure

Both the IMG website and the *SerialCloner* software were also used for protein sequence searching and data analysis related to proteins. The program TMHMM (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) was used to predict transmembrane segments in a protein sequence, and Protter –based on Phobius predictions (<http://phobius.sbc.su.se>)– was used for their representation. To predict the subcellular localization of proteins, the PsortB program (<https://www.psort.org/psortb/>) was used. To predict signal peptides, the SignalP5 program (<http://www.cbs.dtu.dk/services/SignalP/>) was used. To predict coiled-coil domains, the Coils program ([https://embnet.vital-it.ch/software/COILS\\_form.html](https://embnet.vital-it.ch/software/COILS_form.html)) was used. To predict the three-dimensional structure of some proteins, Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>) was used. To align the sequences of proteins, Clustal W (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) was used. Finally, to perform phylogenetic trees, Phylogeny (<http://www.phylogeny.fr>) was used.





### 3. Results & Discussion

---

#### Chapter 1

#### Intercellular communication in *Anabaena*

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e2000022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular





ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

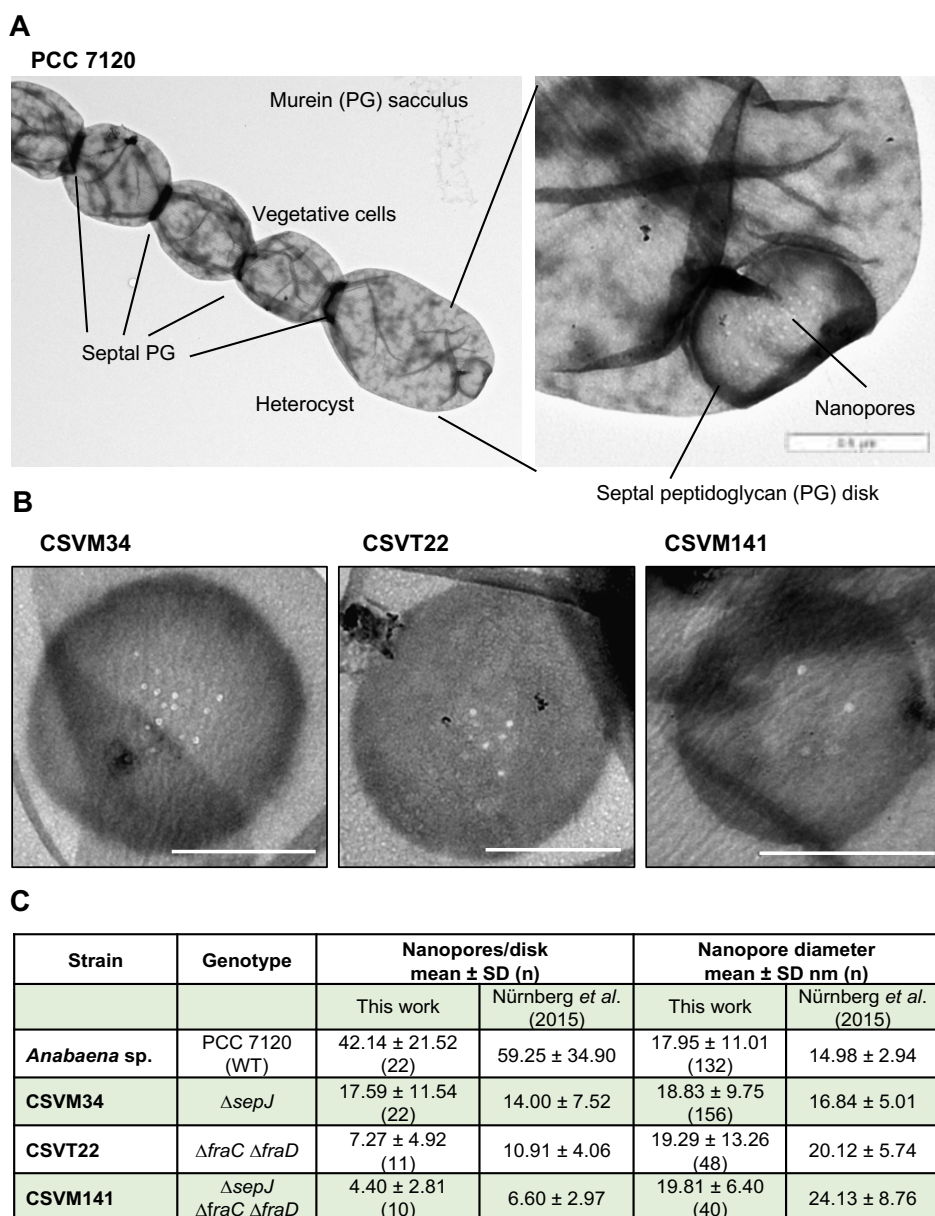
### 3.1 Intercellular communication in *Anabaena*

*Anabaena* belongs to the group of filamentous heterocystous cyanobacteria, and heterocyst differentiation requires the transfer of nutrients and regulators between cells within the filament (Herrero *et al.*, 2016). As described in the Introduction, intercellular communication is likely carried out by the septal junctions that may traverse the septal peptidoglycan (PG) through holes called nanopores (Mullineaux *et al.*, 2008; Mariscal *et al.*, 2011; Nürnberg *et al.*, 2015). SepJ, FraC and FraD are some of the proteins that have been identified in the intercellular septa and whose lack affects filament integrity and the intercellular transfer of fluorescent markers (Flores *et al.*, 2007; Mullineaux *et al.*, 2008; Merino-Puerto *et al.*, 2010; Nürnberg *et al.*, 2015). With the aim of establishing a possible relationship between intercellular molecular transfer and nanopores, we have performed further analysis and reevaluated data obtained for nanopores and calcein transfer in filaments of *Anabaena* and mutants of these septal proteins. Recently a new protein called SepI has been described that also appears to be related to intercellular communication (Springstein *et al.*, 2020). SepI is located to the Z-ring and in the septal regions of *Anabaena*, and the phenotypic characterization of a *sepI* mutant shows shorter filaments than in the wild type. These observations together with the fact that computer predictions show that SepI has a coiled-coil and a linker domain, strongly resembling SepJ (Springstein *et al.*, 2020), makes us to include SepI in this study. Finally, some specific mutants of the permease domain of SepJ (strain CSVM90, producing SepJ $_{\Delta 463-748}$ ; CSFR12, SepJ $_{\Delta 498-507}$ ; CSFR14, SepJ $_{E663A}$ ) were also evaluated to better understand the role of SepJ in intercellular communication (Ramos-León *et al.*, 2018).

#### 3.1.1 Nanopores

Nanopores appear to be a fundamental structure for intercellular molecular exchange. Here we performed a new study of the number of nanopores of *Anabaena* wild type and mutant strains CSVM34 ( $\Delta sepJ$ ), CSCT22 ( $\Delta fraC \Delta fraD$ ) and CSVM141 ( $\Delta sepJ \Delta fraC \Delta fraD$ ), and the results were compared with those of Nürnberg *et al.* (2015). Filaments of these strains were grown in liquid BG11 medium (containing nitrate as nitrogen source) to about 3-4  $\mu\text{g Chl}/\text{mL}$  and, once collected, they were subjected to sonication to fragment them down to 1-3 cells/filament, since otherwise most septal disks would be hidden between adjacent cells. After sonication, each sample was repeatedly boiled in the presence of different concentrations of detergent (SDS) and treated with  $\alpha$ -chymotrypsin to remove all protein, resulting in isolated murein sacculi (Lehner *et al.*, 2011). Samples were placed on a grid that was impregnated with a formvar/carbon film and stained with uranyl acetate for visualization by transmission electron microscopy (TEM). Wild-type strain was also grown in liquid BG11<sub>0</sub> medium (which lacks combined nitrogen) for 48 hours. In this case, filaments were fragmented by gentler sonication. Peptidoglycan sacculi were then isolated and imaged as described above. In Fig. 14A, a sacculus that corresponds to several cell units including a terminal heterocyst is shown. Electro-dense septal PG corresponds to the juxtaposition of the PG layer of two adjacent cells. The fact that those cells remain bound after isolation implies that the PG of adjacent cells are chemically linked. In the heterocyst pole, a septal PG disk is clearly visualized showing nanopores. Representative micrographs of septal disks are shown in Fig. 14B.



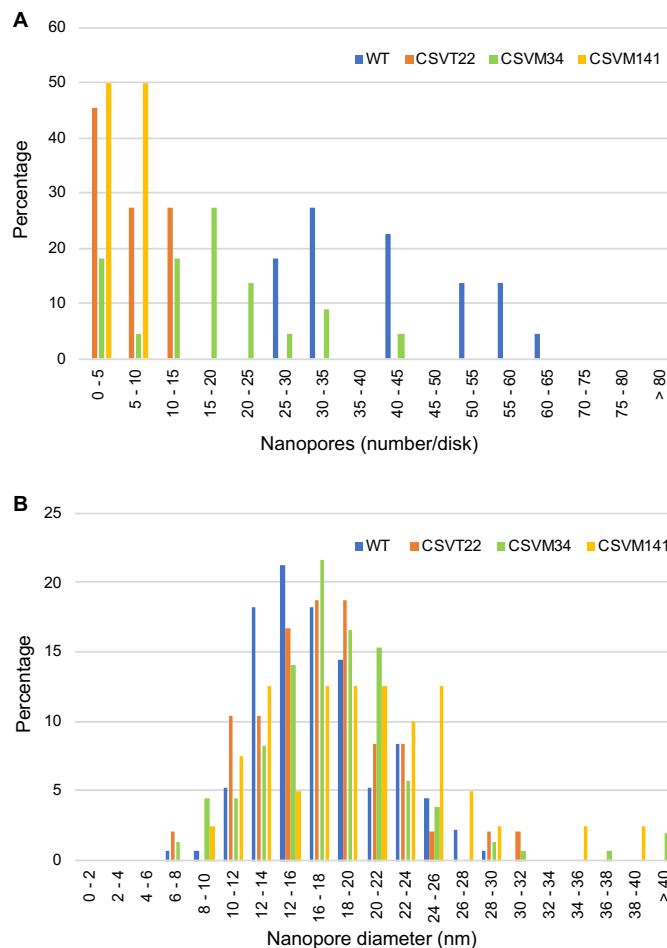


**Fig. 14. Nanopores in septal peptidoglycan disks of *Anabaena*.** (A) Peptidoglycan (murein) sacculus of part of a filament of *Anabaena* grown in BG11<sub>0</sub> medium. The PG was isolated as described in section 2.6 of Materials & Methods, stained with uranyl acetate and visualized by TEM. Note the thickness of septal PG, the presence of one heterocyst, and a septal PG disk with nanopores. (B) Nanopores in disks from *Anabaena* mutants CSVM34, CSVT22 and CSVM141. (C) Genotypes of the strains analyzed and mean and standard deviation of nanopore number and diameter. n, number of septal disks (nanopore counting) or nanopores (diameter) analyzed. Size bars, 500 nm.

Figure 15A shows the frequency distribution of nanopore numbers, showing decreased numbers in strains CSVM34, CSVT22 and CSVM141. This trend of decrease in nanopore numbers is



similar to that described by Nürnberg *et al.* (2015) except for CSVM34 that had about 42% of the nanopore number in the wild type, in contrast to 24% described previously (Fig. 14C). Figure 15B shows the frequency distribution of nanopore diameter for the wild type and the three mutants. In contrast to Nürnberg *et al.* (2015), in the current analysis we did not find any significant difference in nanopore diameters between any of mutants and the wild type (Fig. 14C). Thus, analysis the results with the Student's *t* test (each mutant vs. WT) showed  $P > 0.15$  values in all three cases.



**Fig. 15. Frequency distribution of nanopore numbers and diameters in septal PG disks.** Samples studied were from wild-type *Anabaena* and  $\Delta sepJ$  (strain CSVM34),  $\Delta fraC \Delta fraD$  (CSVT22) and  $\Delta sepJ \Delta fraC \Delta fraD$  (CSVM141) mutants. (A) The number of nanopores is organized in groups of 5 (from 0 to 5; from 5.01 to 10; etc.). (B) Nanopore diameter is organized in groups of 2 nm (from 0 to 2; from 2.01 to 4; etc.)

Given the central role of SepJ in the construction of intercellular septa in *Anabaena*, we investigated *Anabaena* mutants that produce a SepJ protein that lacks most of its permease domain (strain CSVM90) or have a short internal deletion (CSFR12) or an amino acid substitution (CSFR14). Strains CSFR12 and CSFR14 were constructed by the reconstruction of the *sepJ* gene



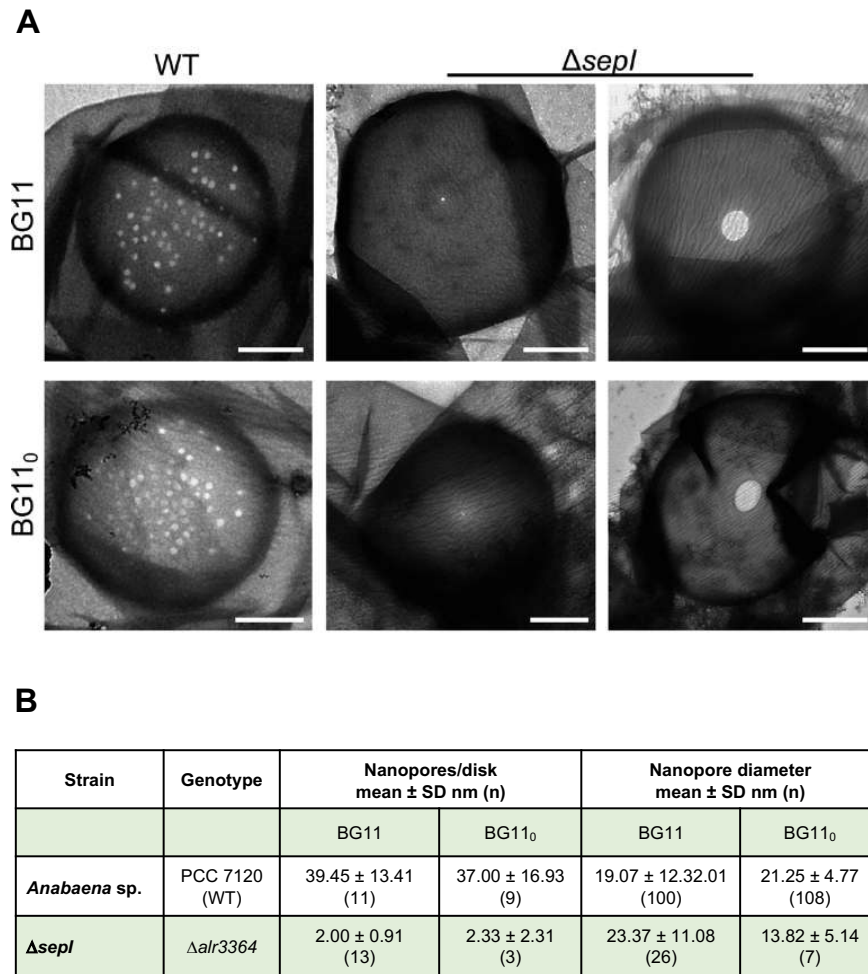
in strain the CSVM90, lacking most of the permease domain, and we have compared these mutants to control strain CSFR11, which produces a reconstituted wild-type SepJ protein. Nanopore parameters are summarized in Table 7. Nanopore number in strain CSVM90 was similar to that of the *sepJ* deletion mutant CSVM34 (see Fig. 14C), which is consistent with the fact that the SepJ protein produced in strain CSVM90 has null activity (Ramos-León *et al.*, 2018); however, nanopore diameter was somewhat smaller in CSVM90 as compared to CSVM34 (Fig. 14C and Table 7). Strains CSFR12 and CSFR14 showed a 40% decrease and a normal number of nanopores, respectively, which is consistent with the observation that they have somewhat altered and normal SepJ-related phenotypes, respectively (Ramos-León *et al.*, 2018).

Strain	Genotype	Nanopores/disk mean $\pm$ SD nm (n)	Nanopore diameter mean $\pm$ SD nm (n)
CSVM90	<i>sepJ</i> SepJ <sub>D463-748</sub>	16.3 $\pm$ 9.5 (15)	13.4 $\pm$ 2.9 (132)
CSFR11*	Wild-type <i>sepJ</i>	44.8 $\pm$ 6.3 (9)	15.4 $\pm$ 3.5 (57)
CSFR12	<i>sepJ</i> (SepJ <sub>D498-507</sub> )	26.4 $\pm$ 11.5 (9)	17.6 $\pm$ 3.1 (128)
CSFR14	<i>sepJ</i> (SepJ <sub>E663A</sub> )	41.6 $\pm$ 11.5 (11)	17.4 $\pm$ 3.3 (175)

**Table 7. Nanopores in septal peptidoglycan disks of selected *sepJ* mutants.** The table shows the genotypes of the strains analyzed and the mean and standard deviation of nanopore number and diameter. n, number of septal disks (nanopore counting) or nanopores (diameter) analyzed. These data were included in Ramos-León *et al.* (2018).

SepI is a recently identified septal protein that interacts with divisome proteins and may also interact with SepJ (Springstein *et al.*, 2020). We observed that the  $\Delta$ *sepI* mutant shows two different types of perforations in the septal PG (Fig. 16A), one type with a tiny diameter (in “sealed disks”) and another with a very large diameter (defining “unsealed disks”). Although we cannot exclude that unsealed disks correspond to not yet fully divided cells and would eventually become normal septal disks, it should be noted that we rarely observed such large perforations in the wild type. The septal disks with those large perforations were not included in the measurement of nanopore numbers and diameter. The number of nanopores per septal PG disk was strongly reduced in the  $\Delta$ *sepI* mutant and we never observed more than 4 nanopores in any of the isolated sacculi, regardless of the availability of combined nitrogen. This observation should be compared to about 39 and 37 nanopores per disk in the wild type grown under the same conditions, BG11 or BG11<sub>0</sub> medium respectively (Fig. 16B). Nanopore diameter was unaffected in the  $\Delta$ *sepI* mutant in BG11 medium but significantly reduced by about 35 % in the  $\Delta$ *sepI* mutant under nitrogen deprivation (BG11<sub>0</sub> medium; Fig. 16B).





**Fig. 16. Nanopores of the *Anabaena sepl* mutant.** (A) TEM micrographs of peptidoglycan isolated from *Anabaena* wild-type strain and  $\Delta sepl$  mutant from BG11 medium (top) and BG11<sub>0</sub> medium (bottom). (B) Genotypes of the strains analyzed and mean and standard deviation of nanopore number and diameter. n, number of septal disks (nanopore counting) or nanopores (diameter) analyzed. Size bars, 0.5 μm.

### 3.1.2 Intercellular calcein transfer

Intercellular molecular transfer in cyanobacteria has been studied using FRAP (Fluorescence Recovery After Photobleaching) analysis, which basically consists in loading the cyanobacterial filaments with a fluorescent marker, bleaching a specific cell of the filament with the confocal microscope laser, and measuring the fluorescence as it recovers over time (Mullineaux *et al.*, 2008). This method (explained in detail in Materials & Methods, section 2.7.3) allows us to calculate the exchange coefficient ( $E$ ) and the recovery rate constant ( $R$ ) (Mullineaux *et al.*, 2008; Merino-Puerto *et al.*, 2010). The results of these parameters have generally been presented as the mean and standard deviation of the data, assuming that the data fit a normal distribution. However,





it has recently been observed that some *Anabaena* mutants present a substantial number of cells that do not recover after bleaching, presenting an  $R$  value lower than  $0.01 \text{ s}^{-1}$ ; these cells have been defined as non-communicating cells (Flores *et al.*, 2018). The presence of non-communicating cells may not be restricted to some *Anabaena* mutants, but it may happen in the wild type as well.

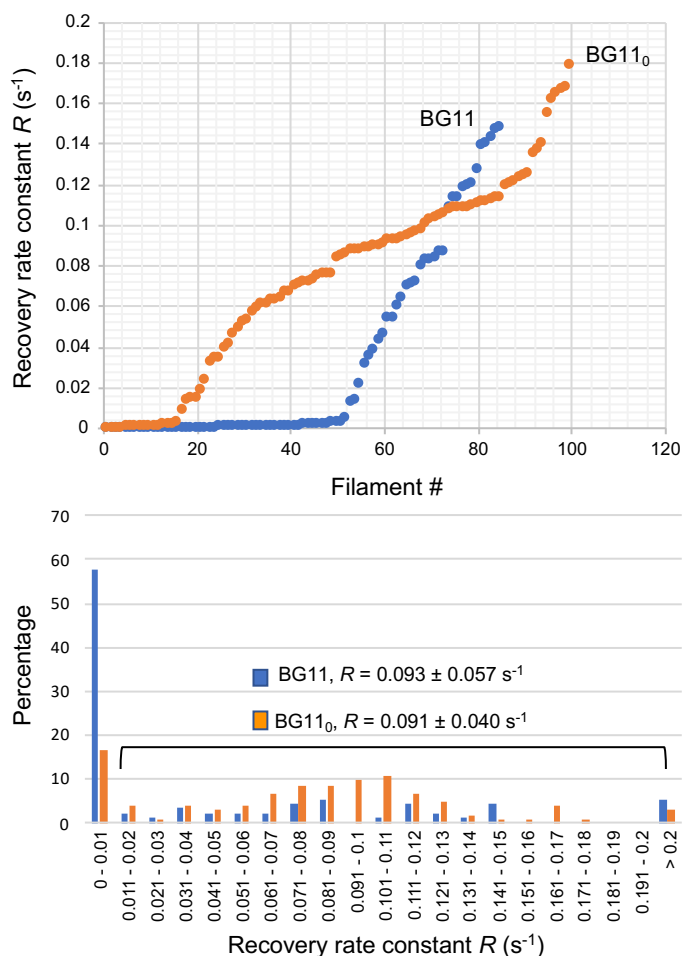
Therefore, we decided to reevaluate the transfer of calcein in different *Anabaena* strains. Those selected for this study were the wild type and mutants CSVM34 ( $\Delta sepJ$ ), CSVT22 ( $\Delta fraC \Delta fraD$ ) and CSVM141 ( $\Delta sepJ \Delta fraC \Delta fraD$ ). All these strains were grown in liquid BG11 medium, washed and re-inoculated at a concentration of  $1 \mu\text{g Chl/mL}$  in flasks containing BG11 or BG11<sub>0</sub> medium respectively, and then incubated for 48 hours under culture conditions. By re-inoculation of the stains, we induce heterocyst differentiation in BG11<sub>0</sub> medium (without combined nitrogen) and ensure the same growth rate in all strains growing in BG11 medium. After incubation, a sample (1 mL) from each strain and condition was taken and incubated with calcein as described in Materials & Methods (section 2.7.3). Then, drops of each treated sample were placed onto plates of BG11 or BG11<sub>0</sub> medium depending on the treatment that the sample had received, and small pieces of agar carrying the filaments were prepared to be visualized under the confocal microscope. The samples were excited at a wavelength of 488 nm to visualize the filaments that had been loaded with calcein, and in homogeneously marked filaments a cell was bleached by increasing the intensity of the laser. After the bleaching, 10 photographs were taken with an interval between photographs of one to two seconds, and the recovery rate constant ( $R$ ) was calculated (Nieves-Mori3n *et al.*, 2017a).

The BG11-grown cultures of the wild type had a substantial number of non-communicating cells, accounting for about 58% of the total number of cells subjected to analysis (Fig. 17). On the other hand, a few cells showed extremely high  $R$  values, for example cells with  $R$  values  $> 0.3 \text{ s}^{-1}$ . Because these high  $R$  values are similar to the recovery times expected for free diffusion (Rudolf *et al.*, 2015), in these cells the cytoplasm might be not fully divided from that of neighboring cells after cell division, and they were not included in further analysis. In contrast to BG11-grown cells, the cultures incubated in BG11<sub>0</sub> medium had less non-communicating cells (about 16%). The estimate of non-communicating cells was also performed individually in each experimental replicate, observing the presence of non-communicating cells in most cultures, especially in BG11 medium (Table 8).

When the  $R$  values of cells grown in BG11 medium or incubated in BG11<sub>0</sub> medium were calculated excluding non-communication cells ( $R < 0.01 \text{ s}^{-1}$ ) and cells that might have not yet divided (those showing  $R > 0.3 \text{ s}^{-1}$ ), very similar recovery constants are obtained (Fig. 17, *lower panel*). The increase in the number of communicating cells in cultures incubated in BG11<sub>0</sub> medium can represent a significant aspect of the acclimation of *Anabaena* to diazotrophic growth. It should be noted, nonetheless, that the fraction of non-communicating cells observed in filaments grown in BG11 or incubated in BG11<sub>0</sub> medium may depend on other growth parameters (temperature, light intensity, CO<sub>2</sub> supply), which could explain why they have not been noticed previously.







**Fig. 17. Intercellular transfer of calcein between vegetative cells of *Anabaena*.** Filaments grown in BG11 medium (blue) or incubated for 48 h in BG11<sub>0</sub> medium (orange). (*Upper panel*) Filaments/cells analyzed by FRAP are organized by increasing value of recovery rate constant. 90 filaments from BG11 medium and 103 filaments from BG11<sub>0</sub> medium were analyzed. Five filaments from BG11 medium giving  $R$  values of 0.247, 0.266, 0.312, 0.436 and 0.505  $s^{-1}$ , respectively, and three filaments from BG11<sub>0</sub> medium giving  $R$  values of 0.227, 0.327 and 0.861  $s^{-1}$ , respectively, are not included to facilitate the visibility of the graph.  $R$  values  $> 0.3 s^{-1}$  were not considered in further analysis because they may correspond to not fully divided cells (see main text). (*Lower panel*) Data from the upper panel organized to show frequency distributions ( $R$  data organized in groups of 0.01  $s^{-1}$ ). The mean and SD are shown for  $R$  values between 0.010 and 0.300  $s^{-1}$ . The data used in this analysis combine data obtained for the wild-type controls in the studies of the  $\Delta sepI$  mutant (Fig. 21 below) and of the *hglK* gene (section 3.2 of this Thesis).

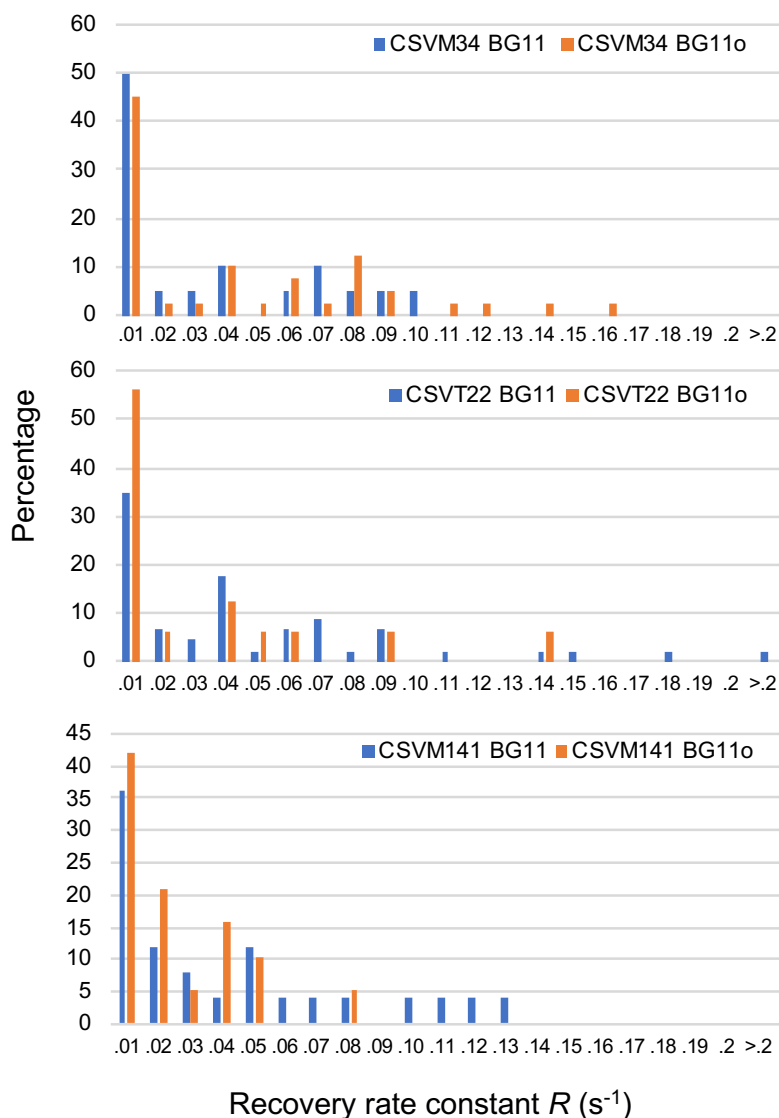


Medium	Percentage of non-communicating cells ( $R < 0.01 \text{ s}^{-1}$ )							
	Replicate							
	1	2	3	4	5	6	7	8
BG11	79	91	45	70	54	30	75	-
BG11 <sub>0</sub>	31	29	9	27	0	0	0	17

**Table 8. Analysis of non-communicating cells in each replicate.** The table shows the percentage of non-communicating cells (recovery rate constant  $R < 0.01 \text{ s}^{-1}$ ) in filaments of *Anabaena* that had been grown in a medium with combined nitrogen (BG11) or in a medium in absence of nitrogen (BG11<sub>0</sub>). Each value corresponds to the percentage of non-communicating cells in a data set from an individual culture after performing the FRAP assay using calcein as a fluorescent marker.

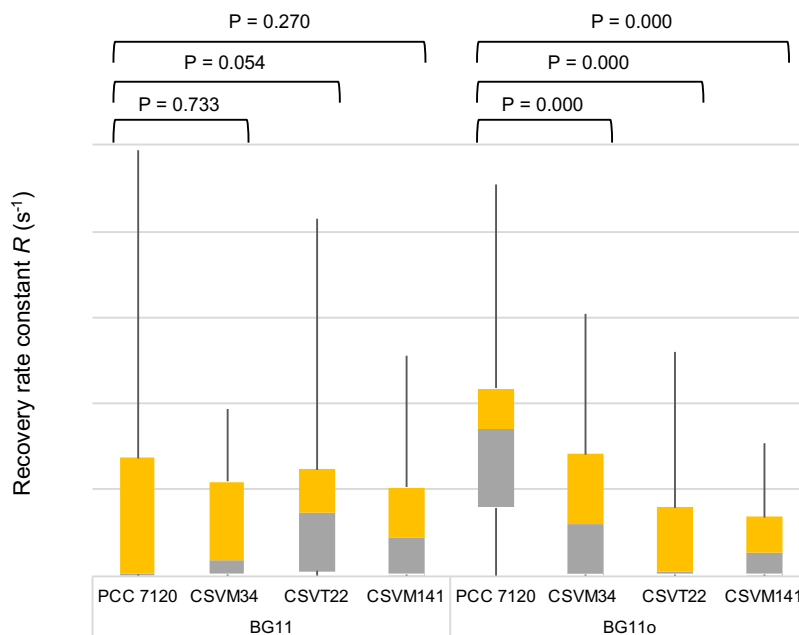
We then addressed calcein FRAP in the deletion mutants of septal protein-encoding genes *sepJ*, *fraC* and *fraD*. As shown in Fig. 18, non-communicating cells were observed in all the mutants not only in BG11-grown filaments but also in filaments that had been incubated under nitrogen deprivation (BG11<sub>0</sub> medium). As the results do not follow a parametric distribution, to compare these data with those for the wild type, we represented the data as box plots and used the Mann-Whitney U test to assess whether the differences observed are significant or not (Fig. 19). This analysis shows that in BG11 medium there were hardly differences between any of the mutants and the wild type, whereas in BG11<sub>0</sub> medium the three mutants were significantly different from the wild type. Thus, perhaps because of the high number of non-communicating cells found in the wild type, results differ from those previously reported for the mutants grown in BG11 medium (Mullineaux et al., 2008; Merino-Puerto et al., 2010). Nonetheless, our current results confirm that the *sepJ* and *fraC-fraD* mutants are impaired in calcein transfer as assessed in filaments that had been incubated in the absence of combined nitrogen.





**Fig. 18. Frequency distribution of calcein transfer values in different *Anabaena* septal protein mutants.** Intercellular transfer of calcein between vegetative cells of *Anabaena* mutant strains CSV22 ( $\Delta fraC \Delta fraD$ ), CSVM34 ( $\Delta sepJ$ ) and CSVM141 ( $\Delta fraC \Delta fraD \Delta sepJ$ ) grown in BG11 medium (left bars, blue) or incubated for 48 h in BG11<sub>o</sub> medium (right bars, orange). Filaments/cells analyzed by FRAP are organized in groups of  $R = 0.01 s^{-1}$ .

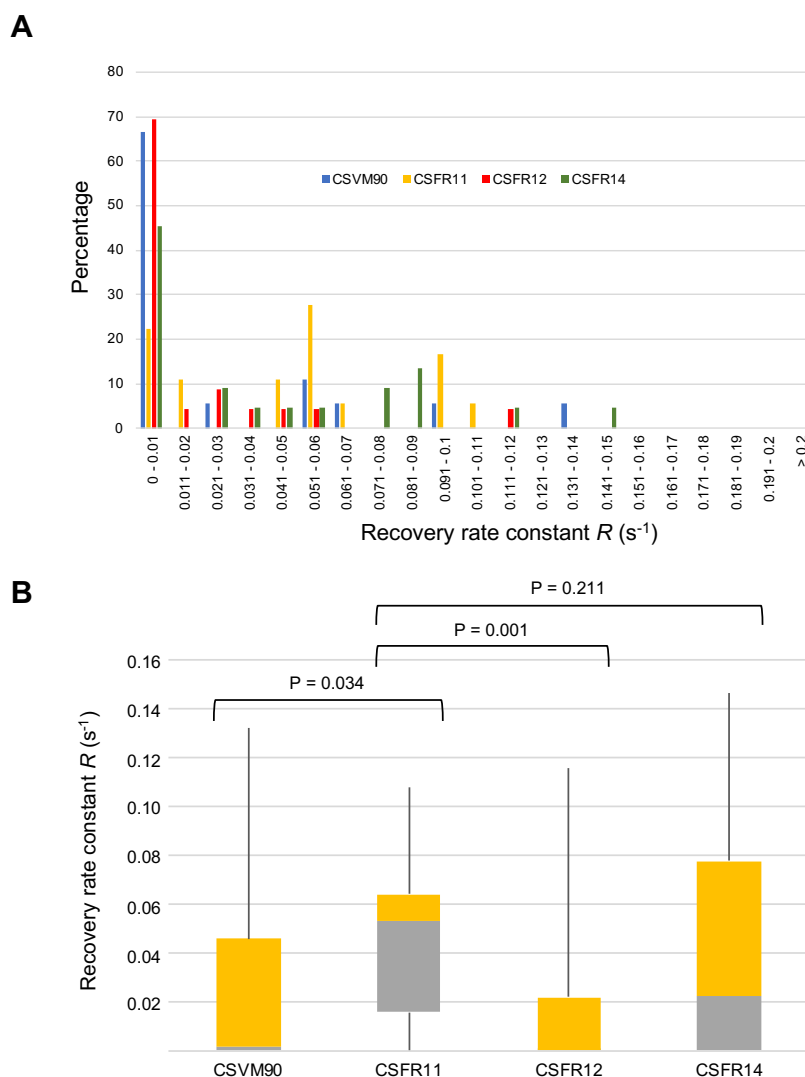




**Fig. 19. Box-plot representation and statistical analysis of intercellular transfer of calcein between vegetative cells.** The graph shows the results of tests performed in *Anabaena* and mutant strains CSVT22 ( $\Delta fraC \Delta fraD$ ), CSVM34 ( $\Delta sepJ$ ) and CSVM141 ( $\Delta fraC \Delta fraD \Delta sepJ$ ) grown in BG11 medium or incubated for 48 h in BG11<sub>0</sub> medium. Grey, quartile group 2 (from Q1 to median); yellow, quartile group 3 (from median to Q3). The results of the non-parametric Mann-Whitney U test are included on the top of the figure (P values).

To address the specific *sepJ* mutants for which nanopores are presented above, we have reevaluated data previously obtained in this laboratory (F. Ramos-León, Ph. Thesis, 2017; summarized in Ramos-León *et al.*, 2018). BG11-grown filaments of strain CSVM90 had a substantial number of non-communicating cells, as well as CSFR12 and CSFR14 (Fig. 20A). In contrast, strain CSFR11, which expresses a reconstituted wild-type SepJ protein, showed a lower number of non-communication cells. Figure 20B shows box plots of these data and the results of Mann-Whitney U tests, which indicate that the recovery rate constant ( $R$ ) was significantly lower in strains CSVM90 and CSVM12 than in strain CSFR11. On the other hand, there was no significant difference between the CSFR14 mutant and the strain that has the reconstituted *sepJ* gene (CSFR11).



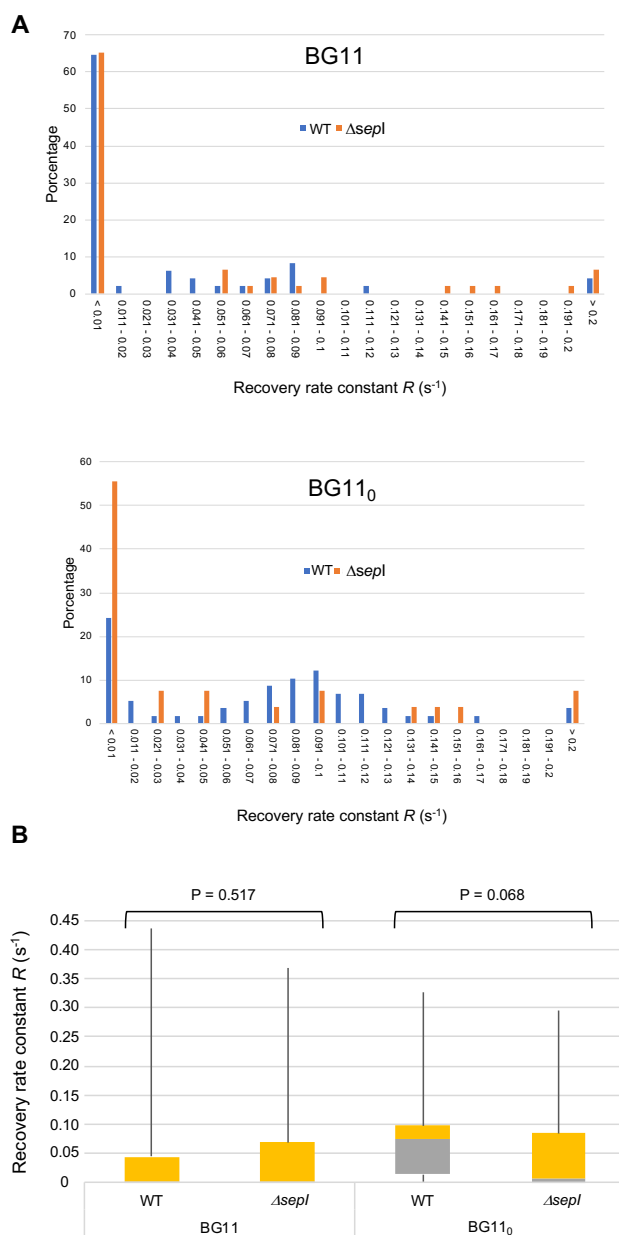


**Fig. 20. Analysis of intercellular calcein transfer between vegetative cells in different *sepJ* mutants.** (A) Frequency distributions of the recovery rate constant ( $R$ , data organized in groups of  $0.01 s^{-1}$ ) in the different *sepJ* mutants grown in BG11 medium. The number of recorded bleached cells were: 18 for CSFR11 (wild-type *SepJ*) and CSVM90 (*SepJ* <sub>$\Delta$ 463-748</sub>); 23 for CSFR12 (*SepJ* <sub>$\Delta$ 498-507</sub>); and 22 for CSFR14 (*SepJ*<sub>E663A</sub>). (B) Box-plot representation of the data in panel A and its statistical analysis by means of the Mann-Whitney U test in which the mutant that has the *sepJ* gene reconstructed (CSFR11) was compared with the rest of the mutants.

As shown above, the *sepI* mutant presents two types of septal disks that we have identified as sealed and unsealed. This observation made it of especial interest to study intercellular molecular transfer in this mutant. Figure 21 shows the data for calcein transfer obtained with the *sepI* mutant grown in BG11 or BG11<sub>0</sub> medium. Cultures from both media contained a substantial number of non-communicating cells. However, whereas the percentage on non-communicating cells was not very different from that of the wild type in filaments from BG11 media (65% for  $\Delta$ *sepI*, 64% for the WT), it was notably increased as compared to the wild type in filaments from BG11<sub>0</sub> media



(54% for  $\Delta sepI$ , 24% for the WT). It would be interesting to investigate whether the non-communicating cells in BG11<sub>0</sub> medium correspond to cells flanked by sealed septal PG disks.



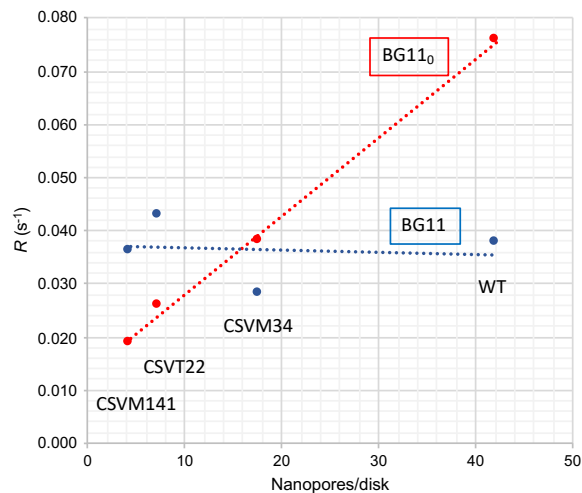
**Fig. 21. Decreased fluorescence recovery rate and presence of non-communicating cells in the  $\Delta sepI$  mutant.** (A) Distribution of cells of *Anabaena* WT (blue) and the  $\Delta sepI$  mutant (orange) showing different  $R$  values for calcein transfer after growth in BG11 medium (top) or 48 h after transfer to BG11<sub>0</sub> medium (bottom). The percentage of cells within  $R$  value intervals of  $0.01 s^{-1}$  (i.e.,  $<0.01$ ; from  $0.011$  to  $0.02$ ; from  $0.021$  to  $0.03$ ; etc.;  $>0.2$ ) is represented. (B) Recovery rate constants of calcein FRAP in vegetative cells of *Anabaena* and the  $\Delta sepI$  mutant grown in BG11 medium and 48 h after transfer to BG11<sub>0</sub> medium. Number of recorded bleached cells (n): WT in BG11 medium, 48;  $\Delta sepI$  in BG11 medium, 46; WT in BG11<sub>0</sub> medium, 57;  $\Delta sepI$  in BG11<sub>0</sub> medium, 27. The significance of the difference between mutant and wild type was assessed by the Mann-Whitney U test (P values indicated on the top). The wild-type values are a subset of those presented in Fig. 17, specifically those run in parallel to the *sepI* mutant.



### 3.1.3 Nanopores and intercellular calcein transfer

Once the number of nanopores and the transfer of calcein were studied, the initial objective of trying to establish a relationship between those two parameters could be addressed. For that, we compared the values of these parameters in *Anabaena* (wild type) and in the deletion mutants of *sepJ*, *fraC* and *fraD*. It is important to note that to calculate the mean values in calcein transfer, those values whose  $R > 0.3 \text{ s}^{-1}$  were excluded for the reasons explained above. Intercellular calcein transfer and nanopore numbers showed an excellent correlation in filaments that had been incubated without combined nitrogen (BG11<sub>0</sub> medium, correlation coefficient = 0.9985) but not in filaments grown in BG11 medium (correlation coefficient = -0.1047) (Fig. 22). Hence, the presence of many non-communicating cells in filaments grown with combined nitrogen impairs the correlation. Nonetheless, because of the excellent correlation found when only few cells are closed, our results support the idea that nanopores are the structures that accommodate septal junctions that are active in calcein transfer.

Strain	Nanopores/disk (n)	Calcein transfer between cells	
		BG11 ( $R, \text{ s}^{-1}$ )	BG11 <sub>0</sub> ( $R, \text{ s}^{-1}$ )
<i>Anabaena</i> sp.	42.14	0.038	0.076
CSVM34	17.59	0.028	0.038
CSVT22	7.27	0.043	0.026
CSVM141	4.40	0.036	0.019
<b>Correlation coefficient</b>		<b>-0.105</b>	<b>0.998</b>



**Fig. 22. Correlation between nanopore numbers and intercellular molecular exchange.** Nanopore numbers and intercellular transfer of calcein between vegetative cells of *Anabaena* wild type and mutant strains CSVM34, CSVT22 and CSVM141 grown in BG11 medium or BG11<sub>0</sub> medium. (Upper panel) Summary table of results including nanopore numbers and  $R$  mean values for all the cells analyzed by FRAP except those with  $R > 0.3 \text{ s}^{-1}$ . The correlation coefficient between calcein transfer and nanopore number is shown for each set of calcein transfer data. (Lower panel) Graphic representation of calcein transfer against number of nanopores showing the positive correlation found for the data from BG11<sub>0</sub> medium but not for the data from BG11 medium.





### 3.1.4 Discussion

Intercellular molecular exchange in the filaments of *Anabaena* takes place by diffusion through septal junctions, which are proteinaceous complexes that join the adjacent cells in the filament. The nanopores are likely the structures through which septal junctions traverse the septal PG (Flores *et al.*, 2018). Mutants lacking septal proteins SepJ and/or FraCD are impaired in intercellular molecular exchange and show a decreased number of nanopores (Nürnberg *et al.*, 2015). Here, we have shown that, considering only cells showing an active intercellular exchange (i.e., in cultures incubated in the absence of combined N), intercellular calcein transfer positively correlates with number of nanopores per septal PG disk. This is consistent with the idea that the nanopores hold the septal junctions and suggests that the number of septal junctions corresponds to the number of nanopores.

The number of nanopores per septal PG disk in *Anabaena* sp. strain PCC 7120 is about 40 to 60 (Fig. 14C); see also Mariscal *et al.*, 2016; Nieves-Mori6n *et al.*, 2017b; Nürnberg *et al.*, 2015), but it has been reported to be about 155 in *Nostoc punctiforme* ATCC 29133 (Lehner *et al.*, 2013) or from 100 to 250 (described as “microplasmodesmata”) in the septa between vegetative cells of different heterocystous cyanobacteria including *Anabaena cylindrica* Lemm., *A. variabilis* (IUCB B377), *A. variabilis* Kütz. (ATCC 29413) and *Nostoc muscorum* (Giddings & Staehelin, 1981). It should be understood that a non-communicating cell will have all (or nearly all) of its septal junctions closed in the septa at both cell poles. This suggests the existence of regulatory mechanisms that determine the closing or opening of all the junctions in a cell depending on physiological conditions. Whereas non-communicating cells have previously been produced by a severe treatment of the culture (Weiss *et al.*, 2019) or have been observed as aging heterocysts (Nürnberg *et al.*, 2015) or in mutants of proteins (specifically sucrose transporters) that may affect the regulation of the septal junctions (Flores *et al.*, 2018), our results show that non-communicating cells can coexist with communicating cells in normal cultures. Although their frequency may depend on different growth parameters, the fact that communicating cells are more abundant in the absence than in the presence of combined nitrogen is consistent with the need of intercellular molecular exchange for diazotrophic growth.

We have also studied three *Anabaena* mutants that produce different version of SepJ. Strain CSVM90 produces a SepJ protein that has only one transmembrane segment and is impaired in septal localization (Ramos-Le6n *et al.*, 2018). Consistently, it produces a number of nanopores and shows calcein transfer similar to strain CSVM34 that is a deletion mutant of *sepJ*. Re-building, from CSVM90, a *sepJ* gene with a short internal deletion (SepJ<sub>Δ498-507</sub>) results in strain CSFR12 that produces a low number of nanopores and is impaired in calcein transfer (tested in BG11 medium), whereas re-building a *sepJ* gene encoding a point mutation (SepJ<sub>E663A</sub>) results in strain CSFR14 that produces a normal number of nanopores and is unaffected in calcein transfer (tested in BG11 medium) (see Table 7 and Fig. 20). Although we did not test calcein transfer in BG11<sub>0</sub> medium for these strains, the results are consistent with a correspondence between calcein transfer (and, hence, septal junctions) and nanopores as concluded above.

The study of a mutant of the recently identified gene *sepI* (Springstein *et al.*, 2020) unraveled a novel phenotype related to production of nanopores. This mutant showed “sealed” septal PG disks with a very low number of nanopores and “unsealed” disks with a large perforation. This perforation resembles what is seen in disks in the process of septal PG synthesis (see Fig. S5 in Bornikoel *et al.*, 2018), but we observed few of these disks in the wild type making it possible



that the presence of frequent unsealed disks is a characteristic of the  $\Delta sepI$  mutant. Indeed, this mutant has been shown to be affected in several parameters related to PG biogenesis (Springstein *et al.*, 2020). Regarding the intercellular transfer of calcein, we specifically observed an increased number of non-communicating cells in the  $\Delta sepI$  mutant in BG11<sub>0</sub> medium (also observed for another fluorescent marker, 5-carboxyfluorescein; Springstein *et al.*, 2020). It will be of high interest to study whether this increased number of non-communicating cells is related to the presence of sealed disks with too few nanopores/septal junctions. In contrast, the cells significantly transferring calcein (Fig. 21) might have a normal number of septal junctions, which should be hold in the unsealed perforations of the septal PG disks. This unexpected finding will need further investigation in the future.



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



### 3. Results & Discussion

---

#### Chapter 2

#### The HglK protein

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



### 3.2 The HglK protein

The pentapeptide-repeat HglK protein was initially described by Black *et al.* (1995) as a protein that could be involved in the transport of glycolipids in the heterocysts of *Anabaena*. This hypothesis was proposed based on a TEM study of a  $\Delta hglK$  mutant, which showed that the heterocysts did not possess the glycolipid envelope characteristic of this cell type, but lipid vesicles could be seen inside the heterocysts (Black *et al.*, 1995). Along with this observation, the microscopic images of the mutant also showed other structural anomalies such as the presence of intercellular septa between vegetative cells and between vegetative cells and heterocysts that were wider than those observed in the wild type (Black *et al.*, 1995). The *hglK* gene is found in all cyanobacteria belonging to clade B1 (Shih *et al.*, 2013), which includes heterocystous cyanobacteria and other phylogenetically related strains that do not produce heterocysts (Table 9). The phenotype of wide septa presented by the *hglK* mutant and the general presence of this gene in heterocystous cyanobacteria led us to ask whether HglK could have a role in the septal junctions.

Strain	HglK	PatL	HetL	FraCDE-F
<i>Fischerella</i> sp. PCC 9605	✓	✓		✓✓
<i>Fischerella</i> sp. PCC 9431	✓	✓	✓	✓✓
<i>Fischerella</i> sp. PCC 9339	✓	✓	✓	✓✓
<i>Fischerella</i> sp. JSC-11	✓	✓	✓	✓✓
<i>Chlorogloeopsis</i> sp. PCC 7702	✓			✓✓
<i>Mastigocladopsis repens</i> PCC 10914	✓	✓		✓✓
<i>Calothrix</i> sp. PCC 7103	✓	✓	✓	✓✓
<i>Calothrix</i> sp. PCC 6303	✓	✓		✓
<i>Nostoc</i> sp. PCC 7524	✓	✓		✓✓
<i>Anabaena</i> sp. PCC 7120	✓	✓	✓	✓✓
<i>Anabaena variabilis</i> ATCC 29413	✓	✓	✓	✓✓
<i>Nostoc</i> sp. PCC 7107	✓	✓		✓✓ *
<i>Cylindrospermum stagnale</i> PCC 7417	✓	✓		✓✓
<i>Anabaena cylindrica</i> PCC 7122	✓	✓	✓	✓
<i>Anabaena</i> sp. PCC 7108	✓	✓	✓	✓
<i>Raphidiopsis brookii</i> D9 †	✓			✓
<i>Cylindrospermopsis raciborskii</i> CS-505	✓			✓
' <i>Nostoc azollae</i> ' 0708	✓	✓		✓
<i>Nodularia spumigena</i> CCY9414	✓	✓		✓
<i>Calothrix</i> sp. PCC 7507	✓	✓		✓✓
<i>Microchaete</i> sp. PCC 7126	✓	✓		✓✓
<i>Nostoc punctiforme</i> PCC 73102	✓	✓		✓✓ *
<i>Scytomena hofmanni</i> UTEX 2349	✓	✓		✓
<i>Rivularia</i> sp. PCC 7116	✓	✓		✓
<i>Synechocystis</i> sp. PCC 7509 †	✓			
<i>Gloeocapsa</i> sp. PCC 7428 †	✓			
<i>Chroococciopsis thermalis</i> PCC 7203 †	✓			

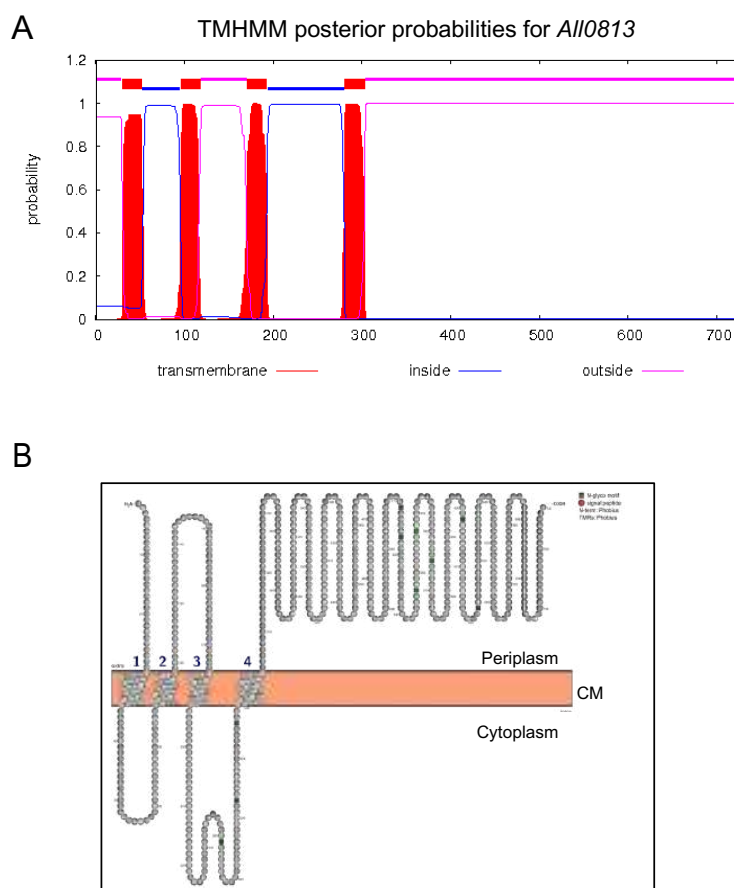
**Table 9. HglK, PatL, HetL and the Fra proteins in cyanobacteria.** The cyanobacteria included in the phylogenetic analysis are those of clade B1 from Shih *et al.* (2013). This clade includes heterocystous cyanobacteria and phylogenetically related strains that do not produce heterocysts (†). The indicated proteins from *Anabaena* sp. strain PCC 7120 were used as query in BLAST analysis. The presence of a



protein in a strain is indicated by a check mark (✓). For the Fra proteins, the presence of a *fraC-fraD-fraE* operon is indicated by one check mark, and the additional presence of the adjacent (reverse) *fraF* gene is indicated by a second check mark. \*In these strains, the *fraF* gene is about 1.6 (PCC 7107) and 5.7 (PCC 73102) kb downstream of the *fraCDE* operon.

### 3.2.1 Subcellular localization of HglK

To address the subcellular localization of HglK, two approaches were tried, one by fusing a fluorescent protein (superfolder GFP) to HglK and the other by fusing HglK with different tags for immunolocalization analysis. With the use of different bioinformatic programs it was predicted that HglK has four transmembrane segments with both the C- and N-termini in the periplasm (Fig. 23). For convenience in the genetic constructs and to avoid as much as possible effects on the targeting of the protein to the membrane, we decided to add the sfGFP and the different tags to the carboxyl end of HglK. Because of its predicted periplasmic localization, we used the superfolder GFP (sfGFP) that can fold efficiently in the periplasm (Dinh & Bernhardt, 2011).

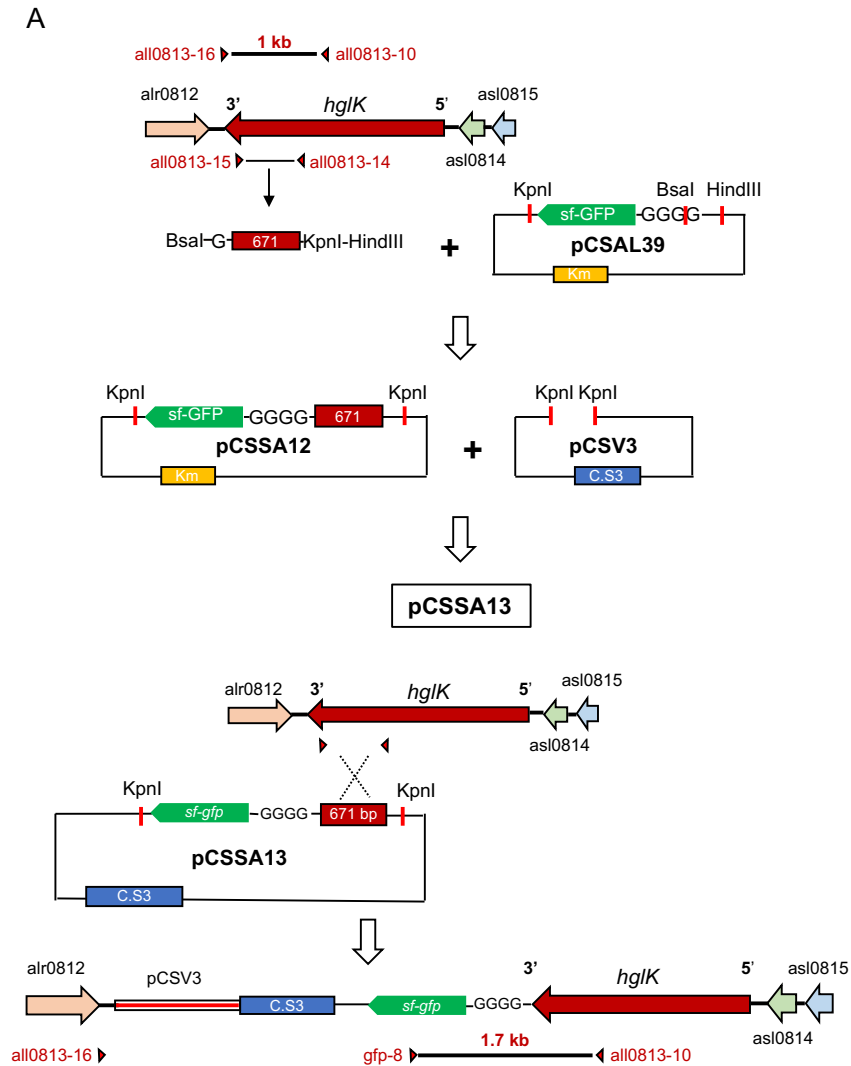


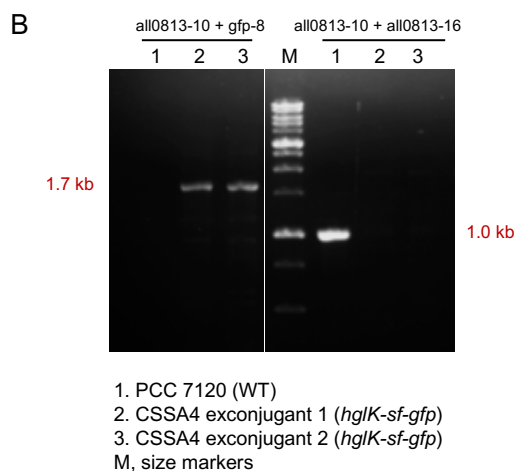
**Fig. 23. Topology of the HglK (All0813) protein.** (A) Transmembrane segments predicted with the TMHMM program (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>). (B) Protter representation based on prediction with Phobius (<http://phobius.sbc.su.se>). CM, cytoplasmic membrane.





Strain CSSA4 was first constructed that expressed HglK fused to the sfGFP. To create this strain, a plasmid that carried a 671-bp 3' fragment of *hglK* fused through a sequence encoding a four-glycine linker to *sf-gfp* was prepared using vector pCSV3 (which contains the C.S3 cassette that provides resistance to Sm and Sp; Valladares *et al.*, 2011). This plasmid was transferred to *Anabaena* wild type by conjugation (Elhai *et al.*, 1997) with selection for resistance to Sm and Sp. The insertion into *hglK* and segregation of chromosomes carrying the fusion was tested by PCR using primers all0813-10 and *gfp-8* to check the insertion of *sf-gfp*, and all0813-16 and all0813-10 to check segregation of the mutated chromosomes (Fig. 24).





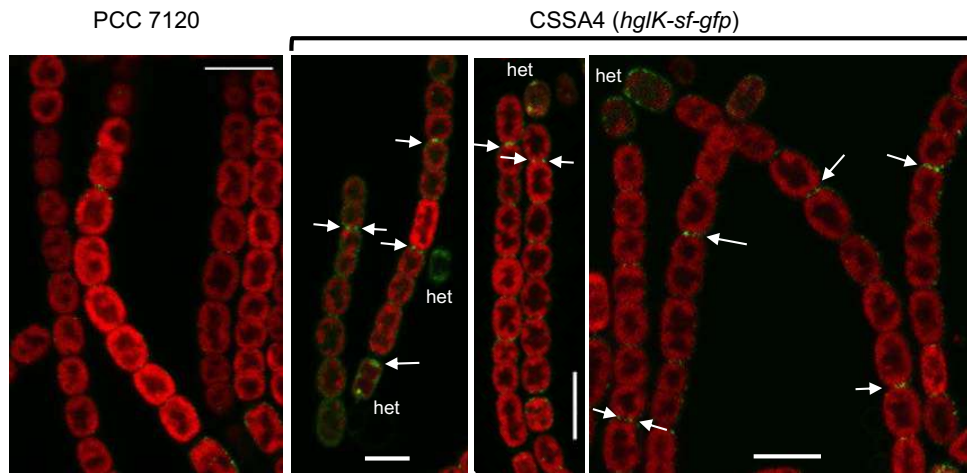
**Fig. 24. Construction of an *Anabaena* strain that produces a fusion of the superfolder GFP to the C-terminus of HglK.** (A) Scheme of the construction of plasmid pCSSA13 and of the recombination between this plasmid and the *Anabaena* chromosome. (B) PCR analysis with the indicated oligodeoxynucleotide primers. The *sf-gfp* fusion was detected in the two exconjugants analyzed but not in the WT, and wild-type chromosomes were not detected in the exconjugants.

The *hglK* gene shows low levels of expression (Black *et al.*, 1995) even after incubation without combined nitrogen, which induces its expression (Table 10). The HglK-sfGFP strain and the wild type (negative control) were grown for 4 days in liquid BG11C medium (supplemented with Sm+Sp for the HglK-sfGFP strain, CSSA4) under bubbling conditions. The cells were then collected, washed and re-inoculated at a concentration of 3 µg Chl/mL in liquid BG11C medium (without combined nitrogen or antibiotics) for induction during 24 hours with bubbling. After the incubation period, HglK-sfGFP could be detected by confocal microscopy in the periphery of the cells (outside of the thylakoid's red fluorescence), mainly in the intercellular septa, indicating that HglK is a cytoplasmic membrane protein with a polar location (Fig. 25). In most cases, the GFP was observed only as one fluorescence spot, but in some cells two spots could be observed. Because the test was performed under nitrogen deprivation, the filaments developed heterocysts, in which HglK-sfGFP could be observed mainly at the poles.

Reference	Fold increase after N step-down				
	Time after N step-down				
	3 h	8 h	12 h	21 h	24 h
Ehira & Omori (2006a)	2.3 x	3.4 x			
Flaherty <i>et al.</i> , 2011*			4.7 x	3.3 x	

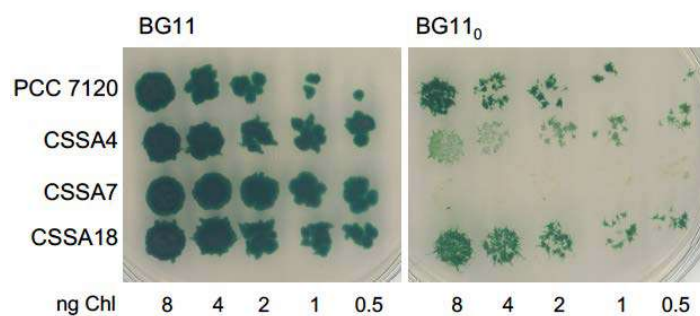
**Table 10. Expression of the *hglK* gene in *Anabaena*.** \*Appreciable expression was observed in the presence of ammonium.





**Fig. 25. Subcellular localization of HglK-sfGFP.** Confocal microscopy of *Anabaena* wild type (PCC 7120) and strain CSSA4 that produces a protein with a fusion of superfolder GFP to the C-terminus of HglK (*hglK-sf-gfp*). Filaments are from cultures incubated for 24 h in bubbled BG11<sub>0</sub> medium. GFP fluorescence (green) and cyanobacterial autofluorescence (red) are shown for each strain. Some intercellular septa in which the GFP fluorescence is clearly observed are indicated by arrows; double-opposite arrows indicate some intercellular septa in which two spots are observed. het, heterocyst. Size bars, 5 μm.

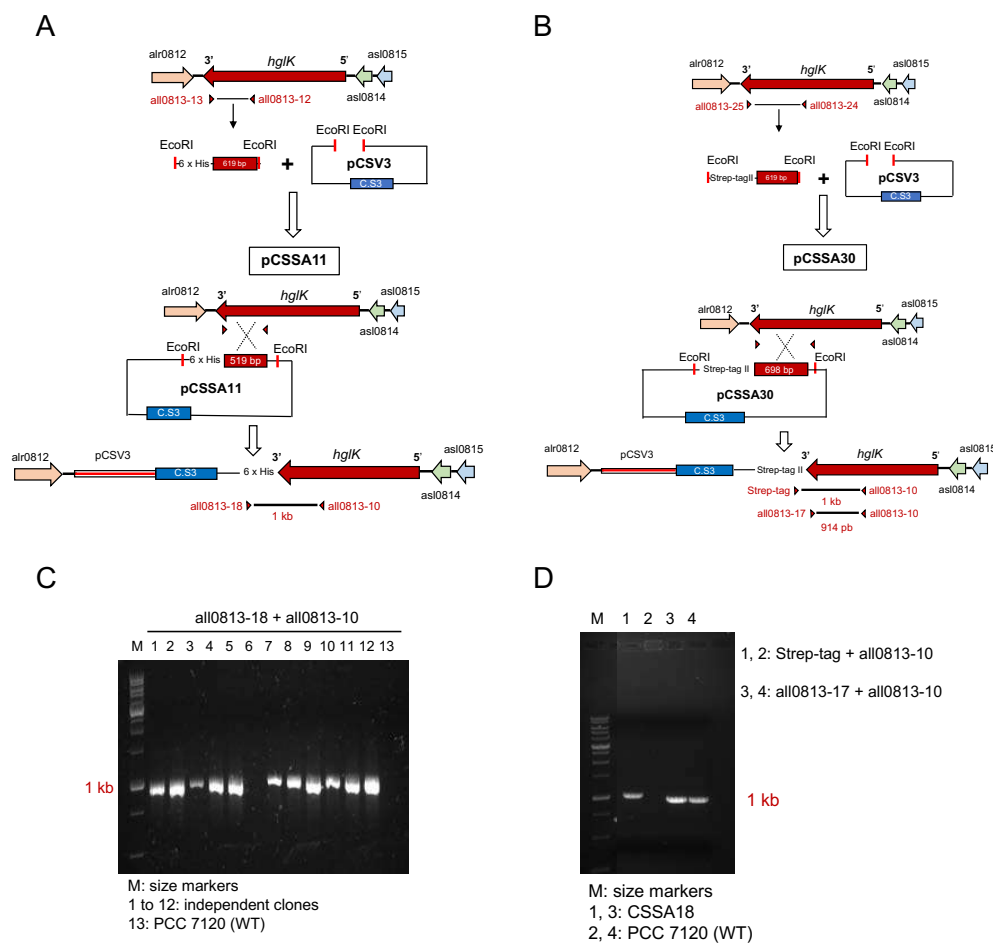
To study whether the added GFP could alter the functionality of HglK, growth tests were performed. The strains were grown in liquid BG11 medium (adding antibiotics for the mutant), washed and resuspended in liquid BG11<sub>0</sub> medium at a concentration of 1 μg Chl/mL, from which serial dilutions were made, and drops were placed on plates of BG11 and BG11<sub>0</sub> media (no plate contained antibiotics since the wild type was also inoculated as a negative control). After 14 days, it was observed that the strain that produced HglK-sfGFP did not grow (or grew very weakly) under diazotrophic growth conditions (Fig. 26). These results indicate that HglK-sfGFP is hardly functional, making it necessary to take another approach to corroborate HglK localization.



**Fig. 26. Growth test of *Anabaena* wild type (PCC 7120) and strains CSSA4 (*hglK-sf-gfp*), CSSA7 (*hglK-6xHis*) and CSSA18 (*hglK-Strep-tag II*).** Filaments grown in BG11 medium (with antibiotics for the mutants) were inoculated with the indicated amounts of Chl in BG11 or BG11<sub>0</sub> medium (without antibiotics) and incubated under culture conditions for 14 days.



As an alternative approach to address the subcellular localization of HglK, two strains were prepared, one producing HglK with a 6xHis tag and another with the Strep-tag II. Both tags were fused to the C-terminus and allowed us to perform immunolocalization analysis. To construct strain CSSA7 that carries the 6xHis tag, we cloned the 3' 519 bp of *hglK* (omitting the stop codon) adding a sequence for 6 histidines. On the other hand, to construct strain CSSA18 that carries the Strep-tag II, we cloned the 3' 698 bp of *hglK* (omitting the stop codon) adding the sequence encoding Strep-tag II. Both fragments were introduced into vector pCSV3 generating plasmids pCSSA11 and pCSSA30, respectively. These plasmids were corroborated by sequencing and transferred to *Anabaena* by conjugation with selection for Sm+Sp (Fig. 27). In order to test whether these tags affect HglK functionality, growth tests were performed (as described above). It was observed that the strain carrying the 6xHis tag did not grow diazotrophically but, on the other hand, the strain carrying the Strep-tag II did grow, indicating that HglK was functional in this case (Fig. 26).



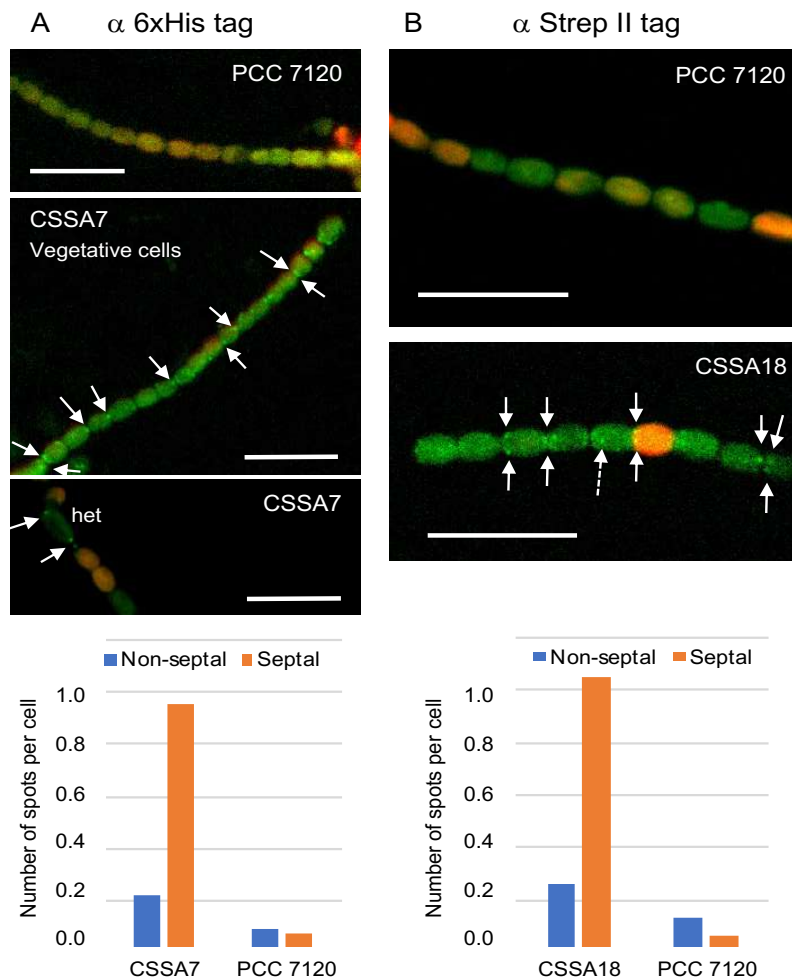
**Fig. 27. Construction of *Anabaena* strains that produce an HglK protein with a 6 x His tag or a Strep-tag II at the C-terminus.** (A) Scheme of the construction of plasmid pCSSA1 and of recombination between this plasmid and the *Anabaena* chromosome. (B) Scheme of the construction of plasmid pCSSA30 and of recombination between this plasmid and the *Anabaena* chromosome. (C) PCR analysis of pCSSA1 exconjugants and the WT with the indicated oligodeoxynucleotide primers. The correct construct was found in all but one (# 6) of the tested exconjugants, one of which was chosen and termed CSSA7. (D) PCR



analysis of pCSSA30 exconjugant CSSA18 and the WT with the indicated oligodeoxynucleotide primers (3, 4: positive controls of DNA loading).

The strains were grown in liquid BG11 medium (with antibiotics, Sm+Sp, for the mutants), and filaments were collected, washed and incubated in liquid BG11<sub>0</sub> medium (without combined nitrogen or antibiotics) for the time indicated in each case. Immunolocalization was performed on those filaments with commercial antibodies that carried a fluorophore and were specific for the 6xHis-tag and the Strep-tag II, respectively. Samples were visualized by fluorescence microscopy with excitation at 488 nm. Strain CSSA7 (HglK-6xHis) produced small fluorescent foci mainly in the intercellular septa and in the poles of the heterocysts (Fig. 28A), and strain CSSA18 (HglK-Strep-tag II) also showed fluorescent foci in or near the intercellular septa (Fig. 28B). With both tagged proteins, the presence of double spots of fluorescence was frequently observed (double arrows in Fig. 28A and B), resembling the double spots occasionally observed with HglK-sfGFP (double arrows in Fig. 25). To confirm that the subcellular location of HglK was mainly septal, the spots observed in the immunolocalization analysis were quantified using the wild type as a control of non-specific binding of antibodies to biological samples (Fig. 28, lower panels). About one septal spot and 0.2 non-septal spots per cell were counted for both HglK-6xHis and HglK-Strep-tag II, whereas the wild-type control showed in both cases about 0.1 septal and 0.1 non-septal spots per cell. If the wild-type controls are subtracted, the number of immunofluorescence spots is about 7.5-fold higher in the septal region than in the rest of the cell for both HglK-6xHis and HglK-Strep-tag II. In summary, in spite of a low expression of the *hglK* gene (Black *et al.*, 1995), these results together with those obtained with HglK-sfGFP indicate that HglK is a protein of predominant localization at the intercellular septa in *Anabaena*.





**Fig. 28. Subcellular localization of 6xHis- and Strep II-tagged HglK.** (A) Immunofluorescence analysis of *Anabaena* wild type (PCC 7120, negative control) or strain CSSA7 that produces to HglK protein with a C-terminal 6xHis tag. Filaments incubated for 48 h in BG11<sub>0</sub> medium were prepared and subjected to immunofluorescence analysis with anti-His-tag antibodies as described in section 2.7.2 of Materials & Methods. No signal is observed at the vegetative cells in the bottom panel because the filament was focused to see the signal at the heterocyst (het). (B) Immunofluorescence analysis of *Anabaena* wild type (PCC 7120, negative control) or strain CSSA18 that produces the HglK protein with C-terminal Strep II tag. Filaments incubated for 24 h in BG11<sub>0</sub> medium were prepared and subjected to immunofluorescence analysis with anti-Strep-tag II antibodies as described in section 2.7.2 of Materials & Methods. Solid single arrows point to some immunofluorescence spots localized at intercellular septa; double-opposite arrows indicate paired spots frequently observed in the regions of the intercellular septa; the dotted arrow indicates an example of immunofluorescence spot localized outside of the septal region. Histograms in the lower part show the number of immunofluorescence spots counted in the septal regions (right columns, orange) or outside the septal region (left columns, blue) in the strains producing the HglK tagged proteins or the wild type (PCC 7120) as a negative control. The number of cells analyzed were: for the His-tag antibodies, 115 (CSSA7) and 99 (PCC 7120); for the Strep-tag antibodies, 111 (CSSA18) and 136 (PCC 7120). Size bars, 10 μm.

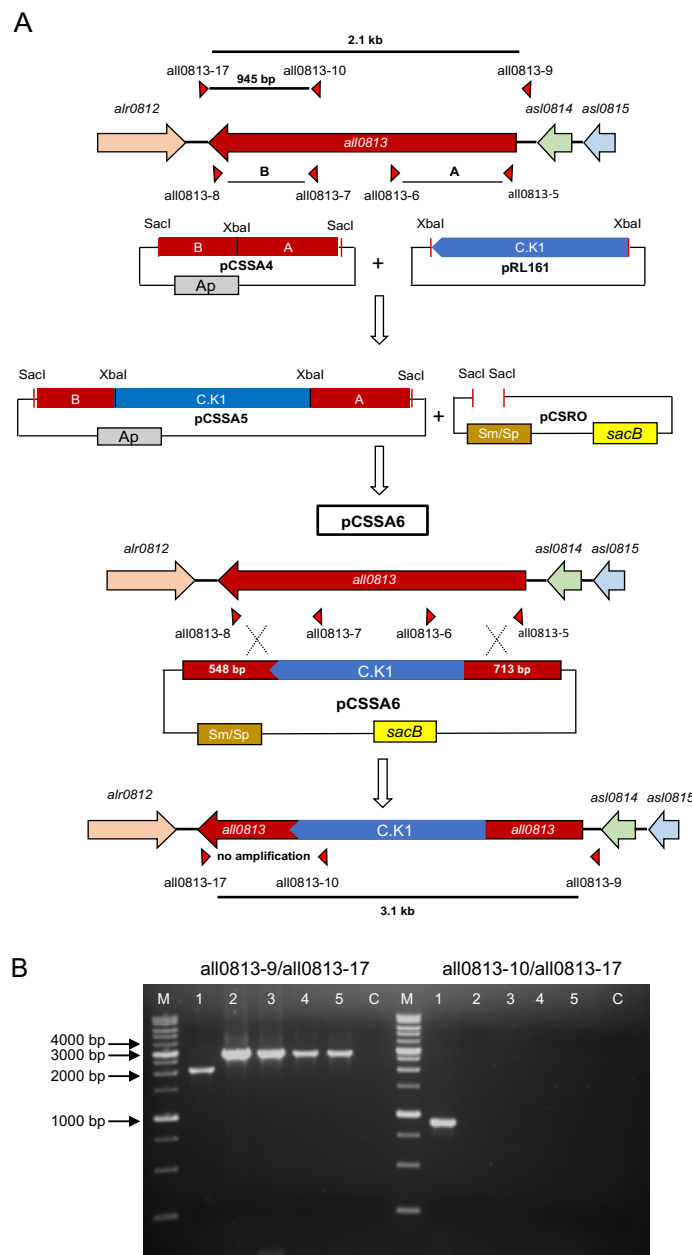


### 3.2.2 Inactivation of *hglK* in different genetic backgrounds

To investigate some aspects of the phenotype of the *hglK* mutant that were not addressed in the first description made by Black *et al.* (1995), a plasmid was constructed that produced the mutation of *hglK* by deletion of an internal fragment and insertion of a gene cassette. This plasmid was transferred by conjugation to *Anabaena* with different genetic backgrounds. For creation of the plasmid, 713 bp of the 5' region of *hglK* and 548 bp of the 3' region were cloned, linked and corroborated by sequencing. Then, gene cassette C.K1 (which bears the *npt* gene that confers resistance to neomycin and its own natural promoter; Elhai & Wolk, 1988a) was introduced, and the whole fragment was transferred to plasmid vector pCSRO (Merino-Puerto *et al.*, 2013) that bears cassette C.S3 and the *sacB* gene for positive selection through sensitivity to sucrose (Cai & Wolk, 1990), thus obtaining plasmid pCSSA6 that carries everything necessary to obtain the *hglK* mutant by deletion/insertion (Fig. 29A). The plasmid was transferred by conjugation to different genetic backgrounds with the objective of producing the replacement of *hglK* by the mutated  $\Delta hglK::C.K1$  construct. The genetic backgrounds used and the strains produced were: *Anabaena* wild type generating mutant CSSA1 ( $\Delta hglK::C.K1$ ); the double  $\Delta fraC \Delta fraD$  mutant (CSVT22) that gave rise to strain CSSA2 ( $\Delta fraC \Delta fraD \Delta hglK::C.K1$ ); the  $\Delta sepJ$  mutant (CSVM34) that generated CSSA3 ( $\Delta sepJ \Delta hglK::C.K1$ ); and, finally, a triple mutant of the genes  $\Delta fraC$ ,  $\Delta fraD$  and  $\Delta sepJ$  (CSVM141) generating CSSA12 ( $\Delta fraC \Delta fraD \Delta sepJ \Delta hglK::C.K1$ ). All these exconjugants were initially selected for their resistance to neomycin and for their ability to grow in a medium supplemented with sucrose, and finally checked to corroborate their genetic structure (Fig. 29B).







**Fig. 29. Construction of *Anabaena* strains with a  $\Delta hglK::C.K1$  mutation.** (A) Scheme of the construction of plasmid pCSSA6 and of the double recombination between this plasmid, which bears  $\Delta hglK::C.K1$ , and the *Anabaena* chromosome. Plasmid pCSSA6 was transferred by conjugation to *Anabaena* with different genetic backgrounds (WT, *fraC fraD*, *sepJ*). The approximate localization of oligodeoxynucleotide primers used in PCR for construction of the mutation and in PCR analysis are shown along with predicted products. (B) PCR analysis with the indicated oligodeoxynucleotide primers and DNA from the following *Anabaena* strains: 1, PCC 7120 (WT); 2, CSSA1 ( $\Delta hglK::C.K1$ ); 3, CSSA2 ( $\Delta fraC \Delta fraD \Delta hglK::C.K1$ ); 4, CSSA3 ( $\Delta sepJ \Delta hglK::C.K1$ ); 5, CSSA12 ( $\Delta fraC \Delta fraD \Delta sepJ \Delta hglK::C.K1$ ). M, size markers; C, negative control with water instead of DNA. Note that no amplification is expected in the mutants with the all0813-10/all0813-17 primer pair because they lack the sequence corresponding to primer all0813-10.

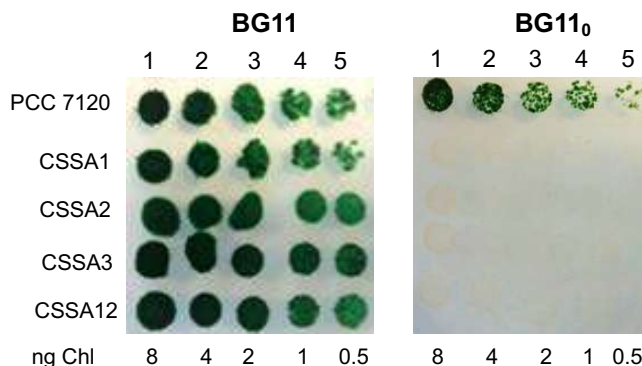


To confirm the inability of the  $\Delta hglK$  mutant to grow under diazotrophic conditions (previously described by Black *et al.*, 1995), growth tests on solid and liquid media were performed with the different mutants. For both tests, all the strains were grown in liquid medium with combined nitrogen (with antibiotics when necessary), washed and resuspended in a medium without combined nitrogen. To perform the tests on solid medium, drops of each strain (at known Chl concentrations) were placed on plates with (BG11) or without (BG11<sub>0</sub>) combined nitrogen and allowed to grow for ten days under normal culture conditions. To perform the test on liquid medium, flasks containing 50 mL of BG11 or BG11<sub>0</sub> medium were inoculated with a concentration of 0.1  $\mu\text{g Chl/mL}$ . Then, the cultures were incubated with shaking and samples were taken every 12 hours to measure the optical density (OD) at 750 nm in order to calculate a growth rate constant. As expected, neither the single  $\Delta hglK$  mutant nor any of the other strains, which carried several mutations including  $\Delta hglK::C.K1$ , was able to grow fixing nitrogen under oxic conditions (Fig. 30). In addition, a study based on optical microscope photographs of CSSA1 and wild-type filaments incubated in BG11<sub>0</sub> for 48 hours showed that the  $\Delta hglK$  mutant produced heterocysts (consistent with the results of Black *et al.*, 1995). The heterocysts in the mutant were 5.4% of the cells while in the wild type the heterocysts were 7% of the cells.

A

Strain	Genotype	$\mu$ (day <sup>-1</sup> )	
		BG11	BG11 <sub>0</sub>
PCC 7120	Wild type	1.142	0.640
CSSA1	$\Delta hglK::C.K1$	1.149	0.024
CSSA2	$\Delta hglK::C.K1 \Delta fraC \Delta fraD$	0.888	0.026
CSSA3	$\Delta hglK::C.K1 \Delta sepJ$	1.051	0.000

B



**Fig. 30. Growth phenotype of *Anabaena* and *hglK* mutant strains.** *Anabaena* wild type (PCC 7120) and the indicated mutants were grown in BG11 medium (with Nm for the *hglK* mutants), harvested, and washed and resuspended in BG11<sub>0</sub> medium. (A) Liquid BG11 or BG11<sub>0</sub> medium (without antibiotics) was inoculated with the indicated strain and incubated under culture conditions. Samples were taken at about 12-h intervals and their OD at 750 nm was determined. The growth rate constant ( $\mu$ ) was calculated from the logarithmic phase of growth. (B) Samples with the indicated amount of Chl were spotted on solid BG11 or BG11<sub>0</sub> medium (without antibiotics), incubated under growth conditions for about 10 days and photographed.

To study how active those heterocysts could be, their nitrogenase activity was determined by the acetylene reduction assay under oxic and anoxic conditions. To carry out the test, the CSSA1 and wild-type strains were grown in liquid BG11 medium and incubated in BG11<sub>0</sub> medium for 48

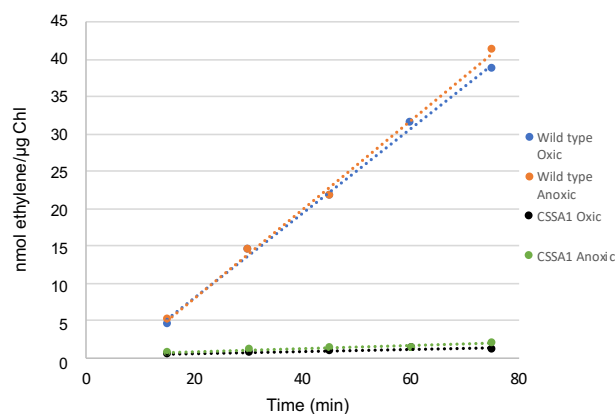


hours. In order to carry out the nitrogenase activity assay under anoxic conditions, DMCU was added to inactivate photosystem II (and therefore to prevent the release of oxygen), and the flasks were bubbled with argon to remove the oxygen from the flask. The CSSA1 activity was about 1.6% and 3.7% the activity of wild type under oxic and anoxic conditions, respectively (Fig. 31). Hence, in contrast to the *hglK* mutant described earlier that was reported to fix nitrogen microaerobically (although no quantitative data were shown; Black *et al.*, 1995), our  $\Delta hglK$  mutant only shows a marginal increase of nitrogenase activity when assayed under anoxic conditions, implying a defect in the production of this enzyme.

A

Strain	Genotype	Assay	Nitrogenase activity (nmol ethylene [ $\mu\text{g Chl}$ ] <sup>-1</sup> h <sup>-1</sup> ) Mean $\pm$ SEM (n)
PCC 7120	Wild type	Oxic	25.18 $\pm$ 7.12 (6)
CSSA1	$\Delta hglK::C.K1$		0.41 $\pm$ 0.19 (8)
PCC 7120	Wild type	Anoxic	25.92 $\pm$ 9.86 (5)
CSSA1	$\Delta hglK::C.K1$		0.95 $\pm$ 0.39 (6)

B

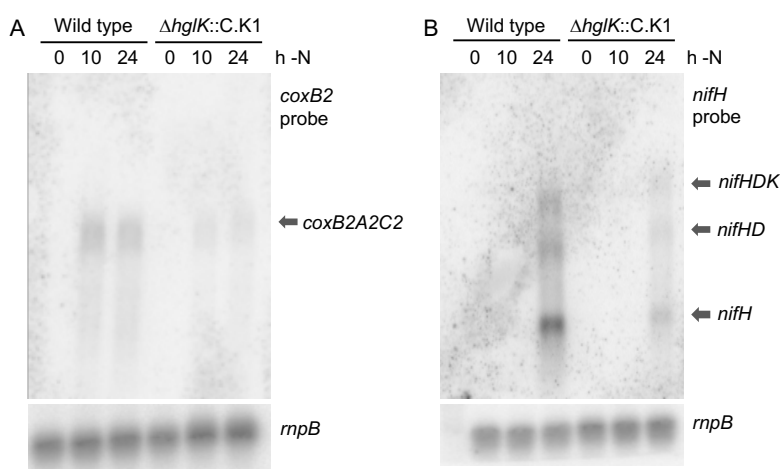


**Fig. 31. Nitrogenase activity of *Anabaena* and the *hglK* mutant.** *Anabaena* wild type (PCC 7120) and strain CSSA1 were grown in BG11 medium (with Nm for the *hglK* mutant), washed and resuspended in BG11<sub>0</sub> medium (without antibiotics), and incubated under growth conditions for 48 h. Nitrogenase activity was determined by the acetylene reduction assay under oxic and anoxic conditions as described in Materials & Methods. (A) Summary of the assays performed in this study. Student's *t* test results: WT vs. CSSA1 (oxic),  $p = 0.0016$ ; WT vs. CSSA1 (anoxic),  $p = 0.021$ ; CSSA1 oxic vs. CSSA1 anoxic,  $p = 0.195$ . (B) The result of a typical assay is shown.

To investigate whether the inactivation of *hglK* has any effect on the expression of genes activated at medium and late steps of heterocyst differentiation, *northern* analysis with probes of the *coxB2* and *nifH* genes was performed. The *cox2* operon encodes a heterocyst-specific cytochrome *c* oxidase that is expressed at mid-differentiation (transcript detected at about 9 hours; Valladares *et al.*, 2003), and the *nifHDK* genes encode the nitrogenase complex and are expressed late in differentiation (transcripts detected at about 18 hours; Wei *et al.*, 1994). The  $\Delta hglK$  mutant and the wild-type strain were grown in bubbled BG11C medium and incubated in bubbled BG11<sub>0</sub>C medium. Samples for RNA isolation were taken at an initial time (0 h), at 10 hours of induction, and at 24 hours of induction. Two membranes were prepared for the *northern* analysis with the



isolated RNA, one was hybridized with a *coxB2* probe which detected a transcript of about 3.7 kb of the *cox2* operon (Fig. 32A) (Valladares *et al.*, 2003), and the other one was hybridized with a *nifH* probe which detected the *nifHDK* (about 4.7 kb), *nifHD* (about 2.8 kb) and *nifH* (about 1.1 kb) transcripts (Fig. 32B) (Wei *et al.*, 1994). The expression of both the *cox2* operon and the *nifHDK* genes was substantially decreased in the  $\Delta hglK$  mutant as compared to the wild type (Fig. 32). After normalization with the *rnpB* control, the levels of the *cox2* operon transcript in the mutant were calculated to be about 15% and 19% those of the wild type after 10 and 24 hours of nitrogen deprivation, respectively, and the levels of the *nifH* transcript in the mutant were about 8.5% those of the wild type after 24 hours of nitrogen deprivation. These results are consistent with the detection of low levels of nitrogenase activity in the  $\Delta hglK$  mutant and indicate that inactivation of *hglK* impairs gene expression associated with heterocyst differentiation.

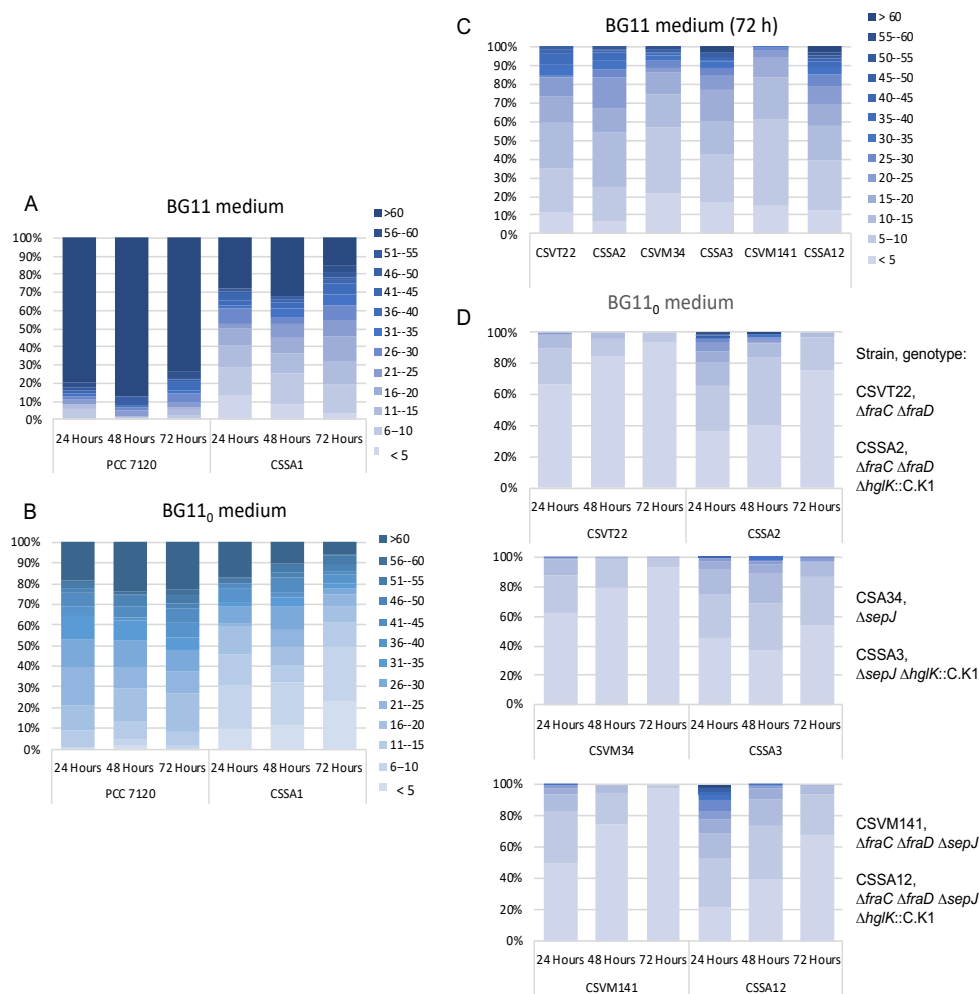


**Fig. 32. Expression of the *cox2* and *nifHDF* gene clusters in *Anabaena* wild type and the *hglK* mutant.** Northern blot analysis with RNA isolated from cultures grown in bubbled BG11C medium (supplemented with Nm for the mutant) and incubated for 10 or 24 h in bubbled BG11<sub>0</sub>C medium (without antibiotics). (A) A probe of the *coxB2* gene was used. Note the presence of signals corresponding to the *cox2* operon transcript (about 3.7 kb; Valladares *et al.*, 2003). (B) A probe of the *nifH* gene used. Note the presence of *nifHDK* (about 4.7 kb), *nifHD* (about 2.8 kb), and *nifH* (about 1.1 kb) transcripts (Wei *et al.*, 1994). Hybridization with a probe of the *rnpB* gene was used as a loading and transfer control.

Mutants of genes such as *sepJ*, *fraC* and *fraD* that encode septal proteins generally show a filament fragmentation phenotype (Bauer *et al.*, 1995; Flores *et al.*, 2007; Merino-Puerto *et al.*, 2010). Therefore, because the HglK protein appears to be localized at the intercellular septa, we investigated the filament length in the  $\Delta hglK$  mutant. To address this question, the mutant and the wild type were grown in liquid BG11 medium (with neomycin for the *hglK* mutant) and re-inoculated in new flasks with liquid BG11 or BG11<sub>0</sub> medium (without the antibiotic) at a concentration of 1  $\mu$ g Chl/mL. After 24, 48 and 72 hours, samples from the different strains and media were taken with great care to prevent disruption of the filaments, and they were photographed with an optical microscope. Finally, the cells in each filament were counted from a total of between 100 and 120 filaments per strain and culture condition. The distribution of filament lengths was analyzed by the Chi-square test to determine if there were statistically-significant differences between the mutant and the wild type. Strain CSSA1 ( $\Delta hglK::C.K1$ ) produced significantly shorter filaments than the wild type (Fig. 33A and B), and this was



especially evident in BG11 medium in which the wild type produced the highest proportion of long filaments (Fig. 33A). On the other hand, to study whether the *hglK* mutation increases fragmentation in combination with other mutations, the protocol described above was carried out with strains CSSA2 ( $\Delta fraC \Delta fraD \Delta hglK::C.K1$ ), CSSA3 ( $\Delta sepJ \Delta hglK::C.K1$ ) and CSSA12 ( $\Delta fraC \Delta fraD \Delta sepJ \Delta hglK::C.K1$ ) compared to their respective parental strains. In no case the inactivation of *hglK* was observed to increase the fragmentation observed in mutants of the *sepJ*, *fraC* or *fraD* genes (Fig. 33C and D). Hence, the inactivation of *hglK* produces filament fragmentation but does not increase the fragmentation resulting from inactivation of *sepJ*, *fraC* or *fraD*.

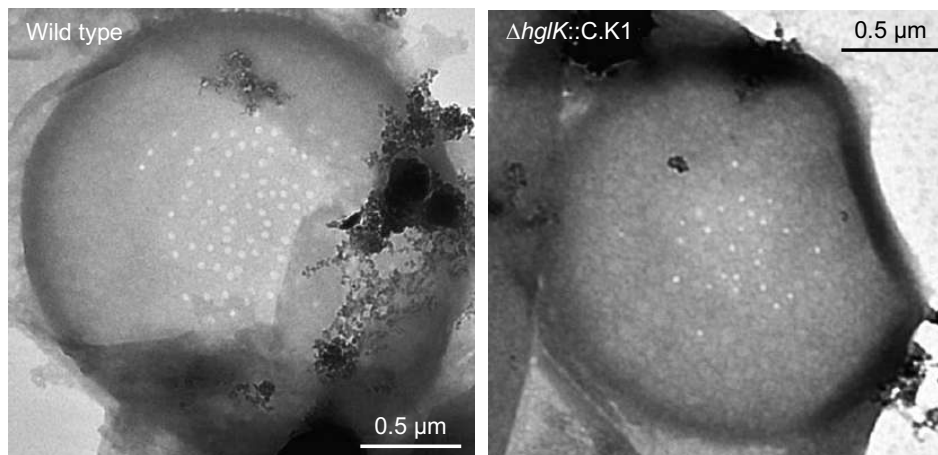


**Fig. 33. Filament length in *hglK* mutant strains with different genetic backgrounds.** The wild type and indicated mutants were grown in BG11 medium (with Nm for the *hglK* mutants), harvested, washed, resuspended in liquid BG11 (A and C) or BG11<sub>0</sub> (B and D) medium (without antibiotics) at 1  $\mu$ g Chl/mL and incubated under culture conditions for the times indicated in hours. Samples were taken with great care, and filament length was determined in 100 to 250 filaments for each strain. The color code on the right indicates number of cells per filament.



### 3.2.3 Intercellular junctions and communication

Because of the broad intercellular septa present in the original *hglK* mutant (Black *et al.*, 1995), and because of the predominant septal localization of HglK (Figs. 25 and 28) and the fragmentation phenotype of our  $\Delta hglK::C.K1$  mutant (Fig. 33), we investigated structures and functions related to intercellular communication. The nanopores found in the septal PG disks between cells are related to intercellular communication (see Results & Discussion Chapter 1). In order to know if the  $\Delta hglK$  mutant was affected in the number of nanopores compared to the wild type, both strains were analyzed in parallel and their nanopores quantified. Strain CSSA1 ( $\Delta hglK::C.K1$ ) and the wild type were grown in liquid BG11 medium (with neomycin for the mutant). When both strains reached about 3-4  $\mu\text{g}$  Chl/mL, the cultures were collected, washed, resuspended in liquid BG11<sub>0</sub> medium (without the antibiotic) and incubated for 48 hours (i.e., an induction under nitrogen deprivation). The cultures were then collected and sonicated to decrease filament length to about 1-3 cells per filament in order to improve the observation of septal disks. Peptidoglycan (PG) was isolated and visualized by TEM as described in Results & Discussion Chapter 1 and Materials & Methods. After quantifying the nanopores in the septal disks, we observed that the  $\Delta hglK$  mutant had about 50% the number of nanopores found in the wild type, and the nanopore diameter was slightly larger in the mutant than in the wild type (Fig. 34). Thus, HglK appears to have a role in the construction of mature septa.



Strain	Nanopore number		Nanopore diameter (nm)	
	Mean $\pm$ SD (n)	<i>p</i>	Mean $\pm$ SD (n)	<i>p</i>
<i>Anabaena</i> (WT)	39 $\pm$ 21 (26)		14.3 $\pm$ 9.3 (140)	
CSSA1 ( $\Delta hglK::C.K1$ )	20 $\pm$ 13 (14)	0.003	15.8 $\pm$ 2.8 (179)	0.048

**Fig. 34. Nanopores in septal peptidoglycan disks of wild-type *Anabaena* and the *hglK* mutant.** Cells were grown in BG11 medium (in the presence of Nm for the mutant), incubated in BG11<sub>0</sub> medium in the absence of antibiotic for 48 h and used for isolation of PG and visualization by TEM as described Materials & Methods. The images at the top show examples of septal disks from the wild type and the *hglK* mutant. The table below summarizes quantitative data: n, number of septal PG disks analyzed to count nanopores or number of nanopores measured. Student's *t*-test *p* (WT vs. mutant) indicated in each case.

We next looked at the intercellular transfer of fluorescent markers studied by FRAP analysis. Because *hglK* is induced in the absence of nitrogen (Table 10), we used in this analysis filaments





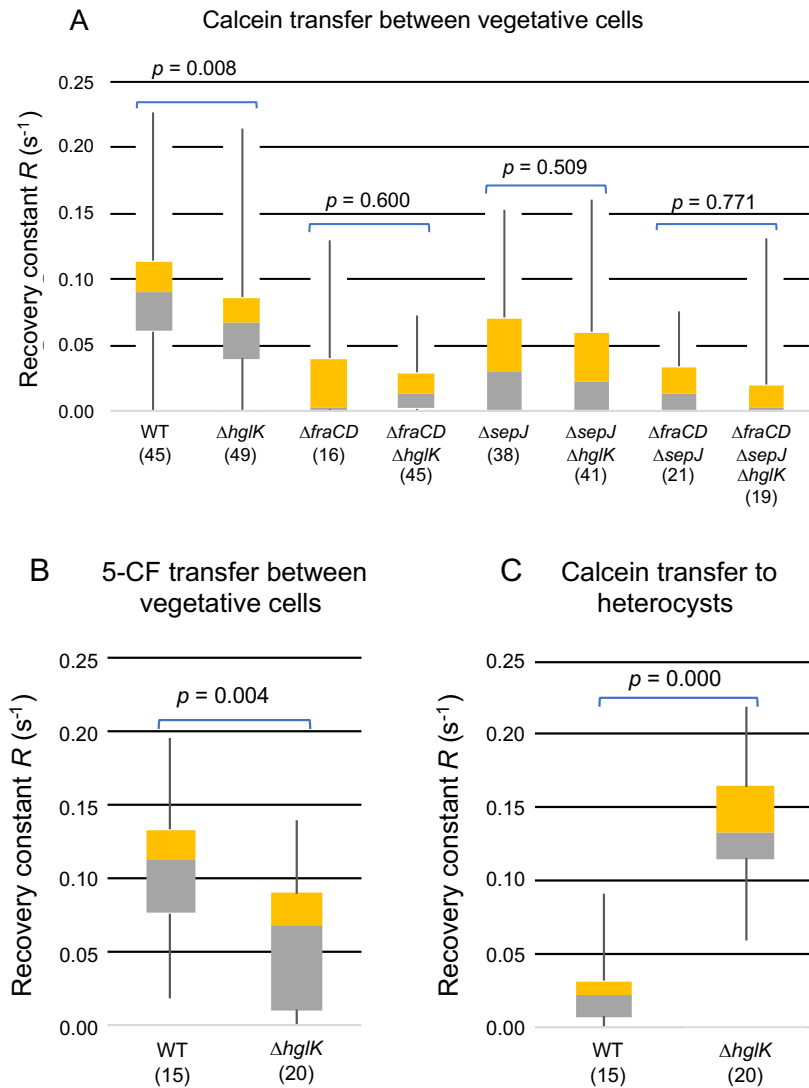
that had been incubated in BG11<sub>0</sub> medium. Filaments of strains CSSA1, CSSA2, CSSA3, CSSA12, CSCT22, CSVM34, CSVM141 and the wild type were grown in liquid BG11 medium (supplemented with neomycin for the strains carrying the  $\Delta hglK::C.K1$  mutation), collected, washed, re-inoculated in liquid BG11<sub>0</sub> medium at a concentration of 1  $\mu\text{g}$  Chl/mL and incubated for 48 h under culture conditions. One mL of each culture was incubated with calcein or 5-carboxyfluorescein (5-CF) and FRAP analysis was carried out as described in Materials and Methods (section 2.7.3). The results obtained are presented in Table 11 and visually in Fig. 35, and the non-parametric Mann-Whitney U test was performed to assess whether the differences observed between strains were significant. The results indicate that transfer of calcein between vegetative cells is decreased by inactivation of *sepJ*, *fraC* and *fraD*, as previously described (Merino-Puerto *et al.*, 2010; Mullineaux *et al.*, 2008). Additionally, transfer of calcein was significantly decreased by inactivation of *hglK*. However, inactivation of *hglK* did not decrease significantly transfer in the *fraC fraD* (CSVT22), *sepJ* (CSVM34) or *fraC fraD sepJ* (CSVM141) backgrounds beyond the decrease caused by these mutations (Fig. 35A, Table 11). The negative effect of inactivation of *hglK* on the molecular exchange between vegetative cells was corroborated by analyzing the transfer of 5-CF, which was also significantly decreased in the mutant (Fig. 35B, Table 11). Thus, *hglK* is required for the normal function of the septal junctions, but its effect is not additive to that of the removal of SepJ or FraC-FraD. In contrast to its negative effect on calcein transfer between vegetative cells, inactivation of *hglK* resulted in a significant increase of calcein transfer from vegetative cells to heterocysts (Fig. 35C, Table 11). This implies the presence of functional junctions in the septa between vegetative cells and heterocysts of the *hglK* mutant.

Fluorescence marker, cell type and genotype	$R$ ( $s^{-1}$ )				Mann-Whitney ( $p$ )
	Parental		+ <i>hglK</i>		
	M	IQR	M	IQR	
<b>Calcein transfer between vegetative cells</b>					
Wild type	0.0901	0.0529	0.0670	0.0469	0.008
$\Delta fraC \Delta fraD$	0.0025	0.0386	0.0126	0.0271	0.600
$\Delta sepJ$	0.0298	0.0700	0.0227	0.0595	0.509
$\Delta fraC \Delta fraD \Delta sepJ$	0.0131	0.0331	0.0028	0.0198	0.771
<b>5-CF transfer between vegetative cells</b>					
Wild type	0.1133	0.0567	0.0681	0.0800	0.003
<b>Calcein transfer to heterocysts</b>					
Wild type	0.021	0.0243	0.077	0.0497	0.000

**Table 11. Intercellular transfer of fluorescent markers in *Anabaena* with different genetic backgrounds.** Cells were grown in BG11 medium (in the presence of Nm for strains carrying the *hglK* mutation), incubated for 48 h in BG11<sub>0</sub> medium without antibiotics, and used in FRAP analysis as described in section 2.7.3 of Materials & Methods. Results of the parental strains and strains with the *hglK* mutation added are shown. M, median; IQP, interquartile range. Each *hglK* mutant was compared to its parental strain by the Mann-Whitney U test ( $p$  indicated in each case).







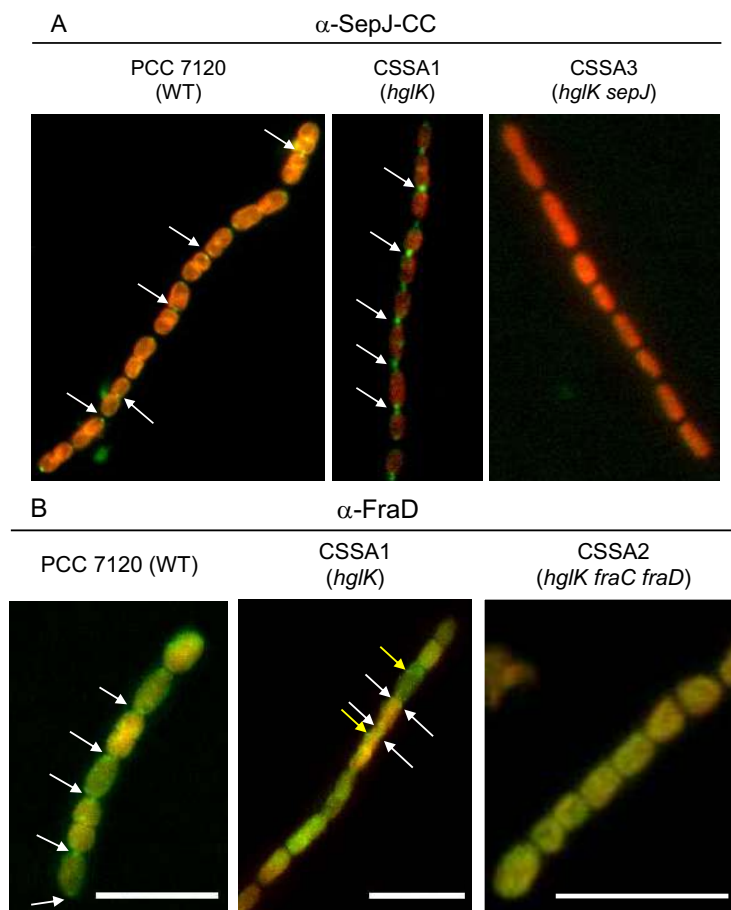
**Fig 35. Effect of the  $\Delta hglK::C.K1$  mutation on the intercellular transfer of fluorescent markers in different genetic backgrounds (box-plot representations of the data in Table 11).** (A) Calcein transfer between vegetative cells of the  $hglK$  mutant in the indicated genetic background and of the corresponding parental strains. (B) 5-carboxyfluorescein (5-CF) transfer between vegetative cells. (C) Calcein transfer from vegetative cells to heterocysts. Data are presented as the recovery rate constant,  $R$ . The number of filaments analyzed is indicated in parenthesis for each strain. Mann-Whitney U test analysis of  $hglK$  versus indicated genetic background was performed and  $p$  is indicated in each case.

### 3.2.4 Subcellular localization of SepJ and FraD

The results shown in the previous section suggest a role of HglK in the septal junctions. We then investigated by immunofluorescence whether the  $\Delta hglK::C.K1$  mutation affected the subcellular localization of septal proteins SepJ and FraD. The strains used for this study were grown in liquid BG11 medium (supplemented with neomycin for the strains carrying the  $hglK$  mutation: CSSA1, CSSA2 and CSSA3), collected, washed and resuspended at a concentration of 1  $\mu\text{g}$  Chl/mL in



liquid BG11<sub>0</sub> medium (without the antibiotic) for 24 hours. For immunolocalization, two types of primary antibodies were used to specifically detect SepJ or FraD: an antibody that recognizes the coiled-coil domain of SepJ (Mariscal *et al.*, 2011) on the CSSA1, CSSA3 and wild-type samples; and an antibody against an extramembrane (periplasmic) fragment of FraD (Merino-Puerto *et al.*, 2011b) on the CSSA1, CSSA2 and wild-type samples. An anti-rabbit antibody carrying a fluorophore was used as a secondary antibody, and the samples were inspected by fluorescence microscopy. For SepJ, septal signals were evident in the wild type and were missing from a negative control strain that lacks the *sepJ* gene, strain CSSA3 ( $\Delta sepJ \Delta hglK::C.K1$ ) (Fig. 36A). In the *hglK* mutant (CSSA1), SepJ was localized at the intercellular septa like in the wild type. Labeling of FraD was unfortunately not as clear as that of SepJ. Nonetheless, FraD spots localized at the intercellular septa could be observed in the wild type that were missing from a *fraD* mutant control, strain CSSA2 ( $\Delta fraC \Delta fraD \Delta hglK::C.K1$ ) (Fig. 36B). In the *hglK* mutant, FraD spots could be observed at the intercellular septa but also out of the septa (Fig. 36B). Quantification of immunofluorescence spots did not show, however, any significant difference between the *hglK* mutant and the wild type in the septal signals of FraD.



**Fig. 36. Subcellular localization of SepJ and FraD in the *hglK* mutant.** Filaments from the indicated strains incubated for 24 h in BG11<sub>0</sub> medium were used in immunofluorescence analysis with antibodies raised against the coiled-coil domain of SepJ (A) or against the extramembrane fragment of FraD (B). White arrows point to some SepJ (A) or FraD (B) spots at intercellular septa; yellow arrows point to some FraD spots outside of the intercellular septa in the *hglK* mutant. Size bars, 10  $\mu$ m.

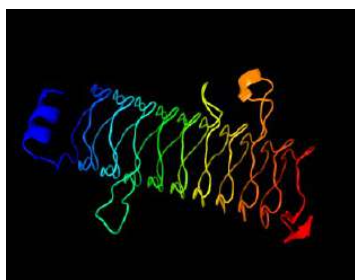


### 3.2.5 Discussion

In this Chapter, we have shown that HglK is a protein predominantly localized in the septal regions of the filaments of *Anabaena*, and that inactivation of the *hglK* gene results in filament fragmentation, a decreased number of septal PG nanopores and decreased activity of intercellular molecular transfer between vegetative cells. However, inactivation of *hglK* does not increase the effects on these parameters that result from inactivation of *sepJ* or *fraC* and *fraD*. Furthermore, inactivation of *sepJ* or *fraC* and *fraD* has a stronger effect on those parameters than inactivation of *hglK* (Nürnberg *et al.*, 2015). Hence, HglK might influence the formation of septal junctions rather than constitute an independent type of junctions. The recently visualized FraD-containing septal junctions evidently comprise additional proteins (Weiss *et al.*, 2019). Whether HglK could be a component of the septal junctions or solely contribute to their assembly remains to be investigated. In any case, HglK could be a structural component of the intercellular septa affecting the formation of septal junctions. In this context, the frequent observation of HglK double spots that are laterally, rather than centrally, localized in the septa (Figs. 25 and 28) is of interest. A role of HglK in the construction of mature septa is consistent with the original observation that inactivation of *hglK* results in the formation of broad intercellular septa (Black *et al.*, 1995).

Our results have shown that inactivation of *hglK* impairs calcein and 5-CF transfer between vegetative cells. In contrast, calcein transfer to heterocysts was increased significantly in the *hglK* mutant. Calcein transfer to the heterocysts appears to be specifically related to SepJ (Mariscal *et al.*, 2016), whose localization is unaltered in the *hglK* mutant (Fig. 36A). The positive effect of inactivation of *hglK* on calcein transfer to the heterocysts can be related to the lack, in the mutant, of cyanophycin plugs (Black *et al.*, 1995), which have been shown to restrict transfer of calcein into the heterocysts (Mullineaux *et al.*, 2008). Whether, additionally, HglK has a regulatory effect on the septal junctions that are functional between vegetative cells and heterocysts is unknown.

As described in the introduction of this Chapter, HglK is a membrane protein with a predicted long pentapeptide-repeat fragment that is likely localized in the periplasm (Fig. 23). As predicted by Phyre2 ([www.sbg.bio.ic.ac.uk/phyre2/](http://www.sbg.bio.ic.ac.uk/phyre2/)), the pentapeptide-repeat fragment of HglK has a structure very similar to that of HetL (Ni *et al.*, 2009). Pentapeptide-repeat proteins adopt a highly regular four-sided, right-handed  $\beta$  helical structure (Fig. 37) that might perform a structural function.



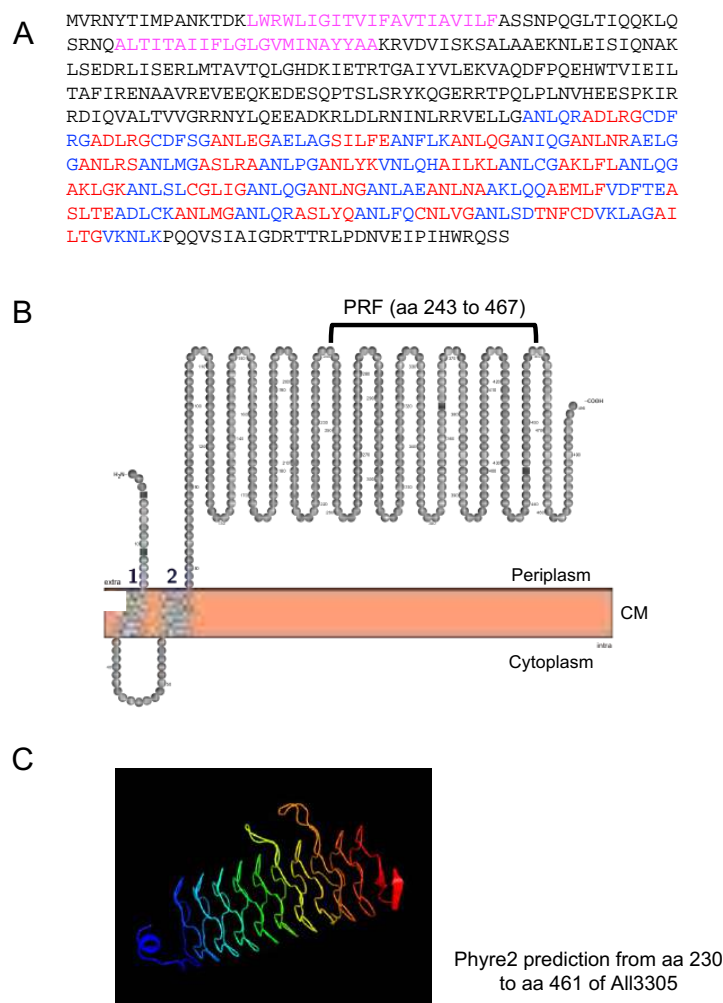
Phyre2 prediction  
from aa 40 to aa  
664 of A110813

**Fig. 37. Three-dimensional structure of the pentapeptide-repeat region of HglK (A110813).** The three-dimensional structure as predicted by Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>) is shown.

There are as many as 32 pentapeptide-repeat proteins in *Anabaena*, and at least five of them bear signal peptides that may direct those proteins to the periplasm (Ni *et al.*, 2009). The limited effect



of inactivation of *hglK* on the parameters that we have studied may reflect that, in *Anabaena*, other proteins have a role redundant with that of HglK. Indeed, inactivation or altered expression of some genes encoding pentapeptide-repeat proteins have been shown to affect filamentation or heterocyst formation related to intercellular molecular transfer in *Anabaena*. Thus, FraF is a pentapeptide-repeat protein encoded in a cluster of genes (*fraC-fraD-fraE* ← *fraF*, where the arrow indicates a change of orientation) that affect filament length (Merino-Puerto *et al.*, 2013), and HetL and PatL are pentapeptide-repeat proteins that affect heterocyst pattern formation (Liu & Golden, 2002; Liu & Wolk, 2011). Interestingly, whereas PatL and FraF are widely distributed in heterocystous cyanobacteria and HetL seems to have a more restricted distribution, HglK is universally present in these organisms suggesting an essential function as mentioned earlier (Table 9). Of further interest, PatL is a predicted membrane-anchored protein that likely has a topology very similar to that of HglK (Fig. 38, compare to Figs. 23 and 37). Whether PatL and HglK have somewhat associated or redundant functions remains to be investigated



**Fig. 38. Overview of the PatL (A113305) protein.** (A) Amino acid sequence with indication of transmembrane segments (magenta) and pentapeptide repeats (alternatively shown in blue and red color), as described by Liu & Wolk (2011). (B) Protter representation based on prediction with the Phobius



program (<http://phobius.sbc.su.se>). The approximate localization of the pentapeptide-repeat fragment (PRF) is indicated. CM, cytoplasmic membrane. (C) Three-dimensional structure of the pentapeptide repeat region as predicted by Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>).

Because the *hglK* mutant was shown to synthesize heterocyst-specific glycolipids (Hgl) but not to deposit an Hgl layer, HglK was suggested to be involved in the export of Hgl or deposition of the Hgl layer (Black *et al.*, 1995). However, the transporter of Hgl to the outside of the outer membrane has now been identified as the ATP-driven efflux pump DevBCA/HgdD (Staron *et al.*, 2011). Here we have shown, on the other hand, that the *hglK* mutant is impaired in heterocyst differentiation, failing to express at normal levels some genes activated at intermediate and late steps of the differentiation process. It is possible, therefore, that rather than being specifically involved in Hgl export or deposition, HglK has a more general role in heterocyst differentiation, its inactivation resulting in a deficiency of several differentiation-related processes. Because HglK appears to influence the formation of the septal junctions, impairment, in the *hglK* mutant, in the intercellular transfer of some compounds could be responsible for the phenotype of deficient differentiation. This would be similar to the situation in the *sepJ* mutant, in which heterocyst differentiation is aborted after synthesis of the heterocyst polysaccharide layer but before synthesis of Hgl (Flores *et al.*, 2007; Nayar *et al.*, 2007). SepJ appears to be involved in the intercellular transfer of the PatS morphogen and of a HetN-dependent regulator of heterocyst differentiation (Mariscal *et al.*, 2016; Rivers *et al.*, 2014), but the physiological compounds whose intercellular transfer may be affected by inactivation of *hglK* are currently unknown.

In summary, our results have shown that the pentapeptide-repeat protein HglK is involved in the formation of fully functional septal junctions and is needed for the differentiation of functional heterocysts in *Anabaena*. The function of a number of pentapeptide-repeat proteins in different aspects of heterocyst differentiation is intriguing and raise the possibility that some of those proteins are (at least partially) redundant for the formation of the septal junctions, which are complex entities that appear to include a number of different proteins.



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



## 3. Results & Discussion

---

### Chapter 3

### HgIK and SepJ proteomics

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**





ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

### 3.3 HglK and SepJ proteomics

The pentapeptide-repeat protein HglK from *Anabaena* is needed for the formation of mature heterocysts (Black *et al.*, 1995). According to bioinformatics predictions, HglK has four transmembrane segments with its amino and carboxyl termini located in the periplasm (Fig. 23). It is the product of the *hglK* gene, which is low-expressed under normal conditions and induced after N-stepdown (Ehira & Ohmori, 2006a; Flaherty *et al.*, 2011). As shown in Results & Discussion Chapter 2, HglK is predominantly located in the cell poles, and it is necessary for the integrity of the filament and for intercellular molecular transfer as probed by fluorescent markers.

SepJ is also a membrane integral protein. Our current model, based on bioinformatics predictions and experimental evidence (Rudolf *et al.*, 2015; Herrero *et al.*, 2016; Ramos-León *et al.*, 2018), considers that SepJ from *Anabaena* contains 11 transmembrane segments leaving the amino terminal part (about 410 amino acid residues) in the periplasm and the carboxyl terminus in the cytoplasm (Fig. 7). SepJ is located in the septum between adjacent cells and, as observed for *hglK* mutants, *sepJ* mutants are impaired in diazotrophic growth, exhibit filament fragmentation and show decreased intercellular transfer of fluorescent markers (Flores *et al.*, 2007; Mullineaux *et al.*, 2008; Mariscal *et al.*, 2011; Ramos-León *et al.*, 2018). SepJ has coiled-coil domains, which are usually involved in protein-protein interactions, and engages in protein complexes that have been observed by blue native (BN)-PAGE (Ramos-León *et al.*, 2017).

Because of the importance of the two proteins for the integrity and function of the cyanobacterial filament, it is possible that they are part or participate in the formation of the protein complexes that constitute the septal junctions. In this Chapter, the objective was to isolate these proteins avoiding denaturation in order to detect possible purification partners by mass spectrometry.

#### 3.3.1 Isolation the HglK and SepJ proteins

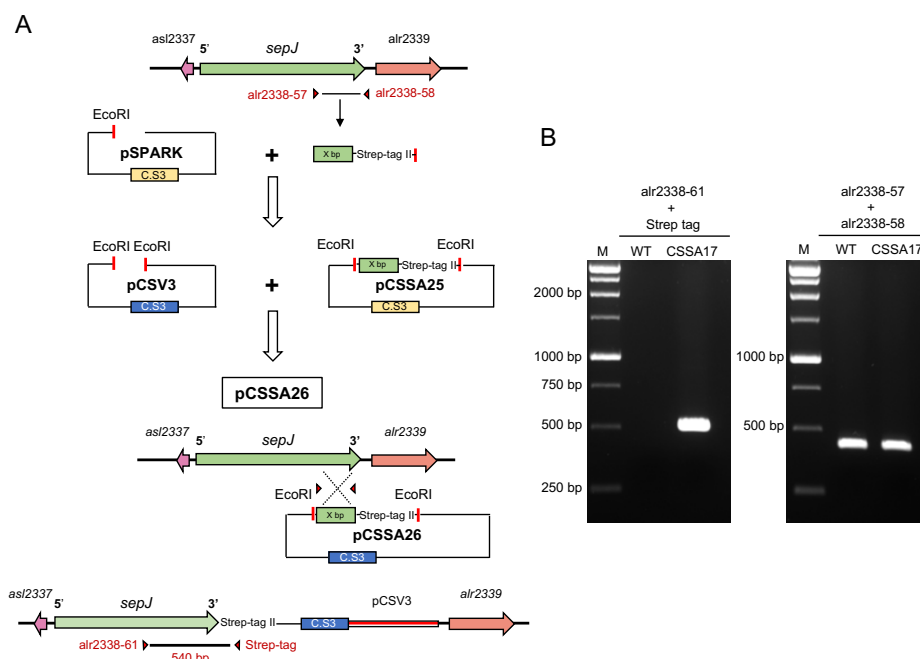
To facilitate the isolation HglK and SepJ, we prepared *Anabaena* strains producing these proteins fused to Strep II or GFP tags. For the isolation using the Strep-tag II, *Anabaena* strain CSSA18 that produces the C-terminally Strep II-tagged HglK (described in Fig. 27) and strain CSSA17 that produces the C-terminally Strep II-tagged SepJ (to be described in the next section) were used. The isolation using GFP antibodies was carried out with *Anabaena* strain CSAM137 (described in Table 2 of Materials & Methods) that produces a SepJ with a deletion of the last nine amino acids at the C-terminus and a C-terminal GFP fusion. An *Anabaena* strain bearing this construct has been shown to exhibit normal SepJ activity (Ramos-León *et al.*, 2017).

##### 3.3.1.1 Isolation using the Strep-tag II

For isolation of proteins with Strep-tag II, the two strains CSSA17 and CSSA18 were used. Strain CSSA17 was constructed by cloning 420 bp from the 3' end of the *sepJ* gene to facilitate recombination of an appropriate construct into the *Anabaena* genome. The amplification was carried out by PCR with primers alr2338-57 and alr2338-58 producing a DNA fragment of 432 bp that includes the 3' part of the *sepJ* gene (except the stop codon) and sequences encoding the Strep-tag II and the target for EcoRI at its 3' end. This fragment was inserted into the pSPARK cloning vector forming plasmid pCSSA25, which was corroborated by sequencing. The construct was extracted from pCSSA25 by EcoRI restriction and introduced into plasmid pCSV3 (which



encodes Sm<sup>R</sup> and Sp<sup>R</sup>) originating plasmid pCSSA26, which was transferred to *Anabaena* by conjugation giving rise to strain CSSA17. The exconjugants obtained were checked by PCR with primers alr2338-60 and Strep tag (reverse), and one exconjugant was chosen as strain CSSA17 (Fig. 39). Construction of strain CSSA18 (HglK-Strep-tag II) is described in Chapter 2 (Fig.27)



**Fig. 39. Construction of *Anabaena* strain CSSA17 that produces SepJ-Strep-tag II.** (A) The schemes of the plasmid generation strategy (top) and of the final genomic structure of the mutant strain (bottom) show the primers used for construction of CSSA17 as well as those used to check the exconjugant by PCR. The schemes also show the targets for the restriction enzyme used. X refers to 432 bp. (B) Electrophoresis of products of the PCRs performed on genomic DNA to check the genotype of strain CSSA17 with the wild type as a control (WT); the primer pair used is indicated above each image. As shown in the PCR performed with primers alr2338-61 and Strep tag (reverse), there is no amplification in the WT since it lacks the sequence of the Strep-tag II. M, size markers.

The strains CSSA17, CSSA18 and wild-type *Anabaena* as a negative control were used to carry out the purification with the Strep-tag II (for full details see Section 2.5.2 in Materials & Methods). The strains were grown in 1.5-L cultures of bubbled liquid BG11C medium supplemented with antibiotics for the strains producing the fusion proteins. Cells were collected by centrifugation and washed with BG11<sub>0</sub>C medium to be finally resuspended in BG11<sub>0</sub>C medium (without antibiotics) to perform an induction for 48 hours without combined nitrogen with bubbling. After induction, the samples were collected by centrifugation and resuspended in lysis buffer, and then they were disrupted by gridding in a mortar with liquid nitrogen. The material was supplemented with 1% DDM and 1 mg/mL of lysozyme to extract membrane proteins and eliminate the cell wall. The material was then incubated with Strep-Tactin magnetic beads. After incubation, the magnetic beads were subjected to different washing steps, and about 84% of the washed beads were stored at -80°C for further processing by mass spectrometry. From the



remaining beads (approximately 16%) proteins were eluted by addition of a buffer containing biotin.

To probe for the quality of the procedure and quantity of the eluted proteins, the lysate before incubation with magnetic beads (input), the flow-through after incubation with beads (not bound proteins), the different washes (wash 1, 2, etc.), the different fractions after elution with biotin (eluted) and the elution with sample buffer (SDS/loading buffer) were collected. These samples were subjected to SDS-PAGE to separate the different proteins. After electrophoresis, proteins were transferred to a PDVF membrane. The membrane was then stained with DB71 (a dye that stains the proteins transferred to the membrane) to test transfer efficiency. In the membrane corresponding to the Strep-Tactin-based purification using extracts of the wild type (Fig. 40A), the molecular weight markers and the lanes corresponding to the input (I) and flow-through (FT) samples were properly stained. Some proteins were also observed in the first wash (W1), which correspond to traces of the most abundant *Anabaena* proteins. Similarly, in the membranes stained with DB71 corresponding to the purification of SepJ-Strep-tag II and HglK-Strep-tag II (Figs. 40C and E respectively), staining was only seen in the lanes corresponding to the molecular size markers, input, flow-through and first wash (the input aliquot was not loaded in Fig. 40C). A staining of the two bait-proteins (SepJ, HglK) was not expected based of the low abundance of these proteins and the low sensitivity of DB71 staining.

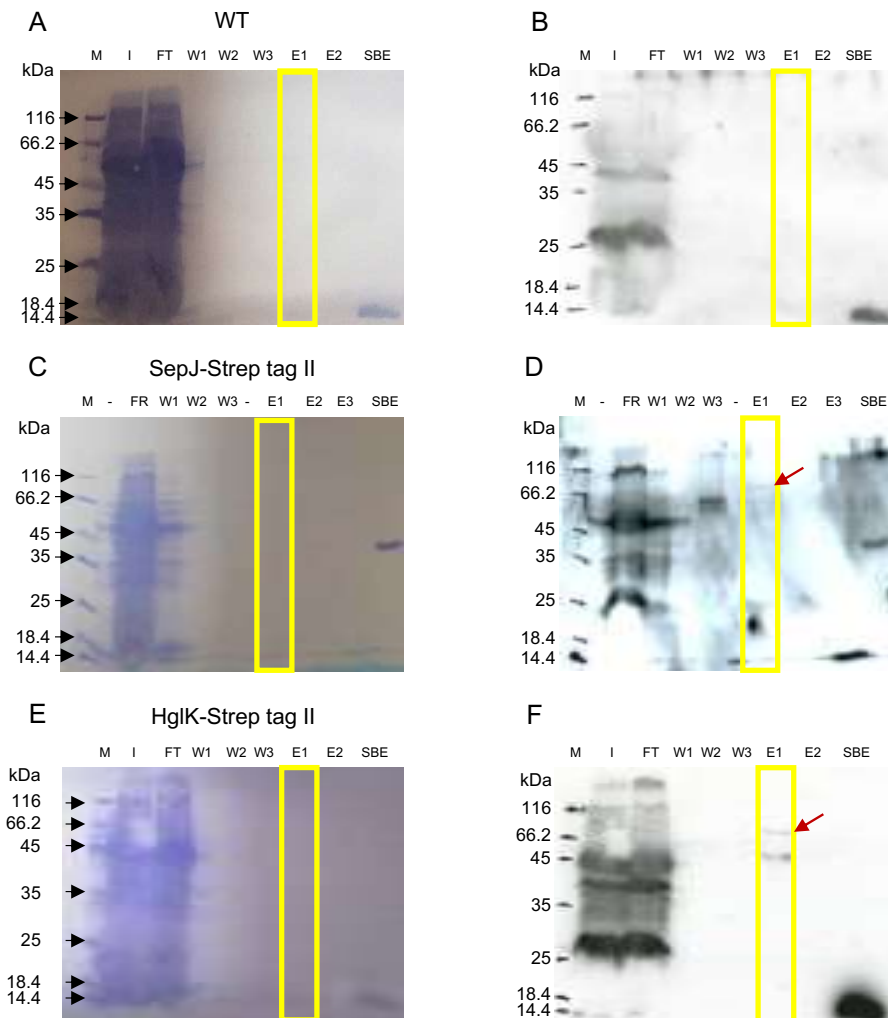
After confirming the efficiency of the transfer of proteins, the membranes were used for *western blotting* (Section 2.5.4.4 of Materials & Methods). In the membrane corresponding to the Strep-Tactin-based purification using the wild-type extracts (Fig. 40B), there were only some bands in the lanes corresponding to the input and flow-through aliquots, probably seen because of high concentration of protein non-specifically reacting with the antibody. Lanes E1 and E2 contained the aliquots eluted from the magnetic beads, which were completely clean with no protein detected, since the wild type lacks the Strep-tag II. A band corresponding to a small protein was seen in the lane corresponding to elution of the magnetic beads with SDS/loading buffer, which may be streptavidin (monomer MW about 13,750 Da) released from Strep-Tactin.

The *western blot* performed on the membrane of the samples from the SepJ-Strep-tag II isolation (see Fig. 40D) showed in the biotin-eluted fraction E1 a protein migrating between 66.2-kDa and 116-kDa (red arrow) consistent with the molecular weight of SepJ-Strep-tag II (81.5 kDa). In the flow-through (lane FT) and first washing (lane W1) samples, non-specific bands were observed that should correspond to proteins reacting non-specifically with the antibody. In the third washing (lane W3), two bands of unknown origin were observed. Finally, two proteins were only eluted from magnetic beads after addition of SDS/loading buffer (lane SBE). One migrates at a molecular weight larger than 116 kDa and the second with a molecular weight of about 40 kDa. Remarkably, the protein was also stained with DB71, and therefore the staining with Strep-Tactin horseradish peroxidase conjugate could be due to non-specific binding caused by a large accumulation of protein.

The *western blot* carried out on the membrane with the different fractions of the HglK-Strep-tag II isolation (see Fig. 40F) showed in the biotin-eluted fraction E1 a band between 66.2 kDa and 116 kDa (red arrow) that likely corresponds to HglK-Strep-tag II (78.6 kDa). A smaller protein between the 45-kDa and 62.2-kDa markers was detected as well. This could result from breakage of HglK somewhere in its amino terminal part. In input and first wash samples (lanes I and W1,



respectively), proteins were observed that may result again from non-specific binding of the antibody to large amounts of protein. Finally, a protein of very small size was seen in the material eluted with SDS/loading buffer (SBE) that may be streptavidin monomer detached from Strep-Tactin.



**Fig. 40. Staining with BD71 and western blot analysis from Strep-tag II isolations.** The images on the left (A, C, E) show the staining of the different membranes with DB71, and the images on the right (B, D, F) show the western blots on the corresponding membranes using Strep-Tactin horseradish peroxidase conjugate to detect the Strep-tag II. A and B correspond to the isolation of the Strep-tag II with the magnetic beads from *Anabaena* wild-type (WT) extracts, used as a negative control. C and D show the results of the Strep-tag II isolation from strain CSSA17 (SepJ-Strep-tag II) extracts, and images E and F show the results of the Strep II tag isolation from strain CSSA18 (HglK-Strep-tag II) extracts. The yellow box indicates the lane that carries the aliquot from the first biotin elution of the Strep-tag II from the magnetic beads, and the red arrow indicates the band that can correspond to SepJ-Strep-tag II (D) or HglK-Strep-tag II (F). The numbers on the left indicate the size of the markers. SepJ-Strep-tag II has a size of 81.48 kDa while HglK-Strep-tag II has a size of 78.58 kDa. In the images, letters correspond to: M, molecular weight marker; I, input; FT, flow-through material; W1, washing 1; W2, washing 2; W3, washing 3; E1, biotin-eluted fraction 1; E2, biotin-eluted fraction 2; SBE, sample buffer-eluted fraction. (-) Empty lane.



### 3.3.1.2 Isolation using SepJ-GFP

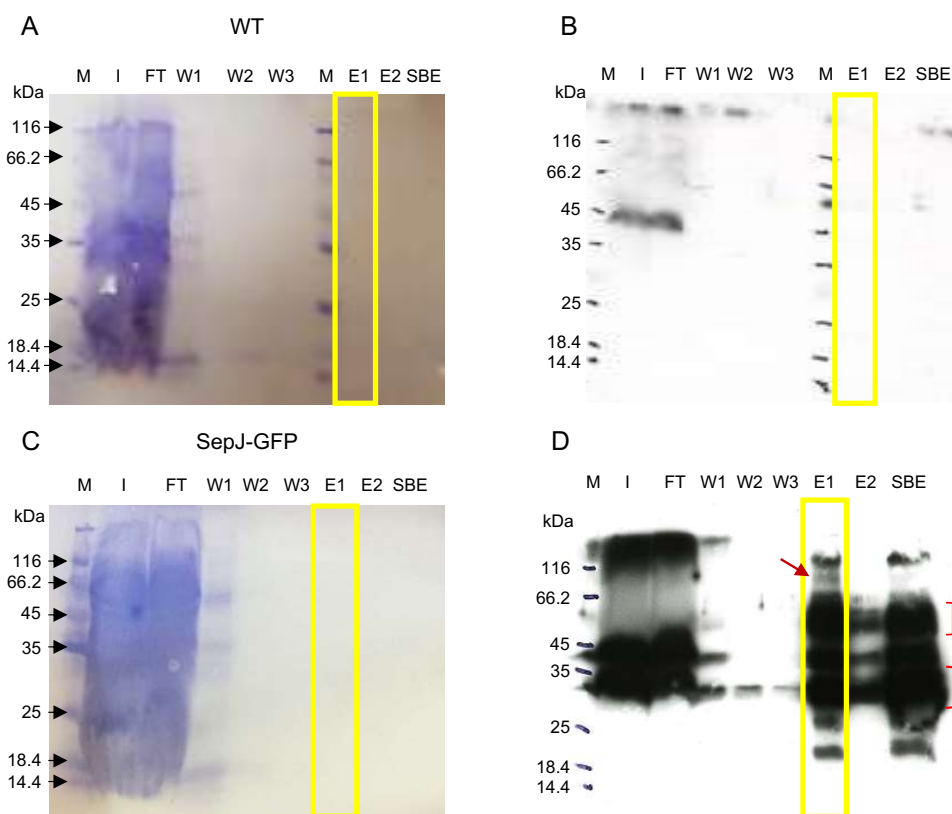
An alternative approach to identify possible protein-complex partners of SepJ was to isolate SepJ-GFP in non-denaturing conditions. Strain CSAM137 that carries the GFP anchored to the C-terminal part of SepJ (described in Table 2 of Materials & Methods) was used, with the wild type as a negative control. Both strains were grown in 1.5-L cultures of bubbled liquid BG11C medium (adding the corresponding antibiotics to the CSAM137 strain). After growth, the filaments were induced for 48 h in BG11<sub>0</sub>C medium (induction under combined nitrogen deprivation). After induction, the samples were collected, resuspended in lysis buffer, disrupted by gridding in a mortar with liquid nitrogen, and treated with DDM and lysozyme. The samples were then incubated with anti-GFP magnetic beads, which were then collected and washed. About 76% of the washed beads were stored at -80°C until processed for mass spectrometry (see details of sample preparation and purification in sections 2.5.1 and 2.5.3 of Materials & Methods). The remaining beads were used to elute the protein with glycine-buffer to check the efficiency of the purification process before carrying out the analysis by mass spectrometry. Aliquots from each of the washes and from the sample before and after its incubation with the magnetic beads were checked by *western blot* analysis together with the glycine-eluted sample.

The proteins of all fractions were separated by SDS-PAGE and transferred to a PDVF membrane followed by *western blotting*. After transfer, the membrane was stained with DB71 to control for efficient protein transfer. The transfer on the membranes corresponding to the aliquots from wild type and CSAM137 purifications (Fig. 41A and C, respectively) was efficient as indicated by the presence of protein markers and proteins in input, flow-through and first wash (lanes I and W1). On the membrane with samples from the wild-type control (Fig. 41A), proteins were only observed in input (I), flow-through (FT) and first wash (lane W1) samples. On the membrane with the different aliquots of the purification from strain CSAM137 (Fig. 41C), in the input and flow-through samples (lanes I and FT, respectively) a large continuous stain indicates a high concentration of protein. On the other hand, in the first wash fraction (lane W1), the protein concentration was lower allowing to differentiate some specific bands, which were similar to those observed in the wild-type control. Collectively, these results confirm protein transfer to the membranes and a lack of binding of substantial amounts of proteins by the anti-GFP antibody.

After incubation with an antibody specific against the GFP and chemiluminiscent development (details of the *western blotting* in Section 2.5.4.5, Materials & Methods) one protein at approximately 45 kDa was detected in input and flow-through material of wild-type extracts (Fig. 41B, lanes I, FT), but other proteins were not detectable. The development of the membrane with the different fractions of the SepJ-GFP purification (Fig. 41D) showed several bands in I, FT and W1, likely reflecting non-specific reaction of the antibody with large amounts of protein. In the first fraction of the glycine buffer-eluate from the magnetic beads (lane E1) a band migrating between the 66.2-kDa and 116-kDa markers (red arrow) that likely corresponds to SepJ-GFP (108.5 kDa) was observed. A protein migrating at the same molecular weight could also be observed in lane SBE, which contains the eluate with SDS/loading buffer. This suggests that the elution of SepJ-GFP with glycine was not entirely efficient. Additional proteins of different sizes appear in lanes E1 to SBE. These proteins can largely correspond to GFP (27 kDa), putatively also including stable GFP dimers, released from SepJ-GFP. Additionally, some of this material could correspond to other degradation products of SepJ-GFP (Ramos-León *et al.*, 2017).







**Fig. 41. Staining with BD71 and western blot analysis of materials from the SepJ-GFP isolation.** The images on the left (A, C) show the staining of the membranes with DB71, and the images on the right (B, D) show the corresponding *western blot* analysis using monoclonal antibodies against GFP to detect the GFP tag. A and B correspond to the isolation with the anti-GFP magnetic beads using *Anabaena* wild type (WT) as a negative control. C and D show the isolation with the anti-GFP magnetic beads using strain CSAM137 (SepJ-GFP) extracts. The yellow box indicates the lane in which the aliquot from the glycine-buffer elution of proteins from the magnetic beads was loaded, and the red arrow indicates the band that can correspond to SepJ-GFP. The numbers on the left indicate protein marker sizes. SepJ-GFP has a size of 108.48 kDa. In the images, letters correspond to: M, molecular weight marker; I, input; FT, flow-through material; W1, washing 1; W2, washing 2; W3, washing 3; E1, glycine-eluted fraction 1; E2, glycine-eluted fraction 2; SBE, sample buffer-eluted fraction. Red brackets denote the position of the GFP (bottom) and possible GFP dimers (top).

### 3.3.2 Mass spectrometry results

As starting material, the Strep-Tactin or anti-GFP magnetic beads that had been stored at  $-80^{\circ}$  after the purification process were taken for subsequent analysis. The samples were processed as described by Brouwer *et al.* (2019) to produce the peptides that were subjected to mass spectrometry. Identification with SepJ as bait was performed three times (each time corresponding to an individual experiment): one time with Strep II tagged protein and two times with GFP tagged protein. On the other hand, the identification with HglK was performed two times with Strep II tagged protein. For SepJ, each of the individual experiments was repeated three times with corresponding repetitions of the control experiments. For HglK, experiment 1





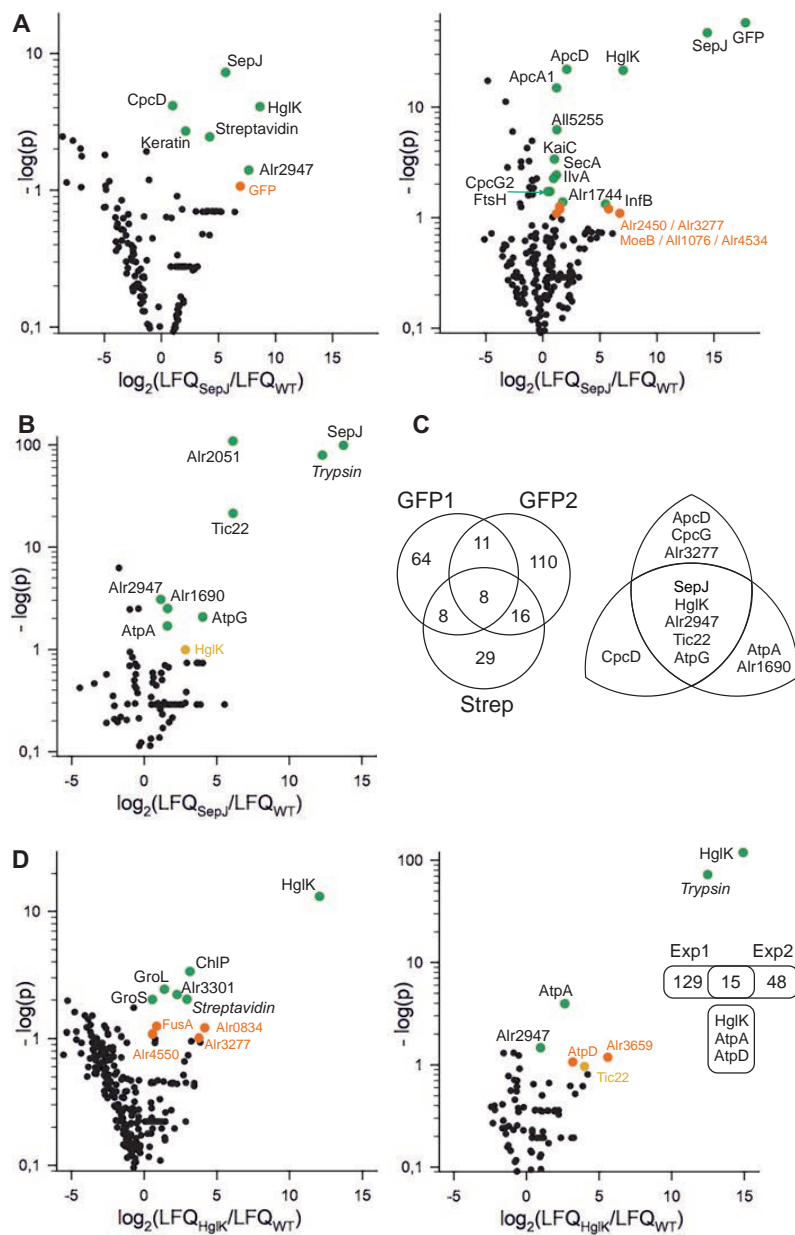
had four repetitions and experiment 2 had five repetitions; each of these experiments had three repetitions of the wild-type controls. (Please click in this link [Proteomic Data](#) to see the Excel file with all proteins detected in these analyses.) For all experiments, the  $\log_2$  value of the fold change of the label-free quantification (LFQ) value between the bait-dependent experiment and the control experiment was calculated, and the p-value was determined. Because both SepJ and HglK are low-abundance membrane proteins, a high degree of contamination may be expected in the samples. This made us to consider proteins with a p-value lower than 0.1 instead of the more restrictive usual  $p < 0.05$  value.

In the analysis of the first experiment of SepJ-GFP (GFP 1), a total of 91 proteins were identified to be higher enriched in the bait-fraction than in the control fraction (see image on the left in Fig. 42A); in the second experiment of SepJ-GFP (GFP 2), 145 proteins were detected (right image in Fig. 42A); and finally, in the analysis of the purified extracts from the SepJ-Strep-tag II sample, 61 proteins were detected (Fig. 42B). After analyzing the results obtained in each experiment, Tables 12 to 14 were created, where all the proteins with a p-value lower than 0.1 are summarized.

To check whether any proteins were identified in the three experiments, an overlap analysis was performed with all the proteins that had been detected in each experiment of SepJ (two experiments from SepJ-GFP and one from SepJ-Strep-tag II) (Fig. 42C). The result of this overlap was that 43 proteins appeared in at least two experiments and 8 proteins were identified in all three analyses. Of the eight proteins detected in all three experiments using SepJ as bait, only four proteins other than SepJ appeared identified with a p-value lower than 0.1: HglK and Alr2947 with  $p < 0.1$  in two experiments, and Tic22 and AtpG with  $p < 0.1$  in one experiment, while the other three proteins did not appear significantly identified in any of the experiments performed. Analyzing the 43 proteins found in at least two independent experiments yielded the following proteins with a p-value lower than 0.1 in at least one of the experiments: CpcD (GFP 1/SepJ-Strep-tag II), ApcD (GFP 1/GFP 2), CpcG (GFP 1/GFP 2), Alr3277 (GFP 1/GFP 2), AtpA (GFP 2/ SepJ-Strep-tag II) and Alr1690 (GFP 2/ SepJ-Strep-tag II).

In reference to the two experiments performed with HglK-Strep-tag II, 144 proteins were detected in the first experiment and 61 proteins in the second experiment to be more abundant in the bait-fraction than in the control fraction (Fig. 42D and Excel file at [Proteomic Data](#)). Of all the proteins detected in each experiment those that were identified with a p-value lower than 0.1 are presented in Tables 15 and 16. As in the case of SepJ, we analyzed the overlap between both experiments. The overlap included 15 proteins of which only three had a p-value lower than 0.1 in at least one of the two experiments. These proteins were HglK, AtpA and AtpD.





**Fig. 42. Analysis of the proteomics results.** Volcano plots for the  $\log_2$ -value of the ratio between the label-free quantification (LFQ) value of protein detection using the strain expressing the bait compared to the wild-type strain as negative control, with the calculated p-value. (A) Shows the results for SepJ-GFP experiment 1 (left) and 2 (right), (B) SepJ-Strep-tag II and (D) HglK-Strep-tag II experiment 1 (left) and 2 (right). Black dots show all identified proteins with  $0.15 < p < 0.815$ ; green dots all identified proteins with  $p < 0.05$ , orange dots with  $0.05 < p < 0.1$  and yellow dots with  $0.1 < p < 0.15$ , the latter for proteins that have been identified in one of the other experiments. The p-value 0.1 was selected because the bait proteins are low-abundant. (C) The overlap of the three SepJ experiments is shown (left). The proteins that are found in at least two experiments and with a  $p < 0.1$  in at least one experiment are indicated. In D, the overlap of the two experiments is shown on the right and the proteins identified with  $p < 0.1$  in the overlap are indicated.



Acc.	ORF	Protein	Bait/Control		Further information
			log <sub>2</sub>	p-value	
Q8YSY2	<i>alr2947</i>		8.63	8.4*10 <sup>-5</sup>	Signal peptide (SPI) containing protein
Q44230	<i>all0813</i>	HglK	7.65	0.04	Chapter 2
Q8YUK6	<i>alr2338</i>	SepJ	5.63	5.8*10 <sup>-8</sup>	Bait
Streptavidin			4.32	3.6*10 <sup>-3</sup>	Streptavidin
P13647		Krt5	2.12	2.0*10 <sup>-3</sup>	Keratin, type II cytoskeletal 5
P07124	<i>asr0531</i>	OpcD	0.99	7.3*10 <sup>-5</sup>	Phycobilisome 8.5-kDa linker polypeptide
Q9U6Y5		GFP	6.92	0.09	

**Table 12. Interaction partners identified with SepJ-GFP as bait —Experiment 1.** Three independent repetitions of SepJ-GFP isolation and three control experiments were analyzed by MaxQuant to calculate the LFQ value for analysis. The accession number (column 1), the ORF (column 2), the protein name (column 3), the log<sub>2</sub> value and the p-value for the LFQ ratio between bait and control (column 4 & 5), and the assigned function (when appropriate) is presented for the proteins identified with  $p < 0.05$  (top proteins) and one with  $0.05 < p < 0.1$  (bottom protein). SPI, signal peptidase I-type of signal peptide.

Acc.	ORF	Protein	Bait/Control		Further information
			log <sub>2</sub>	p-value	
Q9U6Y5		GFP	17.77	0.00	Bait tag
Q8YUK6	<i>alr2338</i>	SepJ	14.43	0.00	Bait
Q44230	<i>all0813</i>	HglK	2.10	1.2*10 <sup>-22</sup>	Chapter 2
P80556	<i>all3653</i>	ApcD	7.04	3.3*10 <sup>-22</sup>	Allophycocyanin subunit alpha-B
P80555	<i>alr0021</i>	ApcA1	1.21	1.3*10 <sup>-15</sup>	Allophycocyanin subunit alpha 1
Q8YLP1	<i>all5255</i>		1.23	6.0*10 <sup>-7</sup>	Polyphenol oxidase
Q8YT40	<i>alr2886</i>	KaiC	1.02	4.4*10 <sup>-4</sup>	Circadian clock protein kinase KaiC
Q8YMS8	<i>alr4851</i>	SecA	1.19	3.8*10 <sup>-3</sup>	Protein translocase subunit SecA
Q8YPG2	<i>alr4232</i>	IlvA	0.95	0.01	L-threonine dehydratase
P29987	<i>alr0535</i>	CpcG2	0.63	0.02	Phycobilisome rod-core linker polypeptide
Q8YXF2	<i>alr1261</i>	FtsH	0.44	0.02	ATP-dependent zinc metalloprotease with TMD
Q8YW72	<i>alr1744</i>		1.75	0.04	Sulfolipid biosynthesis protein
Q8YQJ1	<i>alr3832</i>	InfB	5.49	0.05	Translation initiation factor IF-2
Q8YUA2	<i>alr2450</i>		1.48	0.05	1,4-alpha-glucan branching enzyme activity
Q8YS15	<i>alr3277</i>		5.77	0.06	ATP-dependent RNA/DNA helicases
Q8YT21	<i>all2906</i>	MoeB	1.48	0.06	Molybdopterin biosynthesis protein
Q8YXY0	<i>all1076</i>		1.19	0.08	DNA-binding domain containing protein
Q8YNN0	<i>alr4534</i>		6,76	0.08	Coiled-coil domain containing protein with TMD

**Table 13. Interaction partners identified with SepJ-GFP as bait —Experiment 2.** Three independent repetitions of SepJ-GFP isolation and three control experiments were analyzed by MaxQuant to calculate the LFQ value for analysis. The accession number (column 1), the ORF (column 2), the protein name (column 3), the log<sub>2</sub> value and the p-value for the LFQ ratio between bait and control (column 4 & 5), and the assigned function (when appropriate) is presented for the proteins identified with  $p \leq 0.05$  (top proteins) and  $0.05 < p < 0.1$  (four lower proteins).



Acc.	ORF	Protein	Bait/Control		Further information
			log <sub>2</sub>	p-value	
Q8YVC7	<i>alr2051</i>		6.11	0.00	Gamma-glutamyl-transpeptidase with transmembrane domain
Q8YUK6	<i>alr2338</i>	SepJ	13.73	0.00	Bait
P00761		Trypsin	12.28	7.9*10 <sup>-110</sup>	
Q8Z0I2	<i>alr0114</i>	Tic22	6.12	4.5*10 <sup>-22</sup>	Periplasmic protein translocase
Q8YSY2	<i>alr2947</i>		1.13	8.8*10 <sup>-4</sup>	Signal peptide (SPI) containing protein
Q8YWC3	<i>alr1690</i>		1.60	3.2*10 <sup>-3</sup>	Cell wall-binding protein
P12410	<i>all0008</i>	AtpG	4.03	8.6*10 <sup>-3</sup>	ATP synthase subunit b
P12405	<i>all0005</i>	AtpA	1.59	0.02	ATP synthase subunit alpha
Q44230	<i>all0813</i>	HglK	2.83	0.11	Chapter 2

**Table 14. Interaction partners identified with SepJ-Strep-tag II as bait.** Three independent repetitions of SepJ-Strep-tag II isolation and three control experiments were analyzed by MaxQuant to calculate the LFQ value for analysis. The accession number (column 1), the ORF (column 2), the protein name (column 3), the log<sub>2</sub> value and the p-value for the LFQ ratio between bait and control (column 4 & 5), and the assigned function (when appropriate) is presented for the proteins identified with  $p < 0.05$  (top proteins) and one with  $p = 0.11$  (bottom protein). SPI, signal peptidase I-type of signal peptide.

Acc.	ORF	Protein	Bait/Control		Further information
			log <sub>2</sub>	p-value	
Q44230	<i>all0813</i>	HglK	12.05	7.6*10 <sup>-14</sup>	Bait
Q8Z0G8	<i>alr0128</i>	ChlP	3.15	4.5*10 <sup>-5</sup>	Geranyl-geranyl hydrogenase
Streptavidin			2.94	9.6*10 <sup>-3</sup>	Streptavidin
Q8YRZ1	<i>alr3301</i>		2.63	1.2*10 <sup>-4</sup>	Coiled-coil domain containing protein
Q8YQZ8	<i>alr3661</i>	GroL	1.39	3.8*10 <sup>-3</sup>	60-kDa chaperonin
Q8YQZ9	<i>alr3661</i>	GroS	0.57	9.7*10 <sup>-3</sup>	10-kDa chaperonin
Q8YYL4	<i>alr0834</i>		4.16	0.06	Porin —major outer membrane protein
Q8YS15	<i>alr3277</i>		3.76	0.10	ATP-dependent RNA/DNA helicases
Q8YP62	<i>all4338</i>	FusA	0.86	0.06	Elongation factor G
Q8YNL5	<i>alr4550</i>		0.58	0.08	Porin —major outer membrane protein

**Table 15. Interaction partners identified with HglK-Strep-tag II as bait —Experiment 1.** Four independent repetitions of HglK-Strep-tag II isolation and three control experiments were analyzed by MaxQuant to calculate the LFQ value for analysis. The accession number (column 1), the ORF (column 2), the protein name (column 3), the log<sub>2</sub> value and the p-value for the LFQ ratio between bait and control (column 4 & 5) and the assigned function (when appropriate) is presented for the identified proteins with  $p < 0.05$  (top proteins) and  $0.05 < p \leq 0.1$  (four lower proteins).



Acc.	ORF	Protein	Bait/Control		Further information
			log <sub>2</sub>	p-value	
Q44230	<i>all0813</i>	HglK	14.91	0.00	Bait
P00761		Trypsin	12.47	3.8*10 <sup>-173</sup>	
P12405	<i>all0005</i>	AtpA	2.63	1.2*10 <sup>-4</sup>	ATP synthase subunit
Q8YSY2	<i>alr2947</i>		0.95	0.03	Signal peptide (SPI) containing protein
Q8YR01	<i>alr3659</i>		5.59	0.07	Hemolysin-tyr Ca-binding domain
P06540	<i>all5093</i>	AtpD	3.18	0.09	ATP synthase subunit beta
Q8Z0I2	<i>alr0114</i>	Tic22	3.98	0.11	Periplasmic protein translocase

**Table 16. Interaction partners identified with HglK-Strep-tag II as bait —Experiment 2.** Five independent repetitions of HglK-Strep-tag II isolation and three control experiments were analyzed by MaxQuant to calculate the LFQ value for analysis. The accession number (column 1), the ORF (column 2), the protein name (column 3), the log<sub>2</sub> value and the p-value for the LFQ ratio between bait and control (column 4 & 5) and the assigned function (when appropriate) is presented for the identified proteins with  $p < 0.05$  (top proteins) and  $0.05 < p \leq 0.11$  (three lower proteins). SPI, signal peptidase I-type of signal peptide.

### 3.3.3 Discussion

SepJ and HglK are proteins exclusively or predominantly located at the cell poles in the intercellular septa of *Anabaena* and whose mutation cause alterations both in the integrity of the filament and in the intercellular exchange of fluorescent markers (Flores *et al.*, 2007; Mullineaux *et al.*, 2008; see also Results & Discussion Chapter 2). Another feature that is altered in the mutants is the number of nanopores through which septal junctions appear to traverse the septal peptidoglycan (Results & Discussion Chapter 1). SepJ and HglK are therefore needed in *Anabaena* to produce mature septa, and they could interact with other proteins forming protein complexes involved in septal junction formation or function. In order to search for these possible partners, SepJ and HglK were isolated under non-denaturing conditions with the help of one (HglK) or two (SepJ) of two different tags, the Strep-tag II and the GFP. The different protein isolates were subjected to mass spectrometry with the aim of detecting the proteins that could have co-purified with SepJ and HglK. The mass spectrometry results provided us with a large number of potential partners that were submitted to statistical analysis, using the results of the negative control (wild-type extracts) to correct for unspecific precipitation. To enhance the reliability of the detection of a protein, we concentrated on proteins that were repeatedly detected in different samples (see Fig. 42C). Some of these proteins were probably not SepJ or HglK physiological partners, since they likely are proteins very abundant in the cell that could bind non-specifically to the protein used as bait in the purification processes. Such proteins that might have bound to SepJ or HglK non-specifically include phycobilisome or ATPase subunits such as ApcD, CpcD, CpcG, AtpA, AtpD and AtpG. On the other hand, some other proteins show characteristics that make us to consider them as candidate partners as discussed below.

HglK was detected in the three SepJ isolations, in two of them with  $p < 0.05$  (Tables 12 and 13; Fig. 42A, B), strongly suggesting that both proteins interact *in vivo*. In contrast, SepJ was not detected in the HglK isolations. However, we noted that the precipitation of HglK was less effective when compared to the precipitation of SepJ. This might reflect a lower abundance of



HglK than of SepJ in *Anabaena* cells, which can be deduced from the difficulty of detection of HglK-GFP fusions as compared to SepJ-GFP fusions (Chapter 2, compare to Flores *et al.*, 2007). Also, the *sepJ* transcript is detected more easily than the *hglK* transcript (Black *et al.*, 1995; Flores *et al.*, 2007). Additionally, the selection of the tag might also interfere with the mode of interaction. For example, HglK was precipitated more efficiently with SepJ-GFP than with SepJ-Strep II (Fig. 42A vs. Fig 42B). Thus, it is possible that the Strep-tag II of HglK might have had an influence on the affinity of the HglK-SepJ complex resulting in loss of SepJ during the wash steps of the column material.

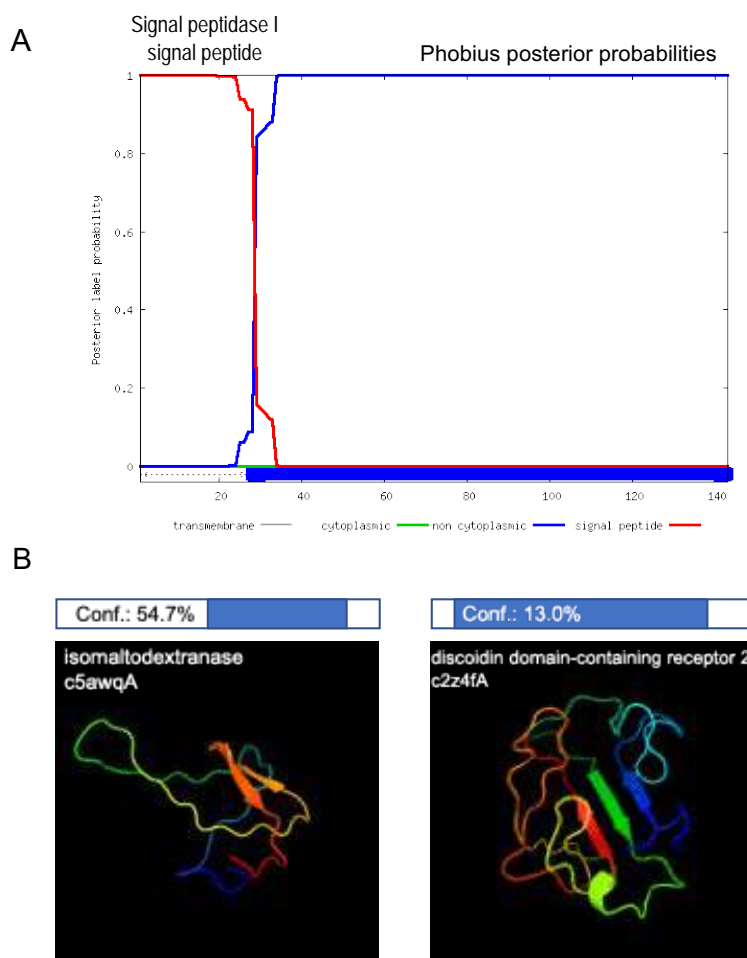
The observed interaction of SepJ and HglK deduced from the presented proteomic analysis is consistent with the phenotypic analysis of *sepJ* and *hglK* mutants. The two mutants show a number of similar characteristics that have led to the suggestion that SepJ and HglK contribute to the maturation of intercellular septal structures in *Anabaena* (Flores *et al.*, 2018 and Results & Discussion Chapter 2). The nature and structural basis of the SepJ-HglK interaction remain to be investigated. We note, nonetheless, that both proteins contain integral-membrane and putative periplasmic domains that could mediate the interaction. On the other hand, SepJ is localized very focused in the center of the septa (Flores *et al.*, 2007), whereas HglK shows a preferential localization in the intercellular septa with possible double spots that are laterally, rather than centrally, localized in the septa (Chapter 2). Perhaps the interaction of the two proteins is limited to the places that delimit their localization.

In addition to HglK, Alr2947 was identified with statistical significance in two SepJ isolations (Tables 12 and 14) and one of the HglK isolations (Table 16). This suggests that Alr2947 could be a common partner for SepJ and HglK, and even a three-part complex involving the three proteins might be considered for future research. Alr2947 is predicted to bear a signal peptidase I-type of signal peptide (Fig. 43A), implying that it is a soluble periplasmic protein. It is a very basic protein (pI, 9.74), and it seems to be rather unstructured (Fig. 43B). Interestingly, it is currently thought that proteins that are unstructured in solution become structured in complexes (Mohan *et al.*, 2006). It is tempting to hypothesize that Alr2947 interacts with periplasmic domains of both SepJ and HglK.

Alr0114 was repeatedly identified in our analyses, with very good statistical significance in one of the SepJ isolations (Table 14) and limited significance in one of the HglK isolations (Table 16). This is the well-known protein Tic22, a periplasmic holdase (chaperons that assist the non-covalent, ATP-independent folding of proteins) that is important in outer membrane biogenesis in *Anabaena* (Brouwer *et al.*, 2019; Tripp *et al.*, 2012). Considering that secretion system protein SecA was also found to be significantly enriched in one of the SepJ-based isolates (Fig. 43A, right panel; Table 13), it is tempting to hypothesize that these proteins interact with the translocating protein in the cytoplasm (SecA) or periplasm (Tic22). Here, Tic22 might be involved in the folding of periplasmic domains of both SepJ and HglK.





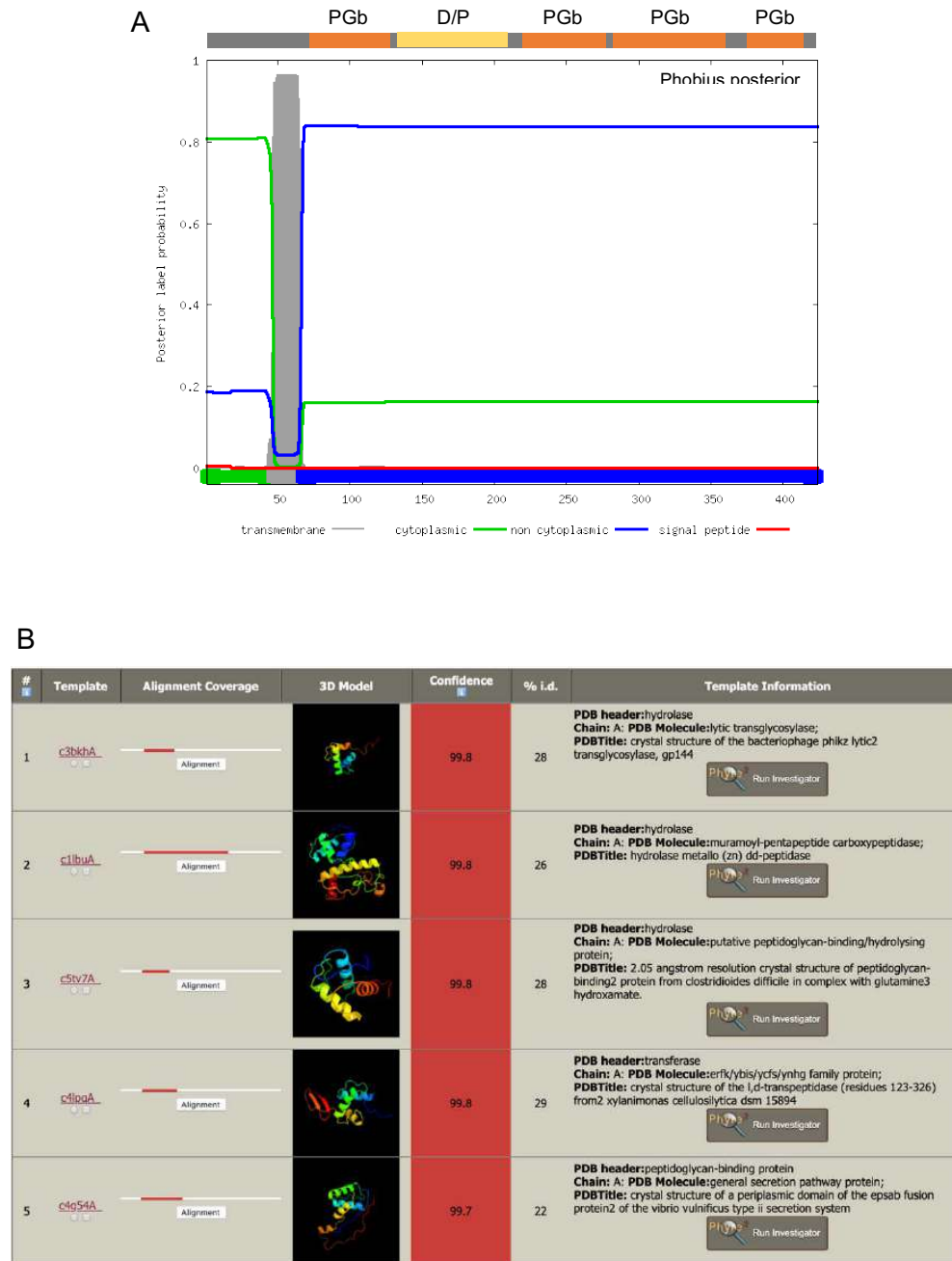


**Fig. 43. Properties of putative SepJ- and HglK-interacting protein Alr2947.** (A) Phobius prediction of a signal peptide and periplasmic location of Alr2947. Signal peptidase I-signal peptide predicted by SignalP-5.0. Note that signal peptidase I releases the protein to the periplasm. Alr2947 is a 143-residue protein, and the processed protein is predicted to have 116 residues. (B) Phyre2 analysis revealed no highly predicted 3D model, showing instead a largely unstructured protein. For the two 3D model examples depicted, the top bar shows in blue the Alr2947 fragments that align with the target sequences. Note the low confidence of the predictions.

Another protein that was found to interact with SepJ is Alr1690 (Table 14). This is a cytoplasmic membrane-anchored protein that is largely periplasmic (Fig. 44A). It bears four predicted peptidoglycan-binding domains and, consistently (Fig. 44A), it is predicted to form structures that are found in PG-binding and PG-metabolism proteins (Fig. 44B). Because the possible periplasmic domain of SepJ also shows PG-binding activity (Ramos-León *et al.*, 2017), Alr1690 and SepJ might interact as PG-binding proteins. Further, it has been previously shown that SepJ interacts with Al11861, which is a PG-binding protein called SjcF1 that is involved in the regulation of nanopore size in *Anabaena* (Rudolf *et al.*, 2015).







**Fig. 44. Properties of putative SepJ-interacting protein Alr1690.** (A) Phobius prediction of a transmembrane segment and periplasmic location of Alr1690. The upper scheme shows the approximate location of peptidoglycan-binding motifs (PGb, orange) and a disordered/polar region (D/P, yellow; taken from information in Uniprot). (B) Phyre2 analysis of Alr1690 predicts structural motifs found in PG-binding proteins and PG-metabolism proteins. The left part shows the Alr1690 fragments that align with the target sequences (red lines) and their predicted structures. The right part indicates the proteins to which those fragments show highest structural similarity.



Finally, we note Alr3277, detected with limited statistical significance in isolations of SepJ (Table 13) and HglK (Table 15). Alr3277 is a 179-amino acid residue protein of predicted cytoplasmic localization that shows a domain similar in structure to the late competence protein ComFB from *Bacillus subtilis* (Fig. 45). We do not know the significance that this finding may have.



**Fig. 45. Properties of putative SepJ- and HglK-interacting protein Alr3277.** Phyre2 analysis predicts a domain with a structure similar to that of ComFB from *Bacillus subtilis*. The left part shows the Alr3277 fragment that aligns with the target sequence (red line) and its predicted structure, and the right part indicates the protein to which that fragment shows highest structural similarity. Alr3277 is composed of 179 amino acid residues and ComFB is a small, 98-residue protein.

In summary, this proteomic analysis has provided us with independent evidence for a possible SepJ-HglK interaction. Additionally, it has identified three periplasmic proteins that may interact with SepJ (Alr1690) or with both SepJ and HglK (Alr2947 and Alr0114/Tic22). Although these possible interactions need further study, the predominance of periplasmic proteins in this analysis is consistent with the important role that SepJ and HglK have for the maturation of intercellular septal structures in *Anabaena*.



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

### 3. Results & Discussion

---

#### Chapter 4

## The SepJ and Fra proteins from marine symbiotic *Richelia* and *Calothrix*

ÁMBITO- PREFIJO

GEISER

Nº registro

00008745e2000022470

CSV

GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

DIRECCIÓN DE VALIDACIÓN

<https://sede.administracionespublicas.gob.es/valida>

FECHA Y HORA DEL DOCUMENTO

12/06/2020 08:18:06 Horario peninsular



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

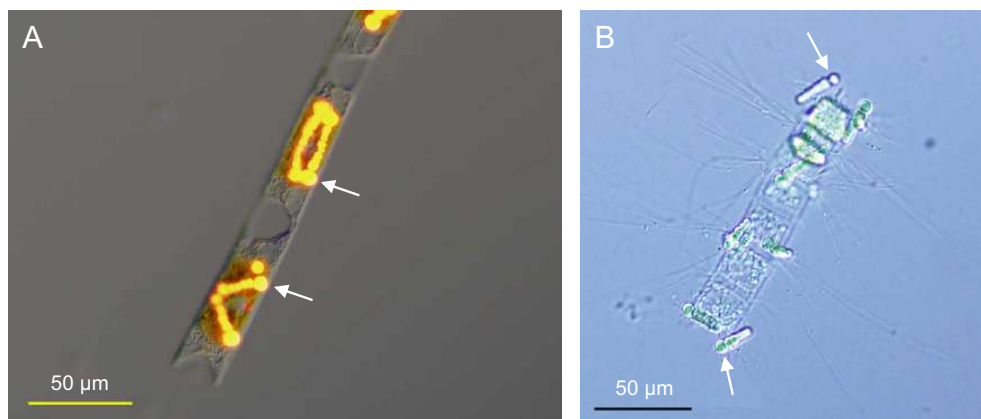
FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



### 3.4 The SepJ and Fra proteins from marine symbiotic *Richelia* and *Calothrix*

The SepJ and Fra proteins have an important role for the multicellular structure and function in *Anabaena* (Herrero *et al.*, 2016). Here we addressed a study of these proteins from two heterocystous cyanobacteria that are marine diatom symbionts, *Richelia intracellularis* and *Calothrix rhizosoleniae* (Hilton *et al.*, 2013). *R. intracellularis* strain HH01 (hereafter RintHH01 or, simply, HH01) is an obligate endosymbiont that lives within the cytoplasm of *Hemiaulus hauckii* (Fig. 46A) (Villareal, 1990; Caputo *et al.*, 2019), and *C. rhizosoleniae* strain SC01 (hereafter CalSC01 or, simply, SC01) is a facultative exosymbiont of *Chaetoceros compressus* (Fig. 46B), i.e., SC01 can live extracellularly attached to the diatom or as a free-living organism (Foster *et al.*, 2010). Genomic sequences of these strains with apparently good coverage have been published and are available (Hilton *et al.*, 2013; <https://img.jgi.doe.gov/>). Two other *Richelia intracellularis* strains for which some genomic sequences are available are HM01, a cytoplasmic endosymbiont of *Hemiaulus membranaceus* with an incomplete sequence available, and RC01, which lives between the cytoplasmic membrane and the frustule (siliceous cell wall) of *Rhizosolenia clevei* and for which a good (though likely incomplete) sequence coverage is available (Hilton *et al.*, 2013; Hilton, 2014). Our approach to study the SepJ and Fra proteins from the symbionts was to express them in *Anabaena* mutants of the corresponding genes to test their functionality.



**Fig. 46.** Micrographs of *Hemiaulus hauckii* containing *Richelia intracellularis* (A) and *Chaetoceros compressus* with *Calothrix rhizosoleniae* attached (B). Arrows point to some heterocysts. Micrographs provided by Rachel A. Foster (Stockholm University, Sweden), courtesy of Tracy A. Villareal (The University of Texas at Austin, USA). Micrograph B published in Hilton *et al.* (2013).

#### 3.4.1 SepJ and Fra proteins encoded in the RintHH01 and CalSC01 genomes.

The genomes of both RintHH01 and CalSC01 contain a *sepJ* gene (2022 bp in RintHH01; 2085 bp in CalSC01; for comparison, *Anabaena sepJ* is 2256 bp) (searched at <https://img.jgi.doe.gov/>). The SepJ proteins encoded in RintHH01 and CalSC01 are 39.4% and 44.5% identical, respectively, to *Anabaena* SepJ, whereas the proteins from RintHH01 and CalSC01 are 61.1% identical to each other. An alignment of these proteins and the SepJ protein sequence also available from *R. intracellularis* HM01 is shown in Fig. 47. The symbiont SepJ proteins can be aligned to the *Anabaena* protein throughout the four domains that have been defined for SepJ: a highly-conserved N-terminal sequence; the coiled-coil domain that contains two strongly



predicted coiled-coil motifs; the linker domain that is rich in Pro, Ser and Thr; and the integral membrane domain that is homologous to proteins in the DME family (Herrero *et al.*, 2016). The least conserved of these domains is the linker domain, which in SepJ from the symbionts is shorter than in SepJ from *Anabaena* (yellow underline in Fig. 47). Nonetheless, the linker domains of CalSC01 and RintHH01 SepJ are enriched in Pro, Ser and Thr, containing, respectively: 19.5% and 19.8% Pro; 14.5% and 9.9% Ser; 18.2% and 8.3% Thr. This conforms to what is expected for SepJ proteins, since the linker domain is highly variable in length and sequence but conserved in amino acid composition (Herrero *et al.*, 2016).

The genome of CalSC01 contains a *fraC-fraD-fraE* gene cluster (Fig. 48A) that is similar to the *Anabaena fraC* operon (Merino-Puerto *et al.*, 2010). In contrast, RintHH01 contains only a *fraC-fraD* cluster (Fig. 48B). This is unusual, since heterocystous cyanobacteria generally contain a complete *fraCDE* cluster (see Table 9 in Results & Discussion Chapter 2). Nevertheless, lack of a *fraE* gene is interesting, since the phenotype of the *fraE* mutant of *Anabaena* is somewhat different from that of *fraC*, *fraD* or double *fraC-fraD* mutants, the three latter mutants showing identical phenotypes (Merino-Puerto *et al.*, 2010, 2011). Additionally, we have found that the subcellular localization of FraE has specific characteristics (see Results & Discussion Chapter 5). The encoded FraC sequences in RintHH01 and CalSC01 are 45.5% and 47.6% identical, respectively, to *Anabaena* FraC and 62.6% identical to each other. A global comparison of these proteins showing overall similarity is presented in Fig. 49A, in which the FraC sequence available from *R. intracellularis* RC01 is also included. The FraD sequences from RintHH01 and SC01 are 45.5% and 52.4% identical to *Anabaena* FraD, respectively, and 61.9% identical to each other. A global comparison of these proteins and that from RC01 is presented in Fig. 49B, in which the insertion or deletion of some short fragments specifically in the periplasmic domain of FraD can be observed. Finally, FraE from CalSC01 is 60.6% identical to *Anabaena* FraE, and similarity is overall although the two proteins may be produced from a different start codon (Fig. 49C).

Based on the presence of these proteins, with their peculiarities, in the symbionts, we asked whether they could be expressed in *Anabaena* and, if so, whether they could carry out their functions in the multicellular phenotype of the cyanobacterium.

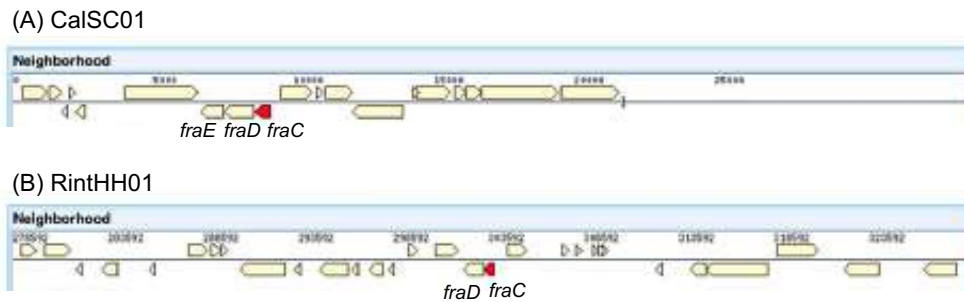






**Fig. 47. Comparison of the SepJ amino acid sequences from *R. intracellulare* (strains HH01 and HM01), *C. rhizosoleniae* (strain SC01) and *Anabaena*.** Clustal W alignment. Blue underline denotes N-terminal sequence; green, coiled-coil domain; yellow, Pro-rich linker domain; red, integral membrane domain.





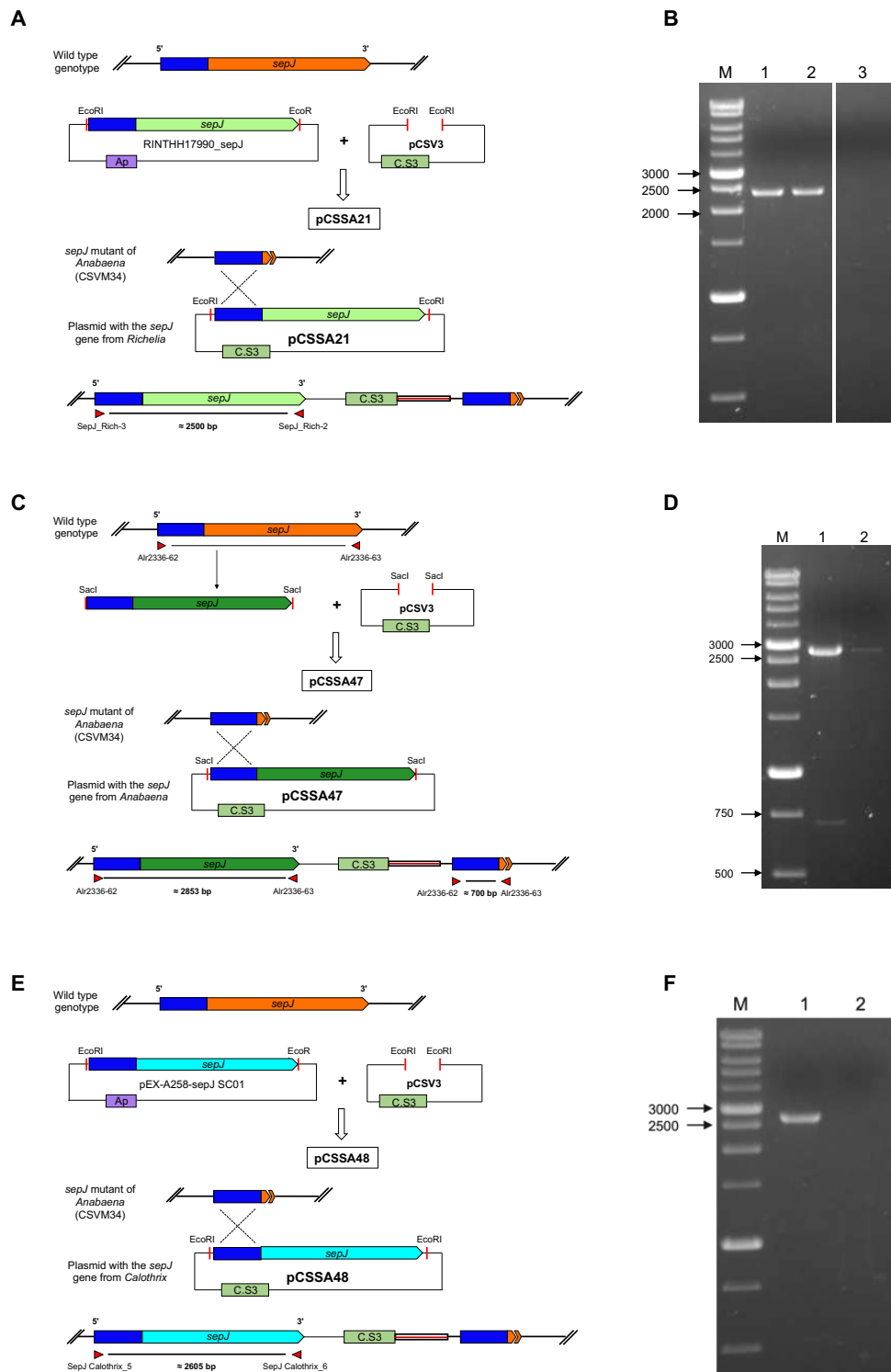
**Fig. 48. Presence of a *fraCDE* gene cluster in *C. rhizosoleniae* SC01 (A) and of a *fraCD* gene cluster in *R. intracellulares* HH01 (B).** Gene clusters identified in Integrated Microbial Genomes (<https://img.jgi.doe.gov/cgi-bin/m/main.cgi>) searching for homologues of *Anabaena fraC* (red color).

### 3.4.2 Expression in *Anabaena* of symbiont *sepJ* genes

To clone in *Anabaena* the *sepJ* genes from the two symbionts, RintHH01 and CalSC01, we chose to try complementation of *Anabaena* strain CSVM34 that has a deletion of most of the *sepJ* gene (deletion corresponds to residues 40 to 747 of the *Anabaena* 751-amino acid SepJ sequence; Mariscal *et al.*, 2011). We prepared constructs consisting of 500 bp of the *Anabaena sepJ* upstream sequence followed by the symbiont custom-synthesized gene, all cloned in plasmid vector pCSAV3 that encodes Sm<sup>R</sup>/Sp<sup>R</sup> and can be transferred from *E. coli* to *Anabaena* by conjugation. This type of construct allows the integration of the whole plasmid into the *Anabaena sepJ* locus by a single recombination event, resulting in a genomic structure that should allow expression of the heterologous gene from the native *sepJ* promoter. Figure 50 shows the construction procedure, which was used for custom-synthesized *sepJ* from RintHH01, custom-synthesized *sepJ* from CalSC01 with codon usage adapted to *Anabaena*, and the PCR-amplified *Anabaena* gene as a positive control. For the construction of the CSSA14 and CSSA28 strains (bearing *sepJ* from RintHH01 and CalSC01, respectively), it was sufficient to digest the custom-synthesized plasmids with the same enzyme as that in plasmid pCSV3 (EcoRI). For the construction of strain CSSA29 (strain carrying *sepJ* from *Anabaena* itself), the primers used to carry out the amplification of the gene (alr2336-62/alr2336-63) included the sequence for SacI, which made it possible its ligation with SacI-digested plasmid pCSV3 (see Fig. 50). Before performing conjugation with the parental strain, CSVM34, all plasmid inserts were corroborated by sequencing. After conjugation, PCR analysis of the exconjugants showed that the constructs were incorporated at the *sepJ* locus as expected (Fig. 50).





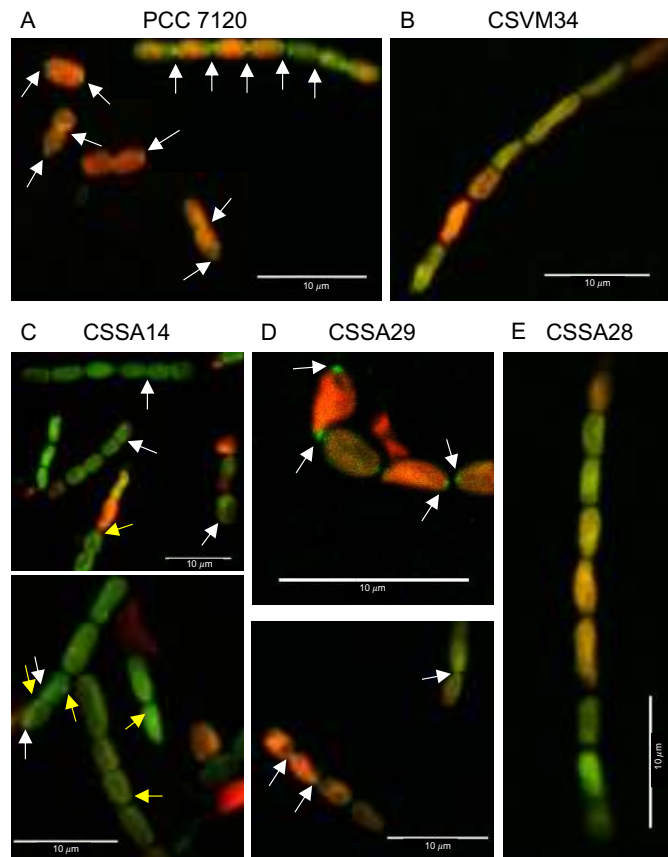


**Fig. 50. Constructs for expression of symbiont *sepJ* genes in *Anabaena*.** The figure shows how the CSSA14, CSSA28 and CSSA29 strains were obtained by single recombination to express SepJ of different cyanobacteria under the natural *Anabaena sepJ* promoter, together with their corresponding genotypic tests by PCR. (A) Construction of *Anabaena* strain CSSA14 in which *sepJ* from RintHH01 is cloned at the natural *sepJ* locus in *Anabaena* strain CSVM34 ( $\Delta sepJ$ ). (B) PCR verification of two independent clones of strain CSSA14 (lanes 1 and 2) where it can be seen the amplification of *Richelia sepJ*, while that amplification did not occur with *Anabaena* wild-type DNA (lane 3). (C) Construction of *Anabaena* strain CSSA29 (positive control) that expresses *sepJ* from *Anabaena* cloned at the *Anabaena*'s natural *sepJ* site in strain CSVM34 ( $\Delta sepJ$ ). (D) PCR test of strain CSSA29 (lane 1) where a band of 2853 bp is seen corresponding to the amplification of the *sepJ* gene plus about 500 bp upstream from the start of the gene. This band of approximately 2853 bp can also be seen in wild-type *Anabaena* (lane 2). In strain CSSA29 (lane 1) there is also a band of approximately 700 bp corresponding to the size of the region including deleted *sepJ*, indicating that this strain is not completely segregated. (E) Construction of *Anabaena* strain CSSA28 in which *sepJ* from CalSC01 (modified to have an *Anabaena*-adapted codon usage) is expressed under the natural *sepJ* promoter of *Anabaena* in strain CSVM34 ( $\Delta sepJ$ ). (F) PCR verification of strain CSSA29 (lane 1) where it can be seen the amplification of *sepJ* from SC01, while that amplification did not occur with *Anabaena* wild-type DNA (lane 2). M, base pair markers.

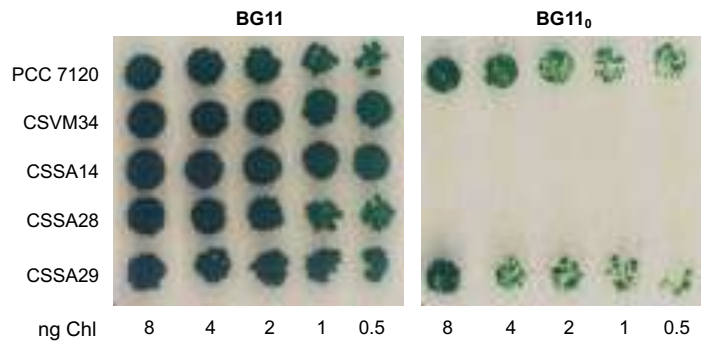
To study whether the heterologous genes were expressed and resulted into a properly located protein, we carried out immunofluorescence analysis with antibodies against the coiled-coil domain of *Anabaena* SepJ (Mariscal *et al.*, 2011; details in Materials & Methods section 2.7.2). The antibodies detected the SepJ protein at the intercellular septa in *Anabaena* and in the self-complemented strain CSSA29, but did not detect any material in the negative control CSVM34 ( $\Delta sepJ$ ), as expected (Fig. 51). In strain CSSA14 (*Anabaena::sepJ<sub>RintHH01</sub>*), some spots were observed that however were not localized at the intercellular septa, and in strain CSSA28 (*Anabaena::sepJ<sub>CalSC01</sub>*) no spots were observed. These negative results provide no evidence for the proper expression and location of the heterologous proteins in *Anabaena*, although some expression is possible at least for HH01 SepJ.

Nonetheless, we carried out growth tests and determined filament length to enquire on any possible function of the heterologous genes in *Anabaena* if, for example, they were expressed at low levels. Growth tests showed that, as is was the case for the parental strain CSVM34, neither strain CSSA14 (*Anabaena::sepJ<sub>RintHH01</sub>*) nor CSSA28 (*Anabaena::sepJ<sub>CalSC01</sub>*) could grow diazotrophically (BG11<sub>0</sub> medium; Fig. 52). In contrast, the self-complemented strain CSSA29 grew diazotrophically as the wild type (PCC 7120), and all strains could grow with combined nitrogen (BG11 medium; Fig. 52). Additionally, we tested filament length in BG11 medium to test whether the symbiont genes could provide filament function in spite of lack of growth without combined nitrogen.





**Fig. 51. Immunolocalization of SepJ.** Localization of SepJ using antibodies against the *Anabaena* SepJ coiled-coil domain in the different strains supplemented with *sepJ* from different cyanobacteria. The white arrows indicate the septal zones where the antibody bound, while yellow arrows indicate the lateral zones. (A) *Anabaena* wild type; (B) strain CSVM34 ( $\Delta sepJ$ ) used as negative control; (C) CSSA14 ( $\Delta sepJ::sepJ_{RimHH01}$ ); (D) CSSA29 ( $\Delta sepJ::sepJ_{PCC7120}$ ); (E) CSSA28 ( $\Delta sepJ::sepJ_{CalSCO1}$ ).

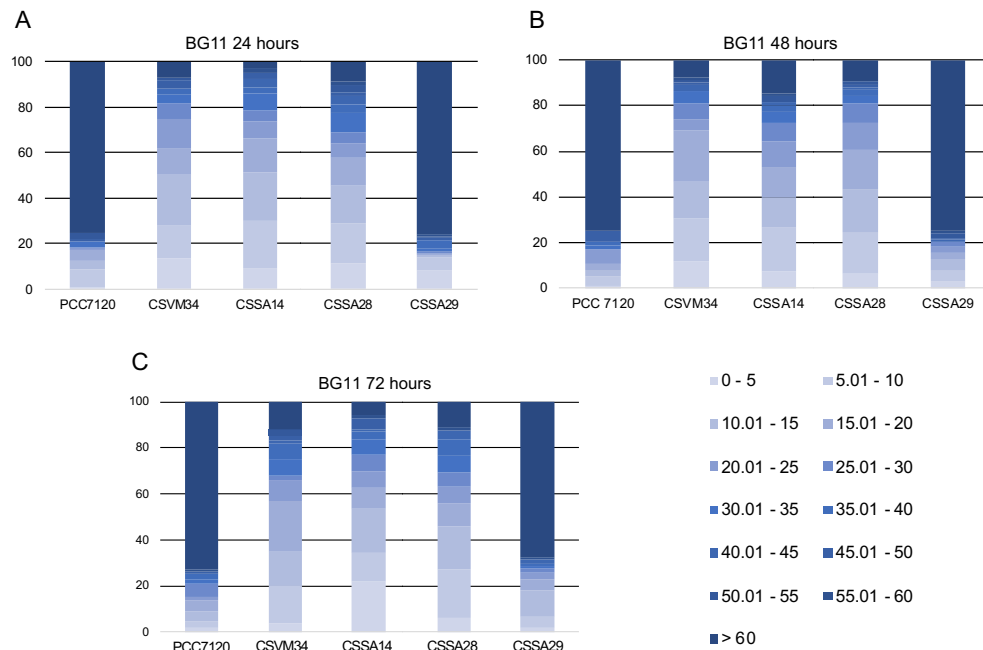


**Fig. 52. Growth test on solid medium.** The figure shows the growth of *Anabaena* wild type (PCC 7120), CSVM34 ( $\Delta sepJ$ ), CSSA14 ( $\Delta sepJ::sepJ_{RimHH01}$ ), CSSA29 ( $\Delta sepJ::sepJ_{PCC7120}$ ) and CSSA28 ( $\Delta sepJ::sepJ_{CalSCO1}$ ) in solid medium with a combined nitrogen source (BG11) or without combined nitrogen (BG11<sub>0</sub>). The spots were inoculated with the amounts of Chl indicated below (ng) and the plates were incubated under standard growth conditions for 7-10 days.





The parental strain CSVM34 ( $\Delta sepJ$ ) produced a very low percentage of long filaments, as previously described (Mariscal *et al.*, 2011), and similar filament lengths were observed with strains CSSA14 and CSSA28 that carried the symbiont genes, whereas only the self-complemented strain CSSA29 showed a distribution of filament lengths similar to the wild type (Fig. 53). In conclusion, we did not find any indication of successful expression and function of the *sepJ* genes from the symbionts in *Anabaena*.



**Fig. 53. Filament length in *Anabaena* strains bearing *sepJ* from the symbionts.** The figure shows the distribution of filaments with the indicated numbers of cells for strains *Anabaena* wild type (PCC 7120), CSVM34 ( $\Delta sepJ$ ), CSSA14 ( $\Delta sepJ::sepJ_{RintHH01}$ ), CSSA28 ( $\Delta sepJ::sepJ_{CalSCO1}$ ) and CSSA29 ( $\Delta sepJ::sepJ_{PCC7120}$ ). Filaments were inoculated at 1  $\mu$ g Chl/mL in liquid medium with a source of combined nitrogen ( $\text{NaNO}_3$ ) and incubated for 24 h (A), 48 h (B) and 72 h (C) under standard growth conditions. Filaments were taken with great care to determine filament length in micrographs taken under the light microscope.

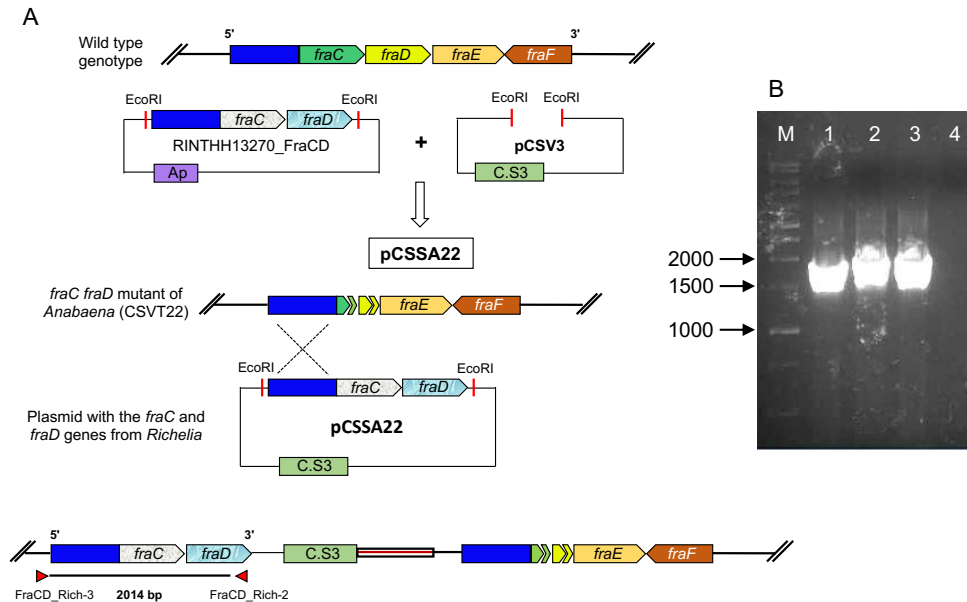
### 3.4.3 Expression in *Anabaena* of symbiont *fra* genes

To clone the *fra* genes from the symbionts in *Anabaena*, we chose to try complementation of *Anabaena* strain CSVT22 that has a deletion of the *fraC* and *fraD* genes (Merino-Puerto *et al.*, 2011b). We first tried complementation with *fraC-fraD* from RintHH01 and, for that, we followed a strategy similar to one described above for *sepJ*, preparing a construct that contained 500 bp of the *Anabaena fraC* upstream sequence followed by the symbiont custom-synthesized *fraC-fraD* genes from RintHH01, all cloned in plasmid vector pCSV3. As in the previous section, for the cloning of the RintHH01 *fraC-fraD* genes from a custom-synthesized plasmid into plasmid pCSV3, both plasmids were digested with EcoRI, and the insert of the plasmid generated was corroborated by sequencing. This construct allows the integration of the whole plasmid just upstream of the *Anabaena fraC* locus by a single recombination event, resulting in a genomic





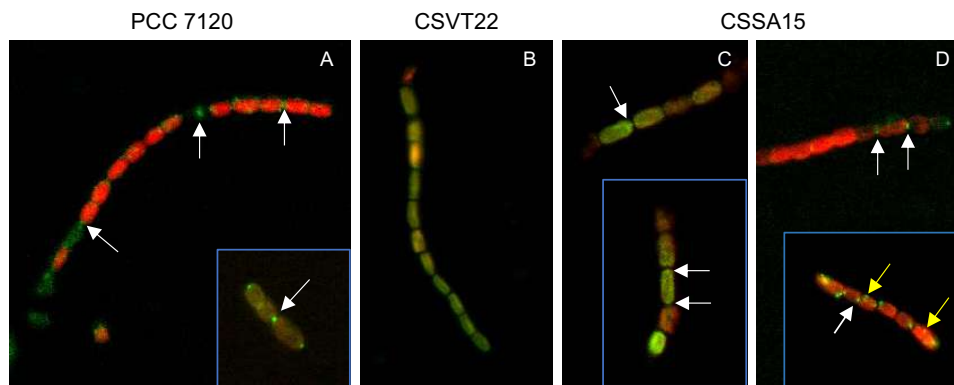
structure that should allow expression of the heterologous *fraC-fraD* genes from the native *fraC* promoter, as well as expression of the last gene in the operon, *fraE*, from the *fraC* promoter (Fig. 54A). Following this procedure, *Anabaena* strain CSSA15 that carries the RintHH01 *fraC-fraD* genes as shown by PCR analysis was obtained (Fig. 54B).



**Fig. 54. Constructs for expression of the *fraCD* genes from RintHH01 in *Anabaena*.** (A) Steps carried out for the insertion of the *fraCD* genes from strain HH01 by single recombination with the genome of *Anabaena* strain CSVT22 ( $\Delta$ *fraCD*). The strain generated, CSSA15, should allow the expression of the *Richelia fraCD* genes using the natural *fraCDE* operon promoter of *Anabaena*. (B) Verification by PCR of some exconjugants. Lanes 1, 2 and 3: DNA from three different clones of CSSA15 which have the insertion of the *Richelia fraCD* genes as shown by the band of 2014 bp, while lane 4 corresponds to the wild-type DNA with which no amplification was obtained. M, base pair markers.

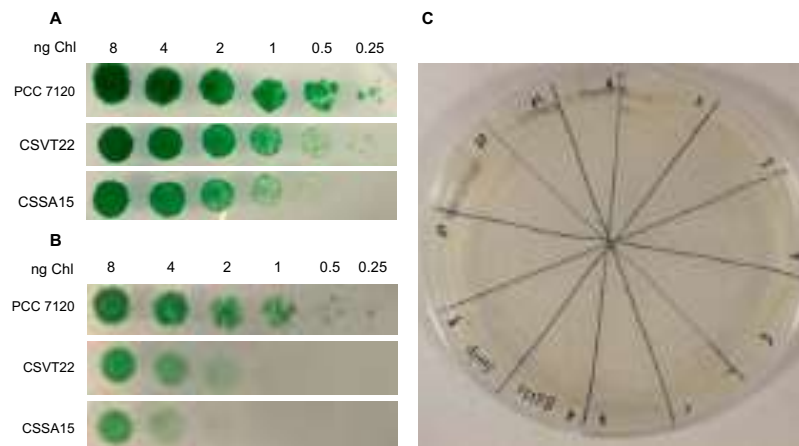
To study whether the heterologous genes were expressed and could result into properly located proteins, we carried out immunofluorescence analysis specifically with antibodies against the extramembrane domain of *Anabaena* FraD (Merino-Puerto *et al.*, 2011b), since no antibodies for FraC are available. The antibodies detected the FraD protein at the intercellular septa in *Anabaena* and not in the negative control, the double  $\Delta$ *fraC*  $\Delta$ *fraD* mutant CSVT22 (Fig. 55A and B). Localization was also obtained, partially at the intercellular septa, in two clones of CSSA15 (*Anabaena::fraCD*<sub>RintHH01</sub>) that were analyzed, indicating that at least the FraD gene from RintHH01 was produced in *Anabaena* (Fig. 55C and D).





**Fig. 55. Immunolocalization of FraD.** The figure shows immunolocalization analysis carried out with antibodies against FraD. White arrows indicate those polar places where the antibody was attached while yellow arrows show lateral places. (A) Immunolocalization carried out in (A) *Anabaena* wild type (PCC 7120), (B) strain CSVT22 ( $\Delta fraC \Delta fraD$ ), (C and D) two clones of strain CSSA15, in which the *fraC fraD* genes from RintHH01 were inserted into strain CSVT22.

As is parental strain CSVT22 (Merino-Puerto *et al.*, 2011b), CSSA15 was however unable to grow diazotrophically (several independent clones tested, see Fig. 56).



**Fig. 56. Growth tests on solid medium with different nitrogen sources.** (A) Solid BG11<sub>0</sub> medium supplemented with ammonium ( $\text{NH}_4^+$ ) as a nitrogen source. (B) Solid BG11 medium (which has  $\text{NaNO}_3$  as a nitrogen source). In A and B, samples of the wild-type *Anabaena* (PCC 7120), and strains CSVT22 ( $\Delta fraC \Delta fraD$ ) and CSSA15 (CSVT22 with the RintHH01 *fraC-fraD* genes) were inoculated with drops. (C) Solid BG11<sub>0</sub> medium (no source of combined nitrogen) in which twelve CSSA15 clones were inoculated by streaking.

Additionally, nitrogenase activity under oxic conditions was tested in several clones of CSSA15, but undetectable or very low activity was found (Table 17). These results show that, in spite of detectable production (at least of FraD), FraC and FraD from RintHH01 did not functionally complement the  $\Delta fraC \Delta fraD$  mutant of *Anabaena*. However, we are not certain if the construct strategy was appropriate. For example, in spite of using the native *Anabaena* promoter, the expression level of RintHH01 *fraCD* genes could be poor and/or incompatible with *fraE* of *Anabaena*.

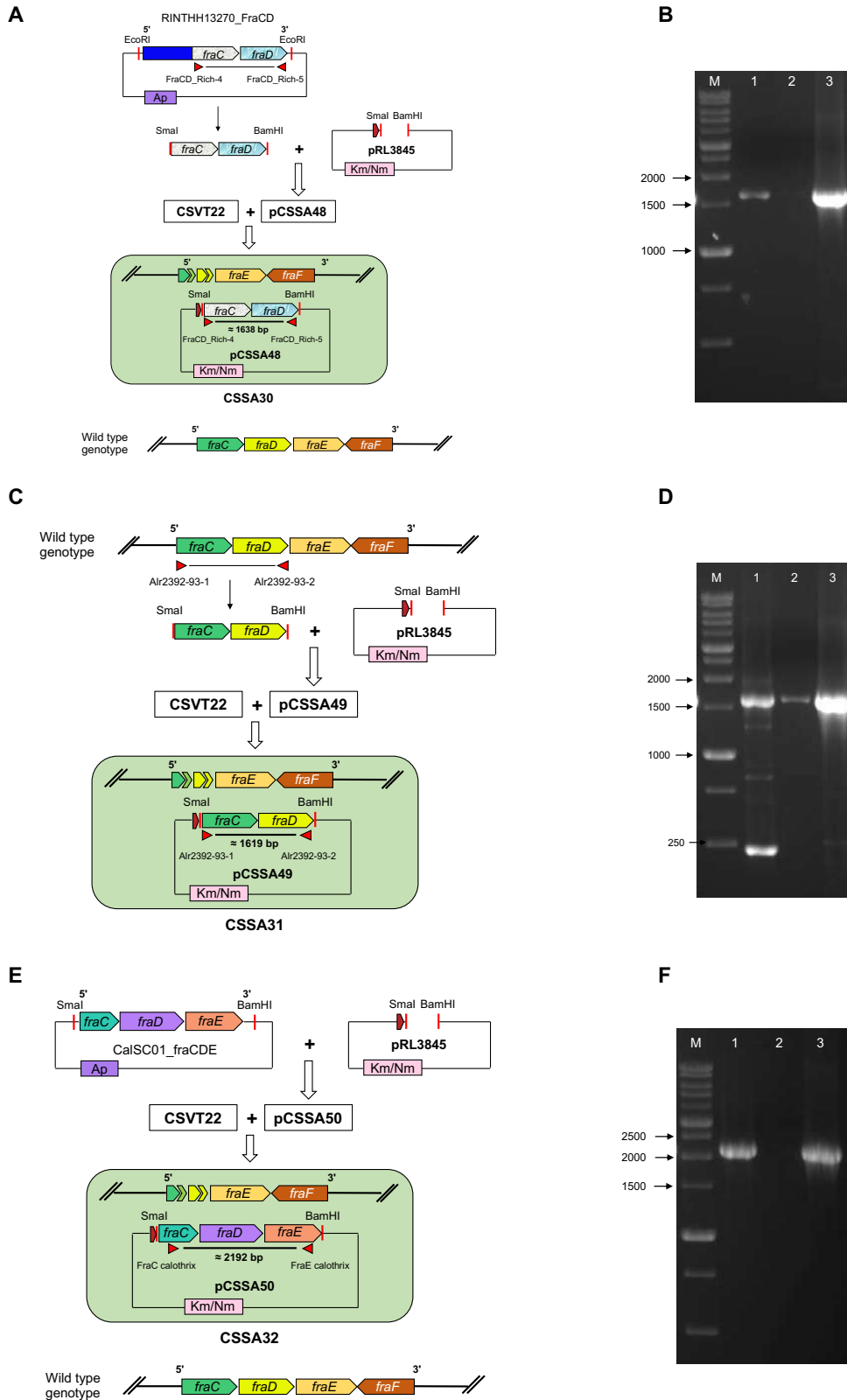


Strain	Nitrogenase activity (nmol ethylene [ $\mu\text{g Chl}$ ] $^{-1}$ h $^{-1}$ )
	Mean $\pm$ SD (n)
PCC 7120	8.628 $\pm$ 5.562 (3)
CSVT22	0.048 $\pm$ 0.046 (3)
CSSA15 #1	0.000 $\pm$ 0.008 (3)
CSSA15 #2	0.282 $\pm$ 0.302 (3)
CSSA15 #6	0.012 ; 0.000 (2)

**Table 17. Nitrogenase activity in strain CSSA15 and control strains.** Nitrogenase activity was determined under oxic conditions as described in Materials & Methods (section 2.8.3). Mean nitrogenase activity, standard deviation and number of replicates (n) for each strain are shown. The strains studied were: wild-type *Anabaena* (PCC 7120), CSVT22 ( $\Delta\text{fraC}$   $\Delta\text{fraD}$ ) and three clones of CSSA15 (CSVT22::*fraC-fraD*<sub>RintHH01</sub>).

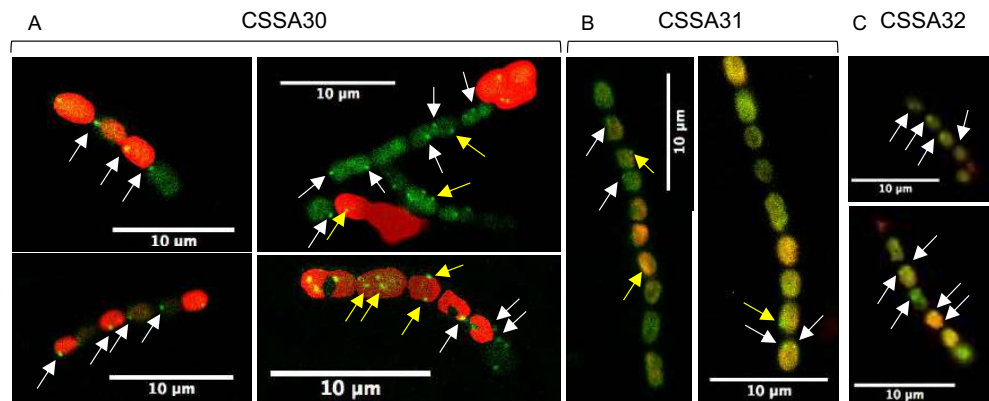
Based on these results, we attempted another strategy that consisted in expression of the *fraC-fraD* genes from RintHH01 or the whole *fraC-fraD-fraE* cluster from CalSC01 cloned downstream from a strong *glnA* promoter in a replicative plasmid (pRL3845, based on the pDU1 replicon of heterocystous cyanobacteria). For the construction of plasmid pCSSA50 composed of the custom-synthesized *fraC-fraD-fraE* cluster from CalSC01 (cloned in a plasmid) and plasmid pRL3845, both plasmids were digested with the same restriction enzymes (SmaI and BamHI) and ligated. On the other hand, for construction of plasmid pCSSA48, it was previously necessary to amplify the *fraCD* genes from RintHH01 (available in plasmids described above) using primers FraCD\_Rich-4 and FraCD\_Rich-5 that provided the restriction sequences necessary to perform the ligation between them and the plasmid pRL3845. This same process was performed with the *fraCD* genes from *Anabaena*, for which primers alr2392-93-1 and alr2392-93-2 were used. Before performing the conjugation, the inserts of all the synthesized plasmids were corroborated by sequencing. The construction of derivatives of *Anabaena* mutant CSVT22 ( $\Delta\text{fraC}$   $\Delta\text{fraD}$ ) carrying the custom-synthesized symbiont genes, as well as the strain carrying the PCR-amplified *Anabaena fraC* and *fraD* genes as positive control, is described in Figure 57. PCR analysis of the exconjugants showed that the following *Anabaena* strains bore the appropriate constructs: CSSA30, RintHH01 *fraC-fraD*; CSSA31, *Anabaena fraC-fraD* (self-complemented strain); CSSA32, CalSC01 *fraC-fraD-fraE* (Fig. 57).





**Fig. 57. Constructs for expression of *fraCD* genes from different cyanobacteria using the *glnA* promoter in a replicative plasmid in *Anabaena*.** Figures A, C and E show the construction of strains CSSA30, CSSA31 and CSSA32 that express the *fraCD* genes (*fraCDE* in strain CSSA32) from different cyanobacteria. These genes are expressed from the *Anabaena glnA* promoter, represented in the figures with a small red box-arrow, in a cyanobacterial replicative plasmid (pRL4835) in *Anabaena* strain CSVT22 ( $\Delta$ *fraCD*). (A) Construction of strain CSSA30 using *fraCD* from RintHH01. Plasmid pCSSA48 was composed by *fraCD* genes from RintHH01 cloned in plasmid pRL4835. Once the plasmid was made, it was transferred to strain CSVT22. (B) Verification by PCR strain CSSA30 that has *Richelia fraCD* as shown by the band of 1638 bp in lanes 1 and 3, corresponding to DNA of strain CSSA30 and plasmid RINTHH13270\_FraCD respectively. Lane 2 corresponds to *Anabaena* wild-type DNA where no amplification is seen. (C) Construction of strain CSSA31 using *fraCD* of *Anabaena* itself to reconstruct the wild-type genotype. The *fraCD* genes from *Anabaena* were cloned in pRL4835 generating plasmid pCSSA49 which was transferred to strain CSVT22. (D) PCR verification with DNA from (1) strain CSSA31, (2) wild-type *Anabaena*, and (3) pCSSA49, in which a correct band of 1619 bp was detected. In lane 1, there is a band with a size lower than 250 bp that corresponds to the region of the *fraCD* genes present in the parental strain (CSVT22). (E) Construction of strain CSSA32 using *fraCDE* of CalSC01. The *fraCDE* genes were extracted from plasmid CalSC01\_fraCDE and introduced into pRL4835, producing plasmid pCSSA50 that was transferred to strain CSVT22. (F) Verification by PCR of strain CSSA32 that has the *fraCDE* genes of *Calothrix* as shown by the band of 2192 bp in lanes 1 and 3 corresponding to DNA of CSSA32 strain and plasmid CalSC01\_fraCDE respectively. Lane 2 corresponds to DNA of wild-type *Anabaena* where no amplification is seen, as expected. M, base pairs marker.

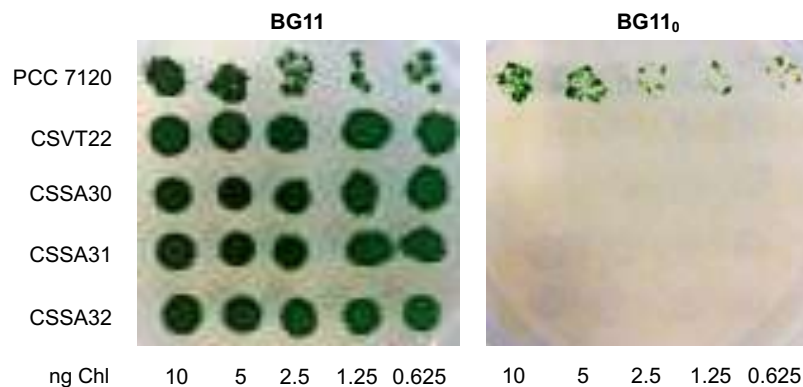
To study whether the heterologous genes were expressed and could result into properly located proteins, we carried out immunofluorescence analysis specifically with antibodies against the extramembrane domain of *Anabaena* FraD (Merino-Puerto *et al.*, 2011b). The positive (wild type) and negative (strain CSVT22) controls have been shown in Fig. 55 above. With the new constructs, the antibodies labeled material in the three test strains, CSSA30, CSSA31 and CSSA32, but the labeled material was only occasionally observed at the intercellular septa, even in the self-complemented strain CSSA31 (Fig. 58). Thus, whereas in the wild type 68% of the fluorescence spots were located in septal areas, in the complemented strains only 27% (CSSA30, CSSA31) and 17% (CSSA32) of the fluorescence spots were located in septal areas. These results show that the constructs are effectively expressed producing substantial amounts of protein, at least in the case of FraD, but perhaps the high expression expected from the *glnA* promoter produces an excessive amount of protein that affects expected localization.



**Fig. 58. Immunolocalization of FraD in strains carrying the *fra* gene constructs in a replicative plasmid.** Immunolocalization analysis carried out with antibodies against FraD. White arrows indicate those septal places where the antibody was attached while yellow arrows indicate lateral places. (A) CSSA30 ( $\Delta$ *fraC*- $\Delta$ *fraD* [pCSSA48], with RintHH01 *fraC fraD*); (B) CSSA31 ( $\Delta$ *fraC*- $\Delta$ *fraD* [pCSSA49], with *Anabaena fraC-fraD*); (C) CSSA32 ( $\Delta$ *fraC*- $\Delta$ *fraD* [pCSSA50], with SC01 *fraC-fraD-fraE*).



Growth tests were then performed, showing that the strains grew well in the presence of combined nitrogen (BG11 medium) but that neither CSV22 nor any of its derivatives could grow diazotrophically (BG11<sub>0</sub> medium) (Fig. 59). Nonetheless, we note that at the beginning of the incubation in BG11<sub>0</sub> medium some growth that failed with time was observed in the complemented strains compared to CSV22 (not shown).



**Fig. 59. Growth test in solid medium with strains carrying the *fra* gene constructs in a replicative plasmid.** The figure shows the growth of *Anabaena* wild type (PCC 7120) and strains CSV22 ( $\Delta fraC \Delta fraD$ ), CSSA30 ( $\Delta fraC \Delta fraD$  [pCSSA48], with RintHH01 *fraC fraD*), CSSA31 ( $\Delta fraC \Delta fraD$  [pCSSA49], with *Anabaena fraC-fraD*), and CSSA32 ( $\Delta fraC \Delta fraD$  [pCSSA50], with SC01 *fraC-fraD-fraE*), in solid medium with a combined nitrogen source (BG11) and in solid medium without combined nitrogen (BG11<sub>0</sub>). The spots were inoculated with the amounts of Chl indicated below and the plates were incubated under standard growth conditions for 10 days.

We then tested nitrogenase activity under both oxic and anoxic conditions and found that the complemented strains did not show improved nitrogenase activity but, rather, showed decreased activity as compared to strain CSV22 under anoxic conditions (Table 18). These results suggest that the *fra* genes cloned under the *glnA* promoter in the replicative plasmid were indeed expressed, producing a deleterious effect on the physiology of *Anabaena*, even in the case of the *Anabaena* genes themselves.

Strain	Nitrogenase activity (nmol ethylene [ $\mu\text{g Chl}$ ] <sup>-1</sup> h <sup>-1</sup> )	
	Mean $\pm$ SD (n)	
	Oxic condition	Anoxic condition
PCC 7120	0.750 $\pm$ 0.606 (3)	2.400 $\pm$ 1.282 (3)
CSV22	0.026 $\pm$ 0.043 (3)	1.448 $\pm$ 0.738 (3)
CSSA30	0.000 $\pm$ 0.000 (3)	0.004 $\pm$ 0.008 (3)
CSSA31	0.000 $\pm$ 0.000 (3)	0.094 $\pm$ 0.054 (3)
CSSA32	0.000 $\pm$ 0.006 (3)	0.018 $\pm$ 0.016 (3)

**Table 18. Nitrogenase activity in strains carrying the *fra* gene constructs in a replicative plasmid.** Nitrogenase activity was determined under oxic or anoxic conditions as described in Materials & Methods (section 2.8.3). Mean nitrogenase activity, standard deviation and number of replicates (n) for each strain are shown. The strains studied were: wild-type *Anabaena* (PCC 7120), CSV22 ( $\Delta fraC \Delta fraD$ ), CSSA30 ( $\Delta fraC \Delta fraD$  [pCSSA48], with RintHH01 *fraC fraD*), CSSA31 ( $\Delta fraC \Delta fraD$  [pCSSA49], with *Anabaena fraC-fraD*) and CSSA32 ( $\Delta fraC \Delta fraD$  [pCSSA50], with SC01 *fraC fraD fraE*).





### 3.4.4 Discussion

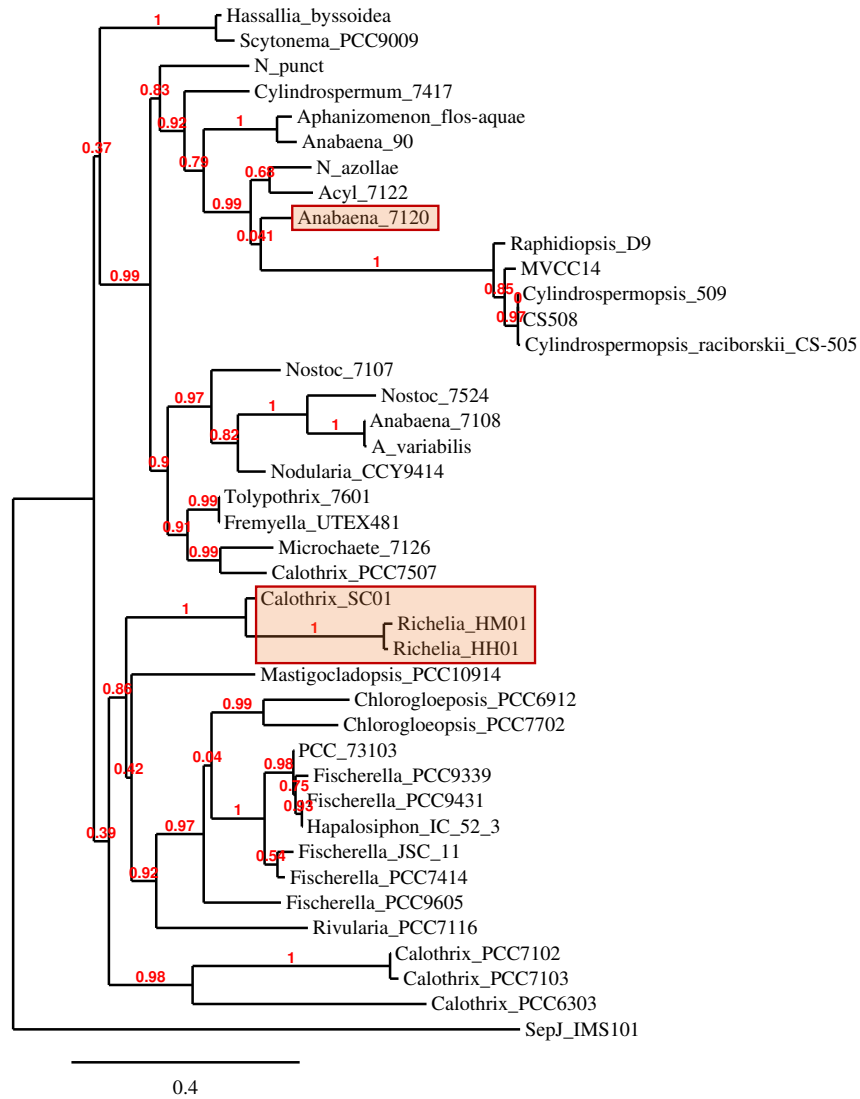
The diatom symbionts *R. intracellularis* and *C. rhizosoleniae* are heterocystous cyanobacteria and, as such, are expected to perform intercellular exchange of nutrients between vegetative cells and heterocysts. It is therefore expected that they express functional SepJ and Fra genes. Previous results showed that the SepJ protein from the filamentous, non-heterocystous cyanobacterium *Trichodesmium erythraeum* complemented at least the filamentation function of SepJ in *Anabaena* grown in the presence of combined nitrogen (Mariscal *et al.*, 2011). Previous results also showed that antibodies raised against the coiled-coil domain of *Anabaena* SepJ recognize septal material in the heterocystous, branching cyanobacterium *Mastigocladus laminosus* (Nürnberg *et al.*, 2014). Because, therefore, SepJ appears to be a very conserved protein, we addressed the study of SepJ from the symbionts by complementation of an *Anabaena*  $\Delta$ sepJ mutant. We failed, however, in obtaining any complementation of SepJ function by the *sepJ* genes from RintHH01 or CalSC01, even though the latter was adapted to the codon usage of *Anabaena*. A phylogenetic analysis of SepJ shows that the symbiont proteins cluster together, but they fall quite apart from *Anabaena* SepJ (Fig. 60). Nonetheless, *T. erythraeum* SepJ (indeed used as an outgroup in the phylogenetic tree) is further apart and, as mentioned above, provides *Anabaena* at least with filamentation function. As discussed earlier in this Thesis, SepJ may function inbuilt in protein complexes, and interactions with other proteins in the complex may depend on subtle structural features that may be affected by differences in amino acid sequence or composition. Such differences might affect the production or, if formed, the function of such complexes.

The FraC and FraD proteins have been recently shown to be part of the septal junctions that have been visualized by cryoET (Weiss *et al.*, 2019), and their functionality may therefore depend also in subtle interactions with septal junction partners. Phylogenetic analysis of FraC and FraD shows that they indeed fall apart from many other FraC/FraD proteins of heterocystous cyanobacteria including *Anabaena* (Figs. 61 and 62). FraE from CalSC01 is also distant from FraE of the most typical *Anabaena* strains (Fig. 63).

These observations may explain the failure of *fraC-fraD* from RintHH01 to complement the *Anabaena*  $\Delta$ fraC  $\Delta$ fraD mutant when expressed from the *Anabaena* *fraC* promoter (strain CSSA15 in Fig. 56 and Table 17). Expression of RintHH01 *fraC-fraD*, CalSC01 *fraC-fraD-fraE* and the own *Anabaena* *fraC-fraD* genes from the *glnA* promoter in a pDU1-based replicative plasmid appears to have, however, deleterious effects (Table 18), which we wonder whether could result from negative effects of the highly-expressed proteins on the formation of septal junctions. Our studies can be considered a first approach to the study of the *fra* genes from the diatom symbionts. Further attempts including accurate regulation of *fra* gene expression levels may help to succeed in the functional expression of the symbiont genes in *Anabaena*. Nonetheless, an interesting question that remains is whether there could be any substantial difference in the septal junctions of cyanobacteria containing only terminal heterocysts, like *Richelia* and *Calothrix*, as compared to cyanobacteria with intercalary heterocysts. In this context, expanding these studies to the septal junction proteins of conventional *Calothrix* spp. would be of interest. More generally, studies of the SepJ and Fra proteins from different heterocystous cyanobacteria are expected to increase our understanding of the formation of the septal junctions that are characteristic of these organisms.

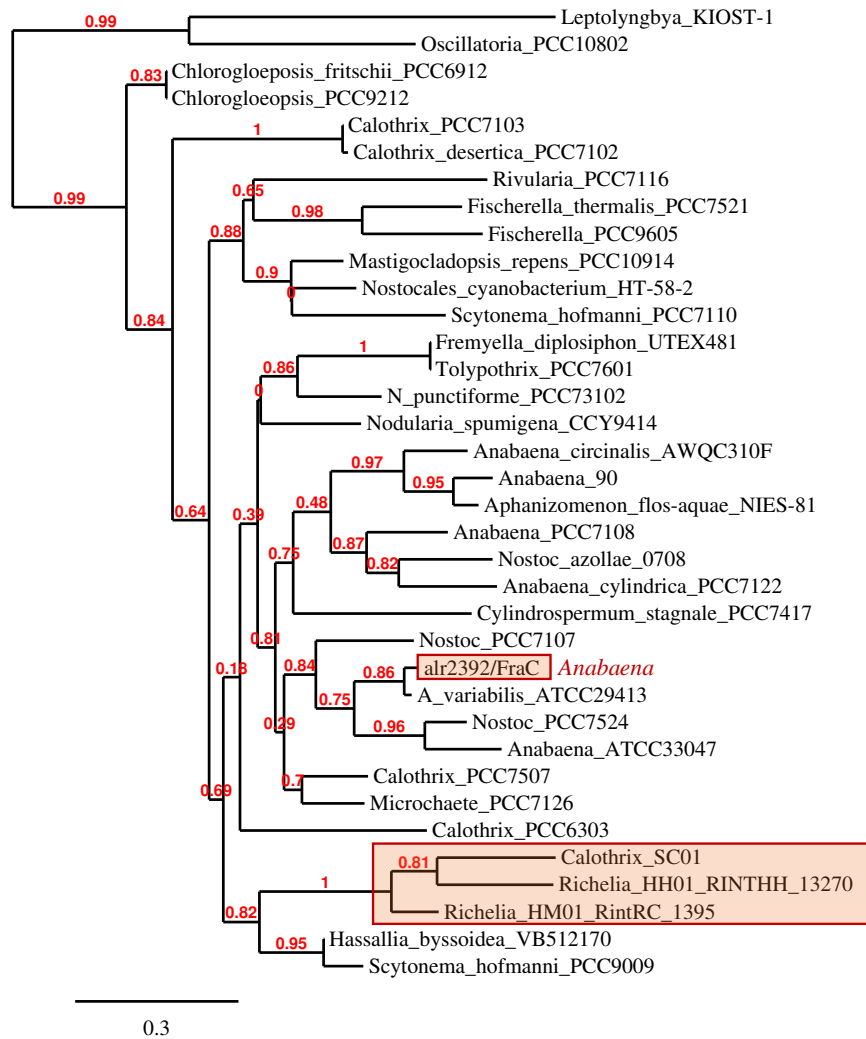






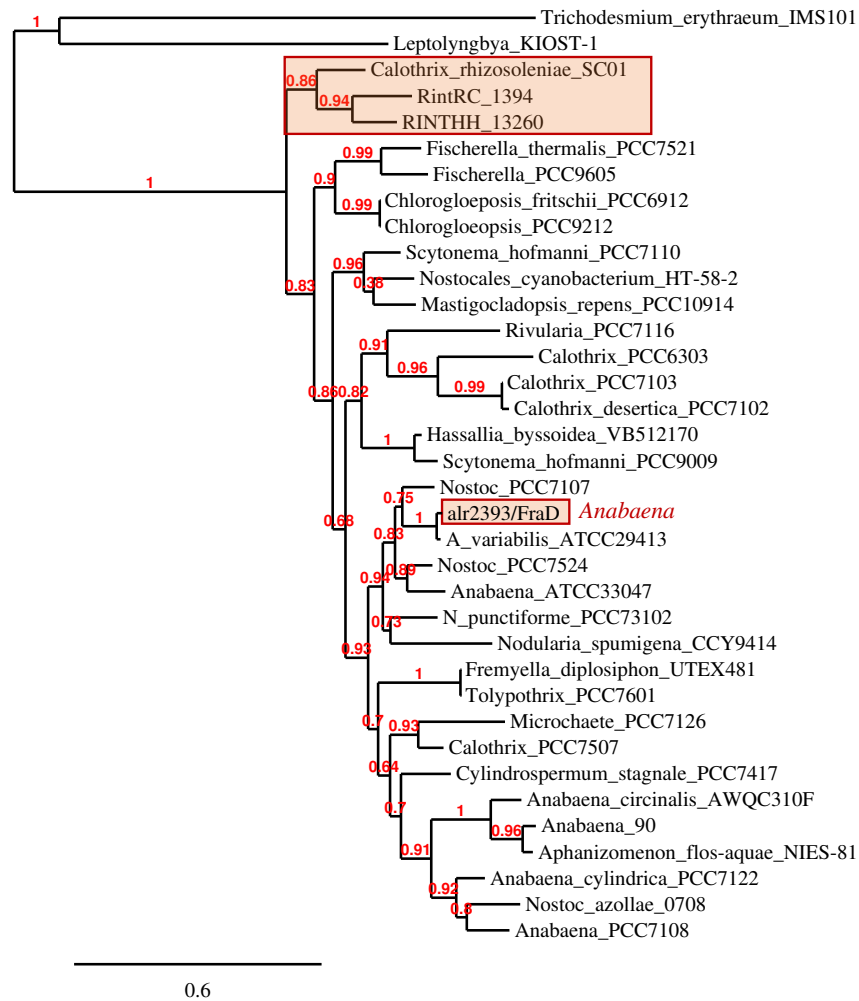
**Fig. 60. Phylogenetic tree of SepJ.** Tree constructed by the Phylogeny program with default parameters (<http://www.phylogeny.fr>). SepJ from *T. erythraeum* (IMS101) was used as an outgroup to root the tree. Note that SepJ from the diatom symbionts cluster together and fall apart from *Anabaena* SepJ.





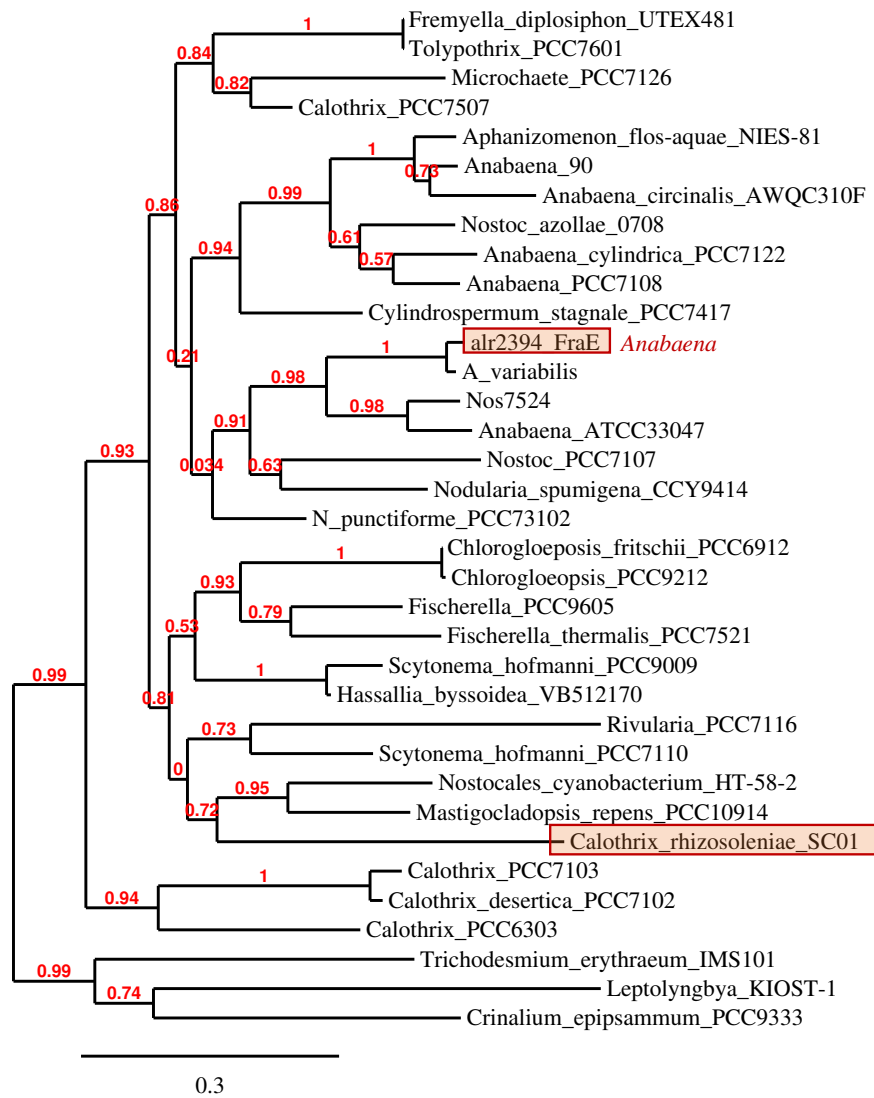
**Fig. 61. Phylogenetic tree of FraC.** Tree constructed by the Phylogeny program with default parameters (<http://www.phylogeny.fr>). FraC from *Leptolyngbya* and *Oscillatoria* strains were used as joint outgroups. Note that FraC from the diatom symbionts cluster together and fall apart from *Anabaena* FraC.





**Fig. 62. Phylogenetic tree of FraD.** Tree constructed by the Phylogeny program with default parameters (<http://www.phylogeny.fr>). FraD from *Leptolyngbya* and *Trichodesmium* strains were used as joint outgroups. Note that FraD from the diatom symbionts cluster together and fall apart from *Anabaena* FraD.





**Fig. 63. Phylogenetic tree of FraE.** Tree constructed by the Phylogeny program with default parameters (<http://www.phylogeny.fr>). FraE from *Crinalium*, *Leptolyngbya* and *Trichodesmium* strains were used as joint outgroups. Note that FraE from CalSC01 falls apart from *Anabaena* FraE.



### 3. Results & Discussion

---

#### Chapter 5

#### Subcellular localization of FraE

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

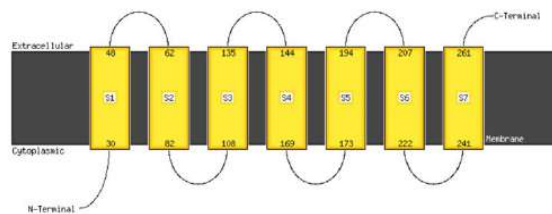
FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



### 3.5 Subcellular localization of FraE

Proteins that have a role in the formation of mature septal junctions such as SepJ, FraC, FraD or, as shown in this Thesis, HglK, have been located to the intercellular septa in *Anabaena* (Flores *et al.*, 2007; Merino-Puerto *et al.*, 2010; see also Results & Discussion Chapter 2). Another protein, FraH, has been shown to be dynamically located, changing from a peripheral localization in vegetative cells of nitrate-grown filaments to the heterocyst poles in diazotrophic filaments (Merino-Puerto *et al.*, 2011a). FraE, encoded in the *fraCDE* operon, is an integral membrane protein that, as in the case of FraC and FraD, is needed to make long filaments under nitrogen deprivation (Merino-Puerto *et al.*, 2010). In contrast to FraC and FraD, it was previously not possible to determine the subcellular localization of FraE. Assuming a cytoplasmic localization of the C-terminus of FraE, a FraE-GFP<sub>mut2</sub> construct was tried, but it did not show GFP fluorescence perhaps indicating that the predicted topology was incorrect (Merino-Puerto, 2011). We have now re-checked the possible topology of FraE, and updated programs suggest a topology with an extra  $\alpha$ -helix positioning the C-terminus in the periplasm (Fig. 64). Hence, in a new attempt to determine the subcellular localization of FraE, we have now used the superfolder-GFP (sfGFP) that, as mentioned earlier (section 3.2.1), folds efficiently in the periplasm.



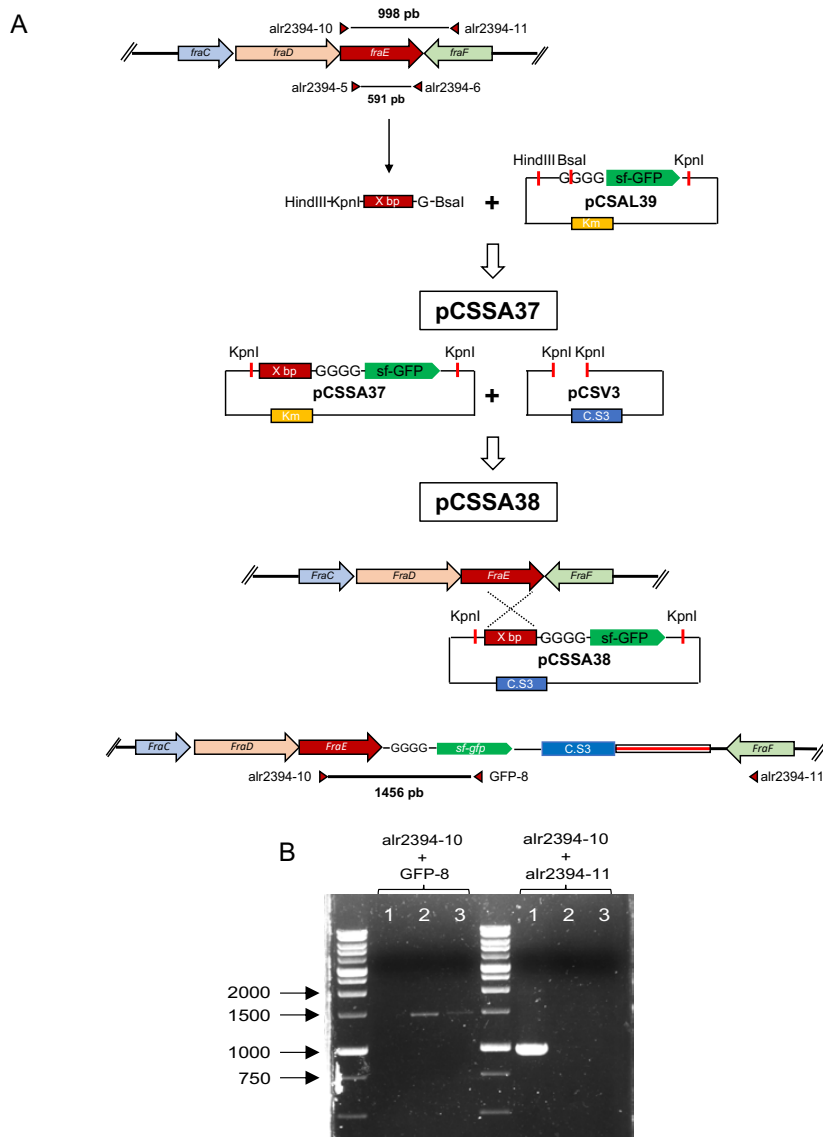
**Fig. 64. Predicted topology of FraE.** Topology predicted from structural analysis performed by the Phyre2 program.

#### 3.5.1 Construction and visualization of FraE-sfGFP

For construction of an *Anabaena* strain bearing a *fraE* gene with the *sf-gfp* fused to its 3' end, 591 bp of *fraE* were amplified by PCR using *Anabaena* DNA as template and primers alr2694-5 and alr2394-6, which lacked the stop codon and included the sequences needed for digestion with restriction enzymes to incorporate the fragment into plasmid pCSAL39 (plasmid details in Table 4 of Materials & Methods) that contains the *sf-gfp* sequence and a sequence encoding a four-glycine linker upstream of *sf-gfp*. Cloning of the PCR product into pCSAL39 produced plasmid pCSSA37 (Fig. 65), which was corroborated by sequencing. Subsequently, the amplified *fraE/4-Gly/sf-gfp* fragment was transferred to plasmid pCSV3 (plasmid details in Table 4 of Materials & Methods), resulting in plasmid pCSSA38 that was transferred to *Anabaena* by conjugation (Fig. 65). Finally, some exconjugants were tested by PCR. The strain carrying the construct to produce the sfGFP attached to the carboxyl terminus of FraE was named CSSA22.



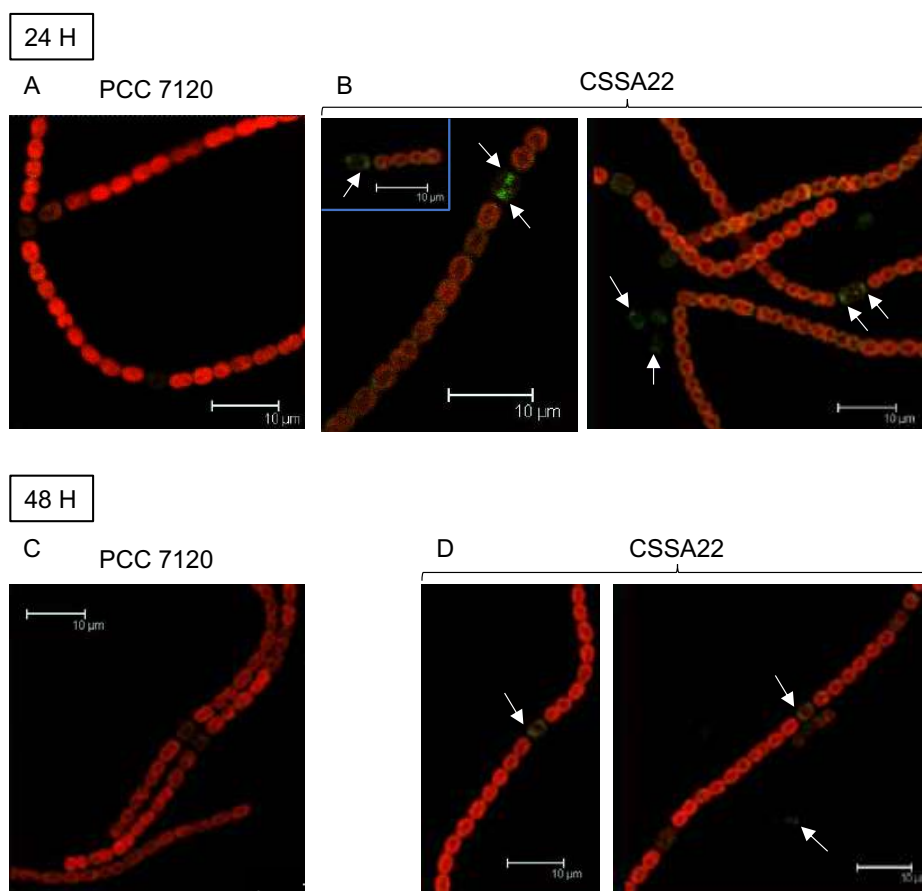




**Fig. 65. Construction and corroboration of *Anabaena fraE-sf-gfp*.** (A) Construction diagram of strain CSSA22 (*fraE-sf-gfp*), indicating the plasmids used. (B) Test of exconjugants by PCR. With primers alr2394-10 and GFP-8 amplification is seen only in the exconjugants that have incorporated *sf-gfp* in their genome, while with primers alr2394-10 and alr2394-11 no amplification is seen in the exconjugants because the pCSV3 plasmid insert makes the fragment too large to be amplified. 1, wild type DNA; 2 and 3, DNA from two independent exconjugants; M, base pair markers.

To study the localization of FraE-sfGFP, an induction for 24 and 48 hours in BG11<sub>0</sub> medium of strain CSSA22 and the wild type as a control was carried out. The analysis of filaments by confocal microscopy showed that FraE was readily visualized at the poles of the heterocysts (Fig. 66).



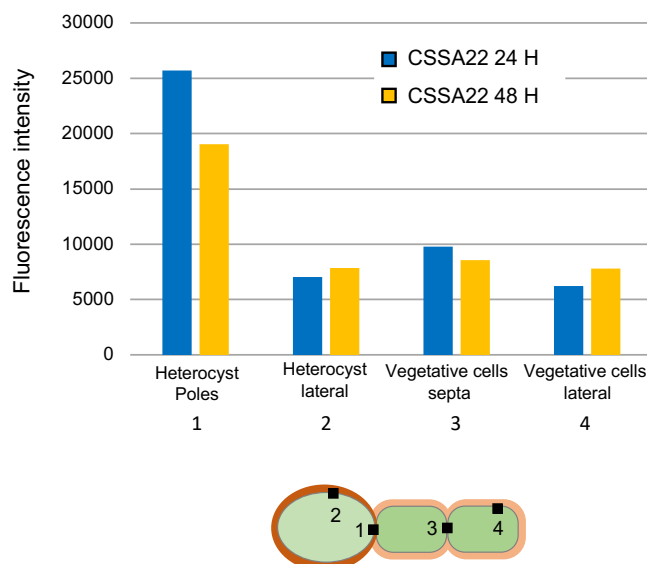


**Fig. 66. Visualization of FraE-sfGFP.** The figure shows overlays of cyanobacterial autofluorescence (red) and sfGFP fluorescence (green) visualized by confocal microscopy. The white arrows indicate the areas of green fluorescence in the heterocysts along of filament. (A and B) Visualization of samples of *Anabaena* wild type and strain CSSA22 (FraE-sfGFP), respectively, after 24 hours of induction in BG11<sub>0</sub> medium. (C and D) *Anabaena* wild type and strain CSSA22 (FraE-sfGFP), respectively, after 48 hours of induction in BG11<sub>0</sub> medium.

However, the quantification of GFP fluorescence showed that FraE-sfGFP was present along the whole filament, but at higher levels in the poles of the heterocysts than in the other studied areas, which were the heterocyst lateral regions and the septa between vegetative cells and the lateral regions of vegetative cells (Fig. 67; see quantitative data including statistical analysis in Table 19). In vegetative cells, FraE-sfGFP was also present at significantly higher levels at the intercellular septa than in lateral location. However, the levels in the vegetative cell septa were only about 1.6 (24 h) and 1.8 (48 h) the lateral levels in vegetative cells, which do not exceed the 2-fold higher levels expected for the juxtaposition of the cytoplasmic membranes of the adjacent cells. Therefore, we conclude that FraE-sfGFP is distributed homogeneously in the periphery of vegetative cells, whereas it is significantly increased at heterocyst poles. Additionally, it was also observed that the intensity of GFP fluorescence at the heterocysts poles decreased significantly



after 48 hours of induction. Such decrease at 48 hours was not observed at significant levels at any of the other locations.



**Fig. 67. Subcellular distribution of FraE-sfGFP.** GFP fluorescence intensity at 24 and 48 hours of induction in BG11<sub>0</sub> medium at the subcellular locations indicated in the scheme. The values shown correspond to the average value at each location, to which the average value of the wild type at the same location was subtracted.


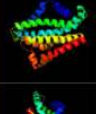

FraE-sfGFP fluorescence at 24 h (relative units; WT background at each position subtracted)				
	Heterocyst poles (1)	Heterocyst lateral (2)	Vegetative cells septa (3)	Vegetative cells lateral (4)
Mean + SEM (n)	25708 ± 1867 (51)	7029 ± 849 (54)	9801 ± 609 (62)	6225 ± 607 (62)
FraE-sfGFP fluorescence at 48 h (relative units; WT background at each position subtracted)				
	Heterocyst poles (1)	Heterocyst lateral (2)	Vegetative cells septa (3)	Vegetative cells lateral (4)
Mean + SEM (n)	19046 ± 1420 (49)	7850 ± 690 (52)	8580 ± 515 (63)	4794 ± 605 (63)

**Table 19. Relative fluorescence from FraE-sfGFP at different subcellular locations.** The table shows the mean intensity of GFP in strain CSSA22 at the poles of heterocysts (1), at lateral position in the heterocysts (2), at the septa between vegetative cells (3) and in lateral position of the vegetative cells (4). Measurements were made after 24 and 48 hours of BG11<sub>0</sub> induction. n, number of zones measured. Student's *t* test analysis showed significant differences ( $p < 0.01$ ) between the following positions: 1 vs 2 (24h, 48h); 1 vs 3 (24 h, 48 h); 1 vs 4 (24 h, 48 h); 2 vs 3 (24 h); 2 vs 4 (48 h); 3 vs 4 (24 h, 48 h); 1 (24 h) vs 1 (48h).



### 3.5.2 Discussion

BLASTp analysis identifies FraE as an integral membrane component (TMD) of an ABC transporter. More precisely, FraE is most similar to the type IV pilus biogenesis protein PilI from *Myxococcus xanthus* (TCDB number 3.A.1.144.5; <http://tcdb.org/>), which is a component of an ABC exporter necessary for pilus assembly and pilus subunit export (Wu *et al.* 1998; reviewed in Herrero *et al.*, 2016). On the other hand, structures recently added to databases has permitted to identify FraE as a protein with structural similarity to the integral membrane component (Wzm) of the ABC transporter involved in export of the O-antigen polysaccharide of Gram-negative bacteria (Fig. 68). This exporter has a channel wide enough to permit the translocation of the linear polysaccharide across the cytoplasmic membrane (Bi *et al.*, 2018). Taking into consideration both similarities (to PilI and to Wzm), we speculate that FraE might be involved in the translocation of a macromolecule outside of the *Anabaena* cytoplasm.

#	Template	Alignment Coverage	3D Model	Confidence	% I.d.	Template Information
1	c5an7D	Alignment		99.9	11	<b>PDB header:</b> transport protein <b>Chain:</b> D; <b>PDB Molecule:</b> transport permease protein; <b>PDBTitle:</b> crystal structure of o-antigen polysaccharide abc-transporter
2	c5an7C	Alignment		99.9	10	<b>PDB header:</b> transport protein <b>Chain:</b> C; <b>PDB Molecule:</b> transport permease protein; <b>PDBTitle:</b> crystal structure of o-antigen polysaccharide abc-transporter
3	c5ni9B	Alignment		99.9	12	<b>PDB header:</b> transport protein <b>Chain:</b> B; <b>PDB Molecule:</b> atp-binding cassette sub-family g member 2; <b>PDBTitle:</b> structure of an abc transporter; part of the structure that could be2 built de novo

**Fig. 68. Top predicted structures of FraE by the Phyre2 program.** The left part shows the FraE fragment that aligns with the target sequence (red line) and its predicted structure. The right part indicates the protein to which that fragment shows highest structural similarity.

This function would be different to that of FraCD, which are components of the septal junctions that can be involved in the intercellular transfer of metabolites such as sucrose (Nürnberg *et al.*, 2015). Hence, the different phenotypes of the *fraC-fraD* and *fraE* mutants (Merino-Puerto *et al.*, 2010, 2011b). The *fraC-fraD* mutants form heterocysts that have a low but measurable nitrogenase activity whereas the *fraC* or *fraD* mutants form heterocysts that show negligible nitrogenase activity (Merino-Puerto *et al.*, 2010). Consistently, whereas the heterocysts of the *fraC* and *fraD* mutants have heterocyst necks with an alteration in the CM facing the CM of the adjacent vegetative cell, the heterocysts of the *fraE* mutant lack an obvious heterocyst neck (Merino-Puerto *et al.*, 2011b). The preferential localization of FraE at the heterocyst poles, i.e., the place of the heterocyst neck, is thus consistent with a role of FraE in the construction of the heterocyst neck. That is, when FraE is missing (such as in the  $\Delta$ *fraE* mutant), the heterocyst neck is not formed. Interestingly, levels of FraE-sfGFP at the heterocyst poles become decreased after 48 h of nitrogen deprivation, consistent with a role in heterocyst formation rather than in heterocyst function. On the other hand, FraE is also observed in the heterocyst lateral membranes and in the vegetative cells' lateral and septal membranes. A role of FraE in the construction of septal junctions along the filament is therefore possible, but this should be a minor role since molecular exchange between vegetative cells is less affected in the *fraE* mutant than in *fraC* or *fraD* mutants (Merino-Puerto *et al.*, 2010).



For some unknown reason, the FraE function may not be needed in *R. intracellularis*, which lacks an evident *fraE* gene (Results & Discussion Chapter 4). The possibility that another membrane protein takes FraE function in the endosymbiont cannot be ruled out, however. Another question is which protein encoded in the *Anabaena* genome could be the ATPase (NBD) partner of FraE to conform a functional ABC exporter. A hypothetical possibility is GlsC, an ABC transporter NBD subunit in *Anabaena* that appears to be a multitask NBD subunit, necessary for glucoside uptake but also for maturation of the intercellular septa as it is necessary to produce a normal number of septal nanopores (Nieves-Mori3n *et al.*, 2017a). Further work will be necessary to establish an association between FraE and GlsC.

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e200022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**

## 4. Summary & Conclusions

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**





#### 4.1 Summary

Multicellularity is a form of biological organization that has evolved independently in many phylogenetic groups. In multicellular organisms, there are generally processes of cell adhesion, communication and differentiation. Multicellularity is observed in several bacterial groups, cyanobacteria being among the oldest multicellular organisms on Earth. Cyanobacteria are characterized by carrying out oxygenic photosynthesis, through which oxygen is released and *assimilatory power* (ATP and reduced ferredoxin/NADPH) is produced; this is then used in the fixation of atmospheric CO<sub>2</sub> and the assimilation of some other oxidized nutrients. Within this phylum both unicellular and multicellular organisms can be found. Those cyanobacteria that behave as multicellular organisms grow forming filaments, and under conditions of nitrogen deficiency, some of them can carry out a process of cellular differentiation producing a cellular type called heterocyst. Heterocysts perform the fixation of atmospheric nitrogen (N<sub>2</sub>), while the rest of the cells of the filament, called vegetative cells, continue with the fixation of CO<sub>2</sub>. Other species of the phylum have the capacity to form akinetes, which are a cellular form of resistance that appears when environmental conditions become adverse, and still others produce hormogonia, small motile filaments with a dispersal function.

This work has been carried out with the model cyanobacterium *Anabaena* sp. PCC 7120, which is a filamentous strain that forms heterocysts. Cyanobacteria possess a Gram-negative cell envelope, and in filamentous cyanobacteria the cytoplasmic membrane individually envelops each cell, while the outer membrane, located outside the peptidoglycan layer, is continuous thus defining a continuous periplasm. When *Anabaena* grows in the absence of combined nitrogen, differentiation of heterocysts occurs, forming a pattern along the filament of one heterocyst every 10-15 vegetative cells. Diazotrophic growth requires an intercellular communication that allows the exchange of nutrients and molecules that regulate differentiation between cells along the filament. This communication takes place through the septal junctions, which are protein structures that communicate adjacent cells directly. Currently, some proteins that are part of the septal junctions are known, including FraC and FraD, and some others that are essential for their correct function or assembly, such as SepJ, SepI, HgIK (demonstrated in this Thesis) or FraE are also known. For the assembly of the septal junctions, perforations in the peptidoglycan between adjacent cells are necessary. These perforations are called nanopores and are located in the peptidoglycan septal disk.

In the first chapter of this work an attempt was made to establish a relationship between nanopores and intercellular communication mediated by the septal junctions, for which we used as tools a fluorescent marker and different *Anabaena* mutants. The fluorescent marker used was calcein, which has the property of diffusing from the outside to the inside of the cell, where it undergoes a process of hydrolysis that prevents it from passing through the cytoplasmic membrane again, remaining inside the cell and emitting fluorescence. By means of FRAP ("Fluorescence Recovery After Photobleaching") experiments it was possible to calculate a quantitative indicator of molecular exchange between adjacent cells of the filament. On the other hand, by isolating the peptidoglycan and visualizing it by transmission electron microscopy, it was possible to quantify the number of nanopores of the strains studied. These tests were carried out on filaments (cultivated in presence or absence of combined nitrogen) of the wild-type strain and mutants of proteins involved in the formation of the septal junctions and of proteins that influence their assembly and functioning. The results obtained allowed us to establish a strong correlation



between the number of nanopores and intercellular exchange of calcein in filaments that had been incubated in a medium without combined nitrogen, but not in those incubated in the presence of combined nitrogen, in which a high proportion of non-communicating cells were observed.

In the second chapter of this Thesis, a characterization of the HglK protein, which belongs to the *pentapeptide-repeat protein* (PRP) family, was performed. HglK is composed of four transmembrane segments in its N-terminal part and of a soluble section of periplasmic location containing the PRP domain in its C-terminal part. Previous studies of HglK showed that mutants that lacked this protein were unable to grow diazotrophically and had an increased spacing between the cells of the filament. In this Thesis, the subcellular localization of HglK was addressed by using *Anabaena* strains that produced HglK fused to superfolder-GFP (sfGFP), to a histidine tail or to the Strep-tag II. By means of confocal and fluorescence microscopy and of immunolocalization assays, it could be established that HglK is predominantly located at the cell poles in *Anabaena*. The study of *hglK* mutants unraveled that HglK is necessary for filament integrity and for normal levels of molecular exchange between the vegetative cells of the filament. However, experiments with mutants of HglK and other septal proteins showed that neither filament fragmentation nor alteration in molecular exchange was greater when HglK inactivation was combined with inactivation of other septal proteins. In the characterization of HglK, it could also be established that the negative diazotrophic growth phenotype of the *hglK* mutant was determined by its inability to fix atmospheric nitrogen that resulted from a defect in gene expression during the differentiation of heterocysts, which was evidenced as a diminished expression of the of the *cox2* operon (cytochrome oxidase of the heterocyst) and of the *nifHDK* genes (nitrogenase) in *hglK* mutants.

The advances obtained through the characterization of HglK and the available knowledge about the SepJ protein suggest the involvement of both proteins in the function or maturation of the intercellular septa of *Anabaena*. SepJ is an integral membrane protein with a focused location at the center of the intercellular septum. SepJ is composed of a permease in its C-terminal part, and a linker (rich in the amino acid proline) and a coiled-coil domain in its N-terminal part, which appears to be periplasmic. In the third chapter of this Thesis we tried to identify proteins with which HglK and SepJ could have some kind of interaction. The experimental approach involved the use of strains of *Anabaena* that produced HglK with the Strep-tag II or SepJ with the Strep-tag II or the GFP. These fusions allowed the purification of the proteins, which was carried out under non-denaturing conditions, trying to avoid the disintegration of possible complexes that could be formed with other proteins. The samples obtained after the purifications were analyzed by mass spectrometry to identify all the proteins present in them. The analysis of the results showed several proteins that could interact with HglK or SepJ. The first notable result is the possible interaction between both SepJ and HglK, although co-purification of the two proteins was only observed in SepJ isolates. The presence of two accompanying common proteins in SepJ and HglK isolates is also noteworthy. These are the periplasmic proteins Alr2937 and Tic22 (Alr0114), the latter of which has an important role in the formation of the outer membrane of *Anabaena*. Finally, Alr1690, a protein anchored to the cytoplasmic membrane with periplasmic peptidoglycan-binding domains, was detected for SepJ, while no protein was found to interact only with HglK.

Continuing with the study of septal proteins, we wanted to deepen our knowledge of the SepJ, FraC and FraD proteins from other filamentous cyanobacteria that form heterocysts, specifically



cyanobacteria that are symbiotic with marine diatoms. In the fourth chapter of this Thesis, we studied these septal proteins encoded in the genomes of the cyanobacteria *Richelia intracellularis* (RintHH01) and *Calothrix rhizosoleniae* (CalSC01). *R. intracellularis* is an obligatory endosymbiont that lives in the cytoplasm of the diatom *Hemiaulus hauckii*, while *C. rhizosoleniae* is a facultative exosymbiont of the diatom *Chaetoceros compressus*, which can proliferate either extracellularly anchored to the diatom or as a free-living organism. In *Anabaena*, *fraC* and *fraD* are part of the *fraCDE* operon which is also present in CalSC01 and, only *fraCD*, in RintHH01. Regarding SepJ, both RintHH01 and CalSC01 contain genes that determine proteins homologous to SepJ of *Anabaena*. *Anabaena sepJ* and *fraC-fraD* mutants show a strong filament fragmentation phenotype and inability to grow diazotrophically, two phenotypes that are easy to study. The RintHH01 *sepJ* and *fraCD* genes and the CalSC01 *sepJ* and *fraCDE* genes were obtained by chemical synthesis, in the case of CalSC01 *sepJ* with a codon usage adapted to *Anabaena*, to be introduced into *Anabaena* mutants carrying the *sepJ* (CSVM34 strain) or *fraCD* (CSVT22 strain) deletions, respectively. Complementation of CSVM34 with RintHH01 *sepJ* and CalSC01 *sepJ* was tried with expression of these genes from the natural *sepJ* promoter of *Anabaena*. However, neither *sepJ* from RintHH01 nor *sepJ* from CalSC01 could complement the mutant phenotype, even partially. However, it was possible to obtain the wild-type phenotype when the mutant was complemented with *sepJ* from *Anabaena*, used as a positive control. On the other hand, the *fraC-fraD* genes of RintHH01 were transferred to strain CSVT22 with two different strategies: in one their expression took place from the natural *fraC* promoter of *Anabaena*, and in the other the expression took place from the *Anabaena glnA* promoter (strong expression under nitrogen deficiency) inserted in a replicative plasmid. Complementation of CSVT22 with the *fraC-fraD-fraE* genes from CalSC01 was only attempted with expression from the *glnA* promoter in a replicative plasmid. Again, none of the attempts with the *fra* genes were successful. Furthermore, gene expression from the *glnA* promoter caused detrimental effects in *Anabaena*, as we found in controls using *Anabaena's* own genes. This suggests that we may have not expressed the genes at levels appropriate for functionality.

The last chapter of this Thesis was devoted to a brief study of the FraE protein, which is encoded in the *fraC-fraD-fraE* operon of *Anabaena*. FraE is a membrane protein necessary for filament integrity. Here we studied its localization, for which an *Anabaena* strain producing a FraE-sfGFP fusion was constructed. Fluorescence analysis tests of GFP carried out with this strain showed that FraE is mainly located at the poles of heterocysts, and that its levels at this location decrease after 48 hours of induction in a medium without combined nitrogen.



## 4.2 Conclusions

1. In the filamentous, heterocystous cyanobacterium *Anabaena* sp. PCC 7120 there is a strong correlation between the number of septal nanopores and intercellular exchange of calcein when the filaments are incubated in a medium without combined nitrogen. This observation supports the idea that the nanopores host the septal junctions that mediate intercellular molecular exchange, which is essential for diazotrophic growth.
2. HglK is a *pentapeptide-repeat protein* necessary for diazotrophic growth in *Anabaena*. The *hglK* mutants show decreased expression of some genes induced during heterocyst differentiation and are affected in filament integrity, nanopore formation and intercellular molecular exchange. The HglK protein has a predominant location at the cell poles. We conclude that HglK contributes to the formation of functional septal structures necessary for the differentiation of mature heterocysts in *Anabaena*.
3. Mass spectrometric analysis of preparations of HglK- and SepJ-tagged proteins isolated from *Anabaena* has identified several proteins that could interact with them. For SepJ, we highlight HglK itself and the peptidoglycan-binding protein Alr1690. In addition, two periplasmic proteins were detected that could interact with both SepJ and HglK: Alr2947 and the chaperone Tic22 (Alr0114). Thus, SepJ and HglK are two septal proteins that may interact with each other and with some other periplasmic proteins.
4. The filamentous, heterocystous cyanobacteria *Richelia intracellularis* and *Calothrix rhizosoleniae*, which are symbiotic with marine diatoms, contain genes coding for SepJ and FraC, FraD and, in the case of *C. rhizosoleniae*, FraE. However, we have not been successful in complementing *sepJ* and *fraC-fraD* mutants of *Anabaena* with the genes of these symbiotic cyanobacteria.
5. The FraE protein has a preferential location at the heterocyst poles in *Anabaena*, which is consistent with the fact that *fraE* mutants of this cyanobacterium produce heterocysts that lack the characteristic polar structure known as *heterocyst neck*.



## 5. References

---

ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



ÁMBITO- PREFIJO

**GEISER**

Nº registro

**00008745e2000022470**

CSV

**GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1**

DIRECCIÓN DE VALIDACIÓN

**<https://sede.administracionespublicas.gob.es/valida>**

FECHA Y HORA DEL DOCUMENTO

**12/06/2020 08:18:06 Horario peninsular**



GEISER-42b6-8a54-06c6-41ac-9516-66b9-50c5-9ad1

- Abed RMM, Dobretsov S & Sudesh K (2009) Applications of cyanobacteria in biotechnology. *J Appl Microbiol* 106: 1-12.
- Bateman A, Murzin AG & Teichmann SA (1998) Structure and distribution of pentapeptide repeats in bacteria. *Protein Sci* 7: 1477-80.
- Bateman A, Birney E, Durbin R, Eddy SR, Howe KL & Sonnhammer EL (2000) The Pfam protein families database. *Nucleic Acids Res* 30: 276-80.
- Bauer CC, Buikema WJ, Black K & Haselkorn R (1995) A short-filament mutant of *Anabaena* sp. strain PCC 7120 that fragments in nitrogen-deficient medium. *J Bacteriol* 177: 1520-6.
- Berendt S, Lehner J, Zhang YV, Rasse TM, Forchhammer K & Maldener I (2012) Cell wall amidase AmiC1 is required for cellular communication and heterocyst development in the cyanobacterium *Anabaena* PCC 7120 but not for filament integrity. *J Bacteriol* 194: 5218-27.
- Bi Y, Mann E, Whitfield C & Zimmer J (2018) Architecture of a channel-forming O-antigen polysaccharide ABC transporter. *Nature* 553: 361- 65.
- Black TA, Cai Y & Wolk CP (1993) Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol Microbiol* 9: 77-84.
- Black K, Buikema WJ & Haselkorn R (1995) The *hglK* gene is required for localization of heterocyst-specific glycolipids in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 177: 6440-8.
- Bonner JT (1998) The origins of multicellularity. *Integr Biol* 1: 28-36.
- Bornikoel J, Carrión A, Fan Q, Flores E, Forchhammer K, Mariscal V, Mullineaux CW, Perez R, Silber N, Wolk CP & Maldener I (2017) Role of two cell wall amidases in septal junction and nanopore formation in the multicellular cyanobacterium *Anabaena* sp. PCC 7120. *Front Cell Infect Microbiol* 7:386.
- Bornikoel J, Staiger J, Madlung J, Forchhammer K & Maldener I (2018) LytM factor Alr3353 affects filament morphology and cell-cell communication in the multicellular cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol* 108: 187-203.
- Bos MP, Robert V & Tommassen J (2007) Biogenesis of the Gram-negative bacterial outer membrane. *Annu Rev Microbiol* 61: 191-214.
- Boyer HW & Roulland-Dussoix DD (1969) A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol* 41: 459-72.
- Brouwer EM, Ngo G, Yadav S, Ladig R & Schleiff E (2019) Tic22 from *Anabaena* sp. PCC 7120 with holdase function involved in outer membrane protein biogenesis shuttles between plasma membrane and Omp85. *Mol Microbiol* 111: 1302-16.
- Buchko GW, Ni S, Robinson H, Welsh EA, Pakrasi HB & Kennedy MA (2006) Characterization of two potentially universal turn motifs that shape repeated five-residues fold-crystal structure of a luminal pentapeptide repeat protein from *Cyanothece* 51142. *Protein Sci* 11: 2579-95.
- Buchko GW, Robinson H, Pakrasi HB & Kennedy MA (2008) Insights into the structural variation between pentapeptide repeat proteins—crystal structure of Rfr23 from *Cyanothece* 51142. *J Struct Biol* 162: 184-92.
- Buikema WJ & Haselkorn R (1991) Isolation and complementation of nitrogen fixation mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 173: 1879-85.
- Burnat M, Herrero A & Flores E (2014) Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocysts forming cyanobacterium. *Proc Natl Acad Sci USA* 111: 3823-8.





- Busby S & Ebricht RH (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* 293:199-213.
- Cai Y & Wolk CP (1990) Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *J Bacteriol* 172: 3138-45.
- Callahan SM & Buikema WJ (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Microbiol* 40: 941-50.
- Cameron JC, Sutter M & Kerfeld CA (2014) The carboxysome: function structure and cellular dynamics. In Flores E & Herrero A (eds.), *The cell biology of Cyanobacteria*. Caister Academic Press, Norfolk, UK, 171-188.
- Caputo A, Nylander JAA & Foster RA (2019) The genetic diversity and evolution and diatom-diazotroph associations highlights traits favoring symbiont integration. *FEMS Microbiol lett* 1:366.
- Cardemil L & Wolk CP (1979) The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Structure of the basic repeating unit. *J Biol Chem* 254: 736-41.
- Castenholz RW (2001) Phylum BX *Cyanobacteria* (*Oxygenic Photosynthetic Bacteria*). *Bergey's Manual of Systematic Bacteriology*, Vol. 1: *The Archaea and the Deeply Branching and Phototrophic Bacteria* 2nd ed, New York, Springer, pp. 473-599.
- Chandler LE, Bartsevich VV & Pakrasi HB (2003) Regulation of manganese uptake in *Synochocystis* 6803 by RfrA, a member of a novel family of proteins containing a repeated five-residues domain. *Biochemistry* 42: 5508-14.
- Claessen D, Rozen DE, Kuipers OP, Sogaard-Andersen L & van Wezel GP (2014) Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol* 12:115-24.
- Corrales-Guerrero L, Mariscal V, Flores E & Herrero A (2013) Functional dissection and evidences for intercellular transfer of the heterocyst-differentiation PatS morphogen. *Mol Microbiol* 88: 1093-105.
- Corrales-Guerrero L, Mariscal V, Nürnberg DJ, Elhai J, Mullineaux CW, Flores E & Herrero A (2014a) Subcellular localization and clues for the function of the HetN factor influencing heterocyst distribution in *Anabaena* sp. PCC 7120. *J Bacteriol* 196: 3452-60.
- Corrales-Guerrero L, Flores E & Herrero A (2014b) Relationships between the ABC-transporter HetC and peptides that regulate the spatiotemporal pattern of heterocyst distribution in *Anabaena*. *PLoS One* 9: e104571.
- Corrales-Guerrero L, Tal A, Arbel-Goren R, Mariscal V, Flores E, Herrero A & Stavans J (2015). Septal fluctuations in expression of the heterocyst differentiation regulatory gene *hetR* in *Anabaena* filaments. *PLoS Genet* 11: e1005031.
- Dinh T & Bernhardt TG (2011) Using superfolder green fluorescent protein for periplasmic protein localization studies. *J Bacteriol* 193: 4984-87
- Ehira S & Ohmori M (2006a) NrrA, a nitrogen-responsive regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 59: 1692-1703.
- Ehira S & Ohmori M (2006b) NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol*, 188: 8520-5.
- Ehira S & Ohmori M (2014) NrrA directly regulates expression of the *fraF* gene and antisense RNAs for *fraE* in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *Microbiology* 160: 844-850.
- Ekman M, Picossi S, Campbell EL, Meeks JC & Flores E (2013) A *Nostoc Punctiforme* sugar transporter necessary to establish a cyanobacterium-plant symbiosis. *Plant Physiol* 161:1984-92.



- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE & Finn RD (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47: D427-32.
- Elhai J & Wolk CP (1988a) A versatile class of positive-selection vectors based on the nonviability of palindrome-containing plasmids that allows cloning into long polilinkers. *Gene* 68: 119-38.
- Elhai J & Wolk CP (1988b) Conjugal transfer of DNA to cyanobacteria. *Methods Enzymol* 167: 747-54.
- Elhai J, Vepritskiy A, Muro-Pastor AM, Flores E & Wolk CP (1997) Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 179: 1998-2005.
- Ermakova M, Battchikova N, Richaud P, Leino H, Kosourov S, Isojärvi J, Peltier G, Flores E, Cournac L, Allahverdiyeva Y & Aro EM (2014) Heterocyst-specific flavodiiron protein Flv3B enables oxic diazotrophic growth of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci USA* 111: 11205-10.
- Ernst A, Black T, Cai Y, Panoff JM, Tiwari DN & Wolk CP (1992) Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. *J Bacteriol* 174: 6025-32.
- Escudero L, Mariscal V & Flores E (2015) Functional dependence between septal protein SepJ from *Anabaena* sp strain PCC 7120 and an amino acid ABC-type uptake transporter. *J Bacteriol* 197: 2721-30.
- Espinosa J, Forchhammer K, Burillo S & Contreras A (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Mol Microbiol* 61: 457-69.
- Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F & Contreras A (2014) PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. *Proc Natl Acad Sci USA* 111: E2423-30.
- Fan Q, Huang G, Lechno-Yossef S, Wolk CP, Kaneko T & Tabata S (2005) Clustered genes required for synthesis and deposition of envelope glycolipids in *Anabaena* sp. PCC 7120. *Mol Microbiol* 58: 227-43.
- Fay P (1992) Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* 56: 340-73.
- Fiedler G, Arnold M, Hannus S & Maldener I (1998) The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 27: 1193-202.
- Flaherty BL, Van Nieuwerburgh F, Head SR & Golden JW (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* 14: 563-75.
- Flores E & Herrero A (2014) The cyanobacteria: morphological diversity in a photoautotrophic lifestyle. *Perspectives in Phycology* Vol. 1: 63-72.
- Flores E, Frías JE, Rubio LM & Herrero A (2005) Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth Res* 83: 117-33.
- Flores E, Herrero A, Wolk CP & Maldener I (2006) Is the periplasm continuous in filamentous multicellular cyanobacteria? *Trends Microbiol* 14: 439-43.
- Flores E, Pernil R, Muro-Pastor AM, Mariscal V, Maldener I, Lechno-Yossef S, Fan Q, Wolk CP & 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: 3884-90.



- Flores E, Herrero A, Forchhammer K & Maldener I (2016) Septal junctions in filamentous heterocyst-forming cyanobacteria. *Trends Microbiol* 24: 79-82.
- Flores E, Nieves-Mori3n M & Mullineaux CW (2018) Cyanobacterial septal junctions: properties and regulation. *Life* 9: 1. doi: 10.3390/life9010001.
- Flores E, Picossi S, Valladares A & Herrero A (2019) Transcriptional regulation of development in heterocyst-forming cyanobacteria. *Biochim Biophys Acta Gene Regul Mech* 1862: 673-684.
- Foster RA, Goebel NL & Zehr JP (2010) Isolation of *Calothrix rhizosoleniae* strain SC01 from Chaetoceros spp. diatoms of the subtropical N. Pacific Ocean. *J Phycol* 46: 1028-37.
- Frías JE, Mérida A, Herrero A, Martín-Nieto J & Flores E (1993) General distribution of the nitrogen control gene ntcA in cyanobacteria. *J Bacteriol* 175: 5710-3.
- Giddings TH & Staehelin LA (1978) Plasma membrane architecture of *Anabaena cylindrica*: occurrence of microplasmodesmata and changes associated with heterocyst development and the cell cycle. *Cytobiologie* 16: 235-49.
- Giddings TH & Staehelin LA (1981) Observation of microplasmodesmata in both heterocyst-forming and non-heterocyst forming filamentous cyanobacteria by freeze-fracture electron microscopy. *Arch. Microbiol* 129: 295-8.
- Giovannoni SJ, Turner S, Olsen GJ, Barners S, Lane DJ & Pace NR (1988) Evolutionary relationships among cyanobacteria and green chloroplast. *J Bacteriol* 170: 3584-92.
- Golden JW & Yoon HS (2003) Heterocyst development in *Anabaena*. *Curr Opin Microbiol* 6: 557-63.
- Gupta M & Carr NG (1981) Enzymes activities related to cyanophycin metabolism in heterocysts and vegetative cells of *Anabaena* spp. *J Gen Microbiol* 125: 17-23.
- Hahn A & Schleiff E (2014) The cell envelope. In Flores E & Herrero A (eds) *The cell biology of Cyanobacteria*. Caister Academic Press, Norfolk, UK, pp. 29-87.
- Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166: 557-80.
- Hegde SS, Vetting MW, Roderick SL, Mitchenall LA, Maxwell A, Takiff HE & Blanchard JS (2005) A fluoroquinolone resistance protein from *Mycobacterium tuberculosis* that mimics DNA. *Science* 308: 1480-3.
- Herrero A, Muro-Pastor AM & Flores E (2001) Nitrogen control in cyanobacteria. *J Bacteriol* 183: 411-25.
- Herrero A, Muro-Pastor AM, Valladares A & Flores E (2004) Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev* 28: 469-87.
- Herrero A, Picossi S & Flores E (2013) Gene expression during heterocyst differentiation. *Adv Bot Res* 65: 281-329.
- Herrero A, Stavans J & Flores E (2016) The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiol Rev* 40: 831-54.
- Higa KC, Rajagopalan R, Risser DD, Rivers OS, Tom SK, Videau P & Callahan SM (2012) The RGSGR amino acid motif of the intercellular signaling protein, HetN, is required for patterning of heterocysts in *Anabaena* sp. strain 7120. *Mol Microbiol* 83: 683-93.
- Hilton JA (2014) Ecology and evolution of diatom-associated cyanobacteria through genetic analyses. Dissertation Thesis, Santa Cruz CA: University of California.



- Hilton JA, Foster RA, Tripp HJ, Carter BJ, Zehr, JP & Villareal TA (2013) Genomic deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont. *Nat Commun* 4: 1767.
- Hohmann-Marriott MF & Blankenship RE (2011) Evolution of photosynthesis. *Annu Rev Plant Biol* 62: 515-48.
- Hoiczyk E & Hansel A (2000) Cyanobacterial cell walls: news from an unusual prokaryotic envelope. *J Bacteriol* 182: 1191-9.
- Hu HX, Jiang YL, Zhao MX, Cai K, Liu S, Wen B, Lv P, Zhang Y, Peng J, Zhong H, Yu HM, Ren YM, Zhang Z, Tian C, Wu Q, Oliveberg M, Zhang CC, Chen Y & Zhou CZ (2015) Structural insights into HetR-PatS interaction involved in cyanobacterial pattern formation. *Sci Rep* 5:16470.
- Huang G, Fan Q, Lechno-Yossef S, Wojciuch E, Wolk CP, Kaneko T & Tabata S (2005) Clustered genes required for the synthesis of heterocyst envelope polysaccharide in *Anabaena* sp. PCC 7120. *J Bacteriol* 187: 1114-23.
- Jakimowicz D & van Wezel GP (2012) Cell division and DNA segregation in *Streptomyces*: how to build a septum in the middle of nowhere. *Mol Microbiol* 85: 393-404.
- Jüttner F (1983) <sup>14</sup>C-labeled metabolites in heterocysts and vegetative cells of *Anabaena cylindrica* filaments and their presumptive function as transport vehicles of organic carbon and nitrogen. *J Bacteriol* 155: 628-33.
- Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakazaki N, Shimpo S, Sugimoto M, Takazawa M, Yamada M, Yasuda M & Tabata S (2001) Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* 8: 205-13.
- Kim Y, Joachimiak G, Ye Z, Binkowski TA, Zhang R, Gornicki P, Callahan SM, Hess WR, Haselkorn R, Joachimiak A (2011) Structure of transcription factor HetR required for heterocyst differentiation in cyanobacteria. *Proc Natl Acad Sci USA* 108: 10109-14.
- Knoll AH (2004) *Life on a Young Planet: The First Three Billion Years of Evolution on Earth* Princeton, NJ Princeton University Press.
- Knoll AH (2008) Cyanobacteria and Earth history. Herrero A & Flores E, *The Cyanobacteria: Molecular Biology, Genomics and Evolution*, Caister Academic Press, Norfolk, UK, pp. 1-19.
- Koebnik R, Locher KP & Van Gelder P (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* 37: 239-53.
- Kuhn I, Peng L, Bedu S & Zhang CC (2000) Developmental regulation of cell division protein FstZ in *Anabaena* sp. strain PCC 7120, a cyanobacterium capable of terminal differentiation. *J Bacteriol* 182: 4640-43.
- Lang NJ & Fay P (1971) The heterocysts of blue-green algae. II. Details of ultrastructure. *Proc Roy Soc Lond* 178: 193-203.
- Lang NJ, Simon RD & Wolk CP (1972) Correspondence of cyanophycin granules with structured granules in *Anabaena cylindrica*. *Archiv für Mikrobiologie* 83: 313-20.
- Lechno-Yossef S, Fan Q, Wojciuch E, Wolk CP (2011) Identification of ten *Anabaena* sp. genes that under aerobic conditions are required for growth on dinitrogen but not for growth on fixed nitrogen. *J Bacteriol* 193: 3482-9.
- Lehner J, Zhang YV, Berendt S, Rasse TM, Forchhammer K & Maldener I (2011) The morphogene AmiC2 is pivotal for multicellular development in the cyanobacteria *Nostoc punctiforme*. *Mol Microbiol* 79: 1655-69.



- Lehner J, Berendt S, Dörsam B, Pérez R, Forschhammer K & Maldener I (2013) Prokaryotic multicellularity: a nanopore array for bacterial cell communication. *FASEB J* 27: 2293-2300.
- Li JH, Laurent S, Konde V, Bédu S & Zhang CC (2003) An increase in the levels of 2-oxoglutarate promotes heterocyst development in the cyanobacterium *Anabaena* sp. PCC 7120. *Microbiology* 149: 3257-63
- Liberton M & Pakrasi HB (2008) Membrane systems in cyanobacteria. In Herrero A & Flores E (eds) *The Cyanobacteria: molecular biology, genomics and evolution*. Caister Academic Press, Norfolk, UK, pp. 271-287.
- Liu D & Golden JW (2002) *hetL* overexpression stimulates heterocyst formation in *Anabaena* sp. strain PCC 7120. *Journal of Bacteriology* 184: 6873-81.
- Liu J & Wolk CP (2011) Mutation 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: 6070-4.
- López-Igual R, Flores E & Herrero A (2010) Inactivation of a heterocyst-specific invertase indicates a principal role of sucrose catabolism in the heterocyst of *Anabaena* sp. *J Bacteriol* 192: 5526-33.
- López-Igual R, Lechno-Yossef S, Fan Q, Herrero A, Flores E, Wolk CP (2012) A major facilitator superfamily protein, HepP, is involved in formation of the heterocyst envelope polysaccharide in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 194:4677-87.
- Luque I & Forchhammer K (2008) Nitrogen assimilation and C/N balance sensing. In: Herrero A & Flores E (eds), *The Cyanobacteria. Molecular biology, genomics and evolution*. Caister Academic Press, UK, pp. 335-382.
- Luque I, Flores E & Herrero A (1994) Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* 13:2862-9.
- Lyons TW, Reinhard CT & Planavsky NJ (2014) The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506: 307-15.
- Mackinney G (1941) Absorption of light by chlorophyll solutions. *J Biol Chem* 140:109-12.
- Maldener I, Hannus S & Kammerer M (2003) Description of five mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst differentiation and identification of the transposon-tagged genes. *FEMS Microbiol Lett* 224: 205-13.
- Maldener I, Summers ML & Sukenik A (2014) Cellular differentiation in filamentous cyanobacteria. In Flores E & Herrero A (eds) *The cell biology of Cyanobacteria*. Caister Academic Press, Norfolk, UK, pp. 29-304.
- Mariscal V (2014) Cell-cell joining proteins in heterocyst-forming cyanobacteria. In Flores E & Herrero A (eds). *The cell Biology of Cyanobacteria*. Caister Academic Press, Norfolk, UK, pp. 293-304.
- Mariscal V & Flores E (2010) Multicellularity in a heterocyst-forming cyanobacterium: pathways for intercellular communication. *Adv Exp Med Biol* 675:123-35.
- Mariscal V, Herrero A & Flores E (2007) Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium. *Mol Microbiol* 65: 1139-45.
- Mariscal V, Herrero A, Nenninger A, Mullineaux CW & Flores E (2011) Functional dissection of the three-domain SepJ protein joining the cells in cyanobacterial trichomes. *Mol Microbiol* 79: 1077-88.
- Mariscal V, Nürnberg DJ, Herrero A, Mullineaux CW & Flores E. (2016) Overexpression of SepJ alters septal morphology and heterocyst pattern regulated by diffusible signals in *Anabaena*. *Mol Microbiol* 10: 968-81





- Martín-Figueroa E, Navarro F & Florencio FJ (2000) The GS-GOGAT pathway is not operative in the heterocyst. Cloning and expression of *glsF* gene from cyanobacterium *Anabaena* sp. PCC 7120. *FEBS Lett* 476: 282-6.
- Meeks JC, Wolk CP, Thomas J, Lockau W, Shaffer PW, Austin SM, Chien WS & Galonsky A (1977) The pathways of assimilation of  $^{13}\text{NH}_4^+$  by the cyanobacterium, *Anabaena cylindrica*. *J Biol Chem* 252: 7894-900.
- Meeks JC & 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.
- Merino-Puerto V (2011) Filament integrity and intercellular communication genes in *Anabaena* sp. PCC 7120. Ph. D. Thesis, Universidad de Sevilla, Seville, Spain.
- Merino-Puerto V, Mariscal V, Mullineaux CW, Herrero A & Flores E (2010) Fra proteins influencing filament integrity, diazotrophy and localization of septal protein SepJ in the heterocyst-forming cyanobacterium *Anabaena* sp. *Mol Microbiol* 75: 1159-70.
- Merino-Puerto V, Mariscal V, Schwarz H, Maldener I, Mullineaux CW, Herrero A & Flores E (2011a) FraH is required for reorganization of intracellular membranes during heterocyst differentiation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 193: 6815-23.
- Merino-Puerto V, Schwarz H, Maldener I, Mariscal V, Mullineaux CW, Herrero A & Flores E (2011b) FraC/FraD-dependent intercellular molecular exchange in the filaments of a heterocyst-forming cyanobacterium, *Anabaena* sp. *Mol Microbiol* 82: 87-98.
- 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: 3957-66.
- Mitschke J, Vioque A, Haas F, Hess WR, Muro-Pastor AM (2011) Dynamics of transcriptional start site selection during nitrogenase stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc Natl Acad Sci USA* 108: 20130-5.
- Mohamed A & Jansson C (1989) Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant Mol Biol* 13: 693-700.
- Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK & Uversky VN (2006) Analysis of molecular recognition features (MoRFs). *J Mol Biol* 362:1043-59.
- Montesinos ML, Muro-Pastor AM, Herrero A & Flores E (1998) Ammonium/methylammonium permeases of a Cyanobacterium. Identification and analysis of three nitrogen-regulated ant genes in *synechocystis* sp. PCC 6803. *J Biol Chem* 273: 31463-70.
- Moslavac S, Bredemeier R, Mirus O, Granvogel B, Eichacker LA & Schleiff E (2005) Proteomic analysis of the outer membrane of *Anabaena* sp. strain PCC 7120. *J Proteome Res* 4: 1330-8.
- Moslavac S, Nicolaisen K, Mirus O, Al Dehni F, Pernil R, Flores E, Maldener I & Schleiff E (2007) A TolC-like protein is required for heterocyst development in *Anabaena* sp PCC 7120. *J Bacteriol* 189: 7887-95.
- Mullineaux CW, Mariscal V, Nennering A, Khanum H, Herrero A, Flores E & Adams DG (2008) Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. *EMBO J* 27: 1299-308.
- Muñoz-Dorado J, Marcos-Torres FJ, García-Bravo E, Moraleda-Muñoz A & Pérez J (2016) Myxobacteria: moving, killing, feeding, and surviving together. *Front Microbiol* 7: 781.
- Muro-Pastor AM, 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 differentiation. *Mol Microbiol* 44: 1377-85.



- Muro-Pastor MI, Reyes JC & Florencio FJ (2001) Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. *J Biol Chem* 274: 38320-85.
- Murray NE, Brammar WJ & Murray K (1977) Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol Gen Genet* 150: 53-61.
- Nakamoto H & Bardwell JC (2004) Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta* 1694: 111-9.
- Nayar AS, Yamaura H, Rajagopalan R, Risser DD & Callahan SM (2007) FraG is necessary for filament integrity and heterocyst maturation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Microbiology* 153: 601-1.
- Ni S, Sheldrick GM, Benning MM & Kennedy MA (2009) The 2A resolution crystal structure of HetL, a pentapeptide repeat protein involved in regulation of heterocyst differentiation on the cyanobacterium *Nostoc* sp. strain PCC 7120. *J Struct Biol* 165: 47-52.
- Nicolaisen K, Hahn A, Schleiff E (2009a) The cell wall in heterocyst formation by *Anabaena* sp. PCC 7120. *J Basic Microbiol* 49: 5-24.
- Nicolaisen K, Mariscal V, Bredemeier R, Pernil R, Moslavac S, López-Igual R, Maldener I, Herrero A, Schleiff E & Flores E (2009b) 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.
- Nieves-Morión M & Flores E (2018) Multiple ABC glucoside transporters mediate sugar-stimulated growth in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *Environ Microbiol Rep* 10: 40-48.
- Nieves-Morión M, Lechno-Yossef S, López-Igual R, Frías JE, Mariscal V, Nürnberg DJ, Mullineaux CW, Wolk CP & Flores E (2017a) Specific glucoside transporters influence septal structure and function in the filamentous heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 14:199.
- Nieves-Morión M, Mullineaux CW & Flores E (2017b) Molecular diffusion through cyanobacterial septal junctions. *mBio* 8: e01756.
- Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67: 593-656.
- Nürnberg DJ, Mariscal V, Parker J, Mastroianni G, Flores E & Mullineaux CW (2014) Branching and intercellular communication in the Section V cyanobacterium *Mastigocladus laminosus*, a complex multicellular prokaryote. *Mol Microbiol* 91: 935-49.
- Nürnberg DJ, Mariscal V, Bornikoel J, Nieves-Morión M, Krauß N, Herrero A, Maldener I, Flores E & Mullineaux CW (2015) Intercellular diffusion of a fluorescent sucrose analog via the septal junctions in a filamentous cyanobacterium. *mBio* 6: e02109.
- Omairi-Nasser A, Haselkorn R & Austin J 2<sup>nd</sup> (2014) Visualization of channels connecting cells in filamentous nitrogen-fixing cyanobacteria. *FASEB J* 28: 3016-22.
- Omairi-Nasser A, Mariscal V, Austin JR<sup>2nd</sup> & Haselkorn R (2015) Requirement of Fra proteins for communication channels between cells in the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci USA* 112: E4458-64.
- Paz-Yepes J, Herrero A & Flores E (2007) The NtcA-regulated amtB gene is necessary for full methylammonium uptake activity in the cyanobacterium *Synochococcus elongatus*. *J Bacteriol* 189: 7791-98.
- Paz-Yepes J, Merino-Puerto V, Herrero A & Flores E (2008) The amt gene cluster of the heterocyst-forming cyanobacterium *Anabaena* sp strain PCC 7120. *J Bacteriol* 190: 6534-9.





- Pernil R & Schleiff E (2019) Metalloproteins in the Biology of Heterocysts. *Life* (Basel) 9: E32.
- Pernil R, Picossi S, Mariscal V, Herrero A & Flores E (2008) ABC-type amino acid uptake transporters Bgt and N-II of *Anabaena* sp. strain PCC 7120 share an ATPase subunit and are expressed in vegetative cells and heterocysts. *Mol Microbiol* 67: 1067-80.
- 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: 5165-72.
- Pfeffer C, Larsen S, Song J, Dong M, Besenbacher F, Meyer RL, Kjeldsen KU, Schreiber L, Gorby YA, El-Naggar MY, Leung KM, Schramm A, Risgaard-Petersen N & Nielsen LP (2012) Filamentous bacteria transport electrons over centimeter distances. *Nature* 491: 218-21.
- 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: 11582-92.
- Picossi S, Montesinos ML, Pernil R, Lichtlé C, Herrero A & Flores E (2005) ABC-type neutral amino acid permease N-I is required for optimal diazotrophic growth and is repressed in the heterocysts of *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 57: 1582-92.
- Picossi S, Flores E & Herrero A (2013) Diverse roles of the GlcP glucose permease in free-living and symbiotic cyanobacteria. *Plant Signal Behav* 8: e27416.
- 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.
- Pierson BK & Castenholz RW (1974) A phototrophic gliding filamentous bacterium oh hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch Microbiol* 100: 5-24.
- Plominsky ÁM, Delherbe N, Mandakovic D, Riquelme B, González K, Bergman B, Mariscal V & Vásquez M (2015) Intercellular transfer along the trichomes of the invasive terminal heterocyst forming cyanobacterium *Cylindrospermopsis raciborskii* CS-505. *FEMS Microbiol Lett* 362: fnu009.
- Ramírez ME, Hebbar PB, Zhou R, Wolk CP, Curtis SE (2005) *Anabaena* sp. Strain PCC 7120 gene *devH* is required for synthesis of the heterocyst glycolipid layer. *J Bacteriol* 187: 2326-31.
- Ramos-León F (2017) Molecular analysis of the cell-cell joining protein SepJ in *Anabaena* sp. PCC 7120. Ph. D. Thesis, Universidad de Sevilla, Seville, Spain.
- Ramos-León F, Mariscal V, Frías JE, Flores E & Herrero A (2015) Divisome-dependent subcellular localization of cell-cell joining protein SepJ in the filamentous cyanobacteria *Anabaena*. *Mol Microbiol* 96: 566-80.
- Ramos-León F, Mariscal V, Battchikova N, Aro EM & Flores E (2017) Septal protein SepJ from the heterocyst-forming cyanobacterium *Anabaena* forms multimers and interacts with peptidoglycan. *FEBS Open Bio* 7: 1515-26.
- Ramos-León F, Arévalo S, Mariscal V & Flores E (2018) Specific mutations in the permease domain of septal protein SepJ differentially affect functions related to multicellularity in the filamentous cyanobacterium *Anabaena*. *Microb Cell* 5:555-65.
- Rippka R (1972) Photoheterotrophy and chemoheterotrophy among unicellular blue-green algae. *Arch Mikrobiol* 87: 93-8.
- Rippka R, Deruelles J, Waterbury JB, Herdman M & Stainer RY (1979) Generic assignment, strain histories and properties of pure culture of cyanobacteria. *J General Microbiol* 111: 1-61.



- Risser DD & Callahan SM (2008) HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 190: 7645-54.
- Risser DD & Callahan SM (2009) Genetic and cytological evidence that heterocyst patterning is regulated by inhibitor gradients that promote activator decay. *Proc Natl Acad Sci USA* 106: 19884-8.
- Rivers OS, Videau P, Callahan SM (2014) Mutation of *sepJ* reduces the intercellular signal range of a *hetN*-dependent paracrine signal, but not a *patS*-dependent signal, in the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 94: 1260-71.
- Rubio LM & Ludden PW (2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Ann Rev Microbiol* 93-111.
- Rudolf M, Tetik N, Ramos-León F, Flinner N, Ngo G, Stevanovic M, Burnat M, Pernil R, Flores E & Schleiff E (2015) The peptidoglycan-binding protein SjcF1 influences septal junction function and channel formation in the filamentous cyanobacterium *Anabaena*. *mBio* 6: e00376-15.
- Ruiz N (2016) Filling holes in peptidoglycan biogenesis of *Escherichia coli*. *Curr Opin Microbiol* 34: 1-6.
- Ruiz N, Kahne D & Silhavy TJ (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol* 4: 57-66.
- Sambrook J & Russell DW (2001) *Molecular Cloning: A laboratory manual*. Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schirmeister BE, Antonelli A & Bagheri HC (2011) The origin of multicellularity in cyanobacteria. *BMC Evol Biol* 11: 45.
- Schirmeister BE, de Vos JM, Antonelli A & Bagheri HC (2013) Evolution of multicellularity coincided with the increased diversification of cyanobacteria and the Great Oxidation Event. *Proc Natl Acad Sci USA* 110: 1791-6.
- Schirmeister BE, Gugger M & Donoghue PC (2015) Cyanobacteria and the Great Oxidation Event: evidence from genes and fossil. *Palaeontology* 58: 769-85.
- Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, 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 KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M & Kerfeld CA (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci USA* 110: 1053-8.
- Simkovsky R, Effner EE, Iglesias-Sánchez MJ & Golden SS (2016) Mutations in Novel Lipopolysaccharide Biogenesis Genes Confer Resistance to Amoebal Grazing in *Synechococcus elongatus*. *Appl Environ Microbiol* 82: 2738-50.
- Snyder DS, Brahamsha B, Azadi P & Palenik B (2009) Structure of compositionally simple lipopolysaccharide from marine *Synechococcus*. *J Bacteriol* 191: 5499-5509.
- Sperandeo P, Martorana AM & Polissi A (2019) Lipopolysaccharide biosynthesis and transport to the outer membrane of Gram-negative bacteria. *Subcell Biochem* 92: 9-37.
- Springstein BL, Arévalo S, Helbig AO, Herrero A, Stucken K, Flores E & Dagan T (2020) A novel septal protein of multicellular heterocystous cyanobacterial is associated with the divisome. *Mol Microbiol* (doi: 10.1111/mmi/14483).
- Stal LJ & Zher JP (2008) Cyanobacterial nitrogen fixation in the ocean: diversity, regulation and ecology. In Herrero A & Flores E (eds) *The Cyanobacteria: molecular biology, genomic and evolution*. Caister Academic Press, Norfolk, UK, pp. 423-446.
- Staron P, Forchhammer K & Maldener I (2011) Novel ATP-driven pathway of glycolipid export involving TolC protein. *J Biol Chem* 286: 38202-10.



- Stewart WD, Fitzgerald GP & Burris RH (1967) *In situ* studies on nitrogen fixation with the acetylene reduction technique. *Science* 158: 536.
- Strohl WR & Larkin JM (1978) Enumeration, isolation, and characterization of *Beggiatoa* from freshwater sediments. *Appl Environ Microbiol* 36: 755-70.
- Szathmary E & Smith JM (1995) The major evolutionary transitions. *Nature* 374: 227-32.
- Thomas J, Meeks JC, Wolk CP, Shaffer PW, Austin SM (1977) Formation of glutamine form ( $^{15}\text{N}$ )ammonia, ( $^{15}\text{N}$ )dinitrogen, and ( $^{14}\text{C}$ )glutamate by heterocyst isolated from *Anabaena cylindrica*. *J Bacteriol* 129: 1545-55.
- Thompson CL, Vier R, Mikaelyan A, Wienemann T & Brune A (2012) “*Candidatus Arthromitus*” revised: segmented filamentous bacteria in arthropod guts are members of *Lachnospiraceae*. *Environ Microbiol* 14:1454-65.
- Tomitani A, Knoll AH, Cavanaugh CM & Ohno T (2006) The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci USA* 103: 5422-47.
- Tripp J, Hahn A, Koenig P, Flinner N, Bublak D, Brouwer EM, Ertel F, Mirus O, Sinning I, Tews I & Schleiff E (2012) Structure and conservation of the periplasmic targeting factor Tic22 protein from plants and cyanobacteria. *J Biol Chem* 287: 24164-73.
- Turner RD, Waldemar V & Foster SJ (2014) Different walls for rods and balls: the diversity of peptidoglycan. *Mol Microbiol* 91: 862-74.
- Typas A, Banzhaf M, Gross CA & Vollmer W (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* 10:123-36.
- Ungerer JL, Pratte BS & Thiel T (2008) Regulation of fructose transport and its effect on fructose toxicity in *Anabaena* spp. *J Bacteriol* 190: 8115-25.
- Valladares A, Montesinos ML, Herrero A & Flores E (2002) An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol* 43:703-15.
- 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: 1239-49.
- Valladares A, Maldener I, Muro-Pastor AM, Flores E & Herrero A (2007) Heterocyst development and diazotrophic metabolism in terminal respiratory oxidase mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189: 4425-30.
- 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: 6126-33.
- Valladares A, Rodríguez V, Camargo S, Martínez-Noël GM, Herrero A & Luque I (2011) Specific role of the cyanobacterial PipX factor in the heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 193: 1172-82.
- Valladares A, Flores E & Herrero A (2016) The heterocyst differentiation transcriptional regulator HetR of the filamentous cyanobacterium *Anabaena* forms tetramers and can be regulated by phosphorylation. *Mol Microbiol* 99: 808-19.
- Vargas WA, Nishi CN, Giarrocco LE & Salerno GL (2011) Differential roles of alkaline/neutral invertases in *Nostoc* sp. PCC 7120: Inv-B isoform is essential for diazotrophic growth. *Planta* 233: 153-62.



- Vázquez-Bermúdez MF, Flores E & Herrero A (2002a) Analysis of binding sites for the nitrogen-control transcription factor NtcA in the promoters of *Synechococcus* nitrogen-regulated genes. *Biochim Biophys Acta* 1575: 95-8.
- Vázquez-Bermúdez MF, Herrero A, Flores E. (2002b) 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. *FEBS Lett* 512: 71-74.
- Vázquez-Bermúdez MAF, Paz-Yepes J, Herrero A, Flores E. (2002c) The NtcA-activated *amt1* gene encodes a permease required for uptake of low concentrations of ammonium in the cyanobacterium *Synechococcus* sp. PCC 7942. *Microbiology* 148: 861-869.
- Vázquez-Bermúdez MF, Flores E & Herrero A (2003) Carbon supply and 2-oxoglutarate effects on expression of nitrate reductase and nitrogen-regulated genes in *Synechococcus* sp. strain PCC 7942. *FEMS Microbiol Lett* 221: 155-9.
- Vetting MW, Hegde SS, Fajardo JE, Fiser A, Roderick SL, Takiff HE & Blanchard JS (2006) Pentapeptide repeat proteins. *Biochemistry* 45: 1-10.
- Vetting MW, Hegde SS, Hazleton KZ & Blanchard JS (2007) Structural characterization of the fusion of two pentapeptide repeat proteins, Np275 and Np276, from *Nostoc punctiforme*: resurrection of an ancestral protein. *Protein Sci* 16: 755-60.
- Villareal TA (1990) Laboratory culture and preliminary characterization of the nitrogen-fixing Rhizosolenia-Richelia symbiosis. *Marine ecology* 11: 117-132.
- Vioque A (1997) The RNase P RNA from cyanobacteria: short tandemly repeated repetitive (STRR) sequences are present within the RNase P RNA gene in heterocystforming cyanobacteria. *Nucleic Acids Res* 25: 3471-3477.
- Vollmer W, Joris B, Charlier P & Foster S (2008) Bacterial peptidoglycan (murein) hidrolases. *FEMS Microbiol Rev* 32: 259-86.
- Walsby AE (2007) Cyanobacterial heterocysts: terminal pores proposed as sites of gas exchange. *Trends Microbiol* 15: 140-9.
- Wang Q, Li H & Post AF (2000) Nitrate assimilation genes of the marine diazotrophic, filamentous cyanobacterium *Trichodesmium* sp. Strain WH9601. *J Bacteriol* 182: 1764-67.
- 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: 8489-93.
- Wang Y, Lechno-Yossef S, Gong Y, Fan Q, Wolk CP, Xu X (2007) Predicted glycosyl transferase genes located outside the HEP island are required for formation of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189: 5372-8.
- Weckesser J, Katz A, Drews G, Mayer H & Fromme I (1974) Lipopolysaccharide containing L-acofriose in the filamentous blue-green alga *Anabaena variabilis*. *J Bacteriol* 120: 672-8.
- Wei TF, Ramasubramanian TS & Golden JW (1994) *Anabaena* sp. strain PCC 7120 ntcA gene required for growth on nitrate and heterocyst development. *J Bacteriol* 176: 4473-82.
- Weiss GL, Kieninger AK, Maldener I, Forchhammer K & Pilhofer M (2019) Structure and function of a bacterial gap junction analog. *Cell* 178: 274-84.
- West SA, Griffin AS, Gardner A & Diggle SP (2006) Social evolution theory for microorganisms. *Nat Rev Microbiol* 4: 597-607.
- Wilson MC, Mori T, Rückert C, Uria AR, Helf MJ, Takada K, Gernert C, Steffens UA, Heycke N, Schmitt S, Rinke C, Helfrich EJ, Brachmann AO, Gurgui C, Wakimoto T, Kracht M, Crüsemann M, Hentschel



- U, Abe I, Matsunaga S, Kalinowski J, Takeyama H & Piel J (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* 506: 58-62.
- Wildon DC & Mercer FV (1963) The ultrastructure of the vegetative cell of blue-green algae. *Aust J Biol Sci* 16: 585-96.
- Wilk L, Strauss M, Rudolf M, Nicolaisen K, Flores E, Kühlbrandt W & Schleiff E (2011) Outer membrane continuity and septosome formation between vegetative cells in the filaments of *Anabaena* sp. PCC 7120. *Cell microbiol* 13: 1744-54.
- Wolk CP, Thomas J, Shaffer PW, Austin SM & Galonsky A (1976) Pathway of nitrogen metabolism after fixation of <sup>13</sup>N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. *J Biol Chem* 251: 5027-34.
- Wolk CP, Ernst A & Elhai J (1994) Heterocyst metabolism and development. In Bryant DA (ed) *The molecular biology of cyanobacteria*. Kluwer Academic Publisher, Holland, 769-823.
- Wu SS, Wu J, Cheng YL & Kaiser D (1998) The *pilH* gene encodes an ABC transporter homologue required for type IV pilus biogenesis and social gliding motility in *Myxococcus xanthus*. *Mol Microbiol* 29:1249-61.
- Xu X, Elhai J & Wolk CP (2008) Transcriptional and developmental responses by *Anabaena* to deprivation of fixed nitrogen. In Herrero A & Flores E (eds) *The Cyanobacteria: molecular biology, genomic and evolution*. Caister Academic Press, Norfolk, UK, pp. 383-422.
- Xu X, Khudyakov I & Wolk CP (1997) Lipopolysaccharide dependence of cyanophage sensitivity and aerobic nitrogen fixation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 179: 2884-91.
- Yoon HS & Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. *Science* 282: 935-8.
- Yoon HS & Golden JW (2001) PatS and products of nitrogen fixation control heterocyst pattern. *J Bacteriol* 183: 2605-13.
- Zhang L, Zhou F, Wang S & Xu X (2017) Processing of PatS, a morphogen precursor, in cell extracts of *Anabaena* sp. PCC 7170. *FEBS Lett* 591: 751-59.
- Zhang LC, Chen YF, Chen WL & Zhang CC (2008) Existence of periplasmic barriers preventing green fluorescent protein diffusion from cell to cell in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 70: 814-23.
- Zhang LC, Risoul V, Latifi A, Christie JM & Zhang CC (2013) Exploring the size limit of protein diffusion through the periplasm in cyanobacterium *Anabaena* sp. PCC 7120 using the 13 kDa iLOV fluorescent protein. *Res Microbiol* 164: 710-7.
- Zhang R, Ni S & Kennedy MA (2019) Type I beta turns make a new twist in pentapeptide repeat proteins: Crystal structure of Alr5209 from *Nostoc* sp. PCC 7120 determined at 1.7 angström resolution. *JBS* doi: 10.1016/j.jysbx.2019.100010.
- Zheng Z, Omairi-Nasser A, Li X, Dong C, Lin Y, Haselkorn R & Zhao J (2017) An amidase is required for proper intercellular communication in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Proc Natl Sci USA* 114: E1405-12.
- Zhu J, Jäger K, Black T, Zarka K, Koksharova O & Wolk CP (2001) Hcwa, an autolysin, is required for heterocyst maturation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 183: 6841-51.

