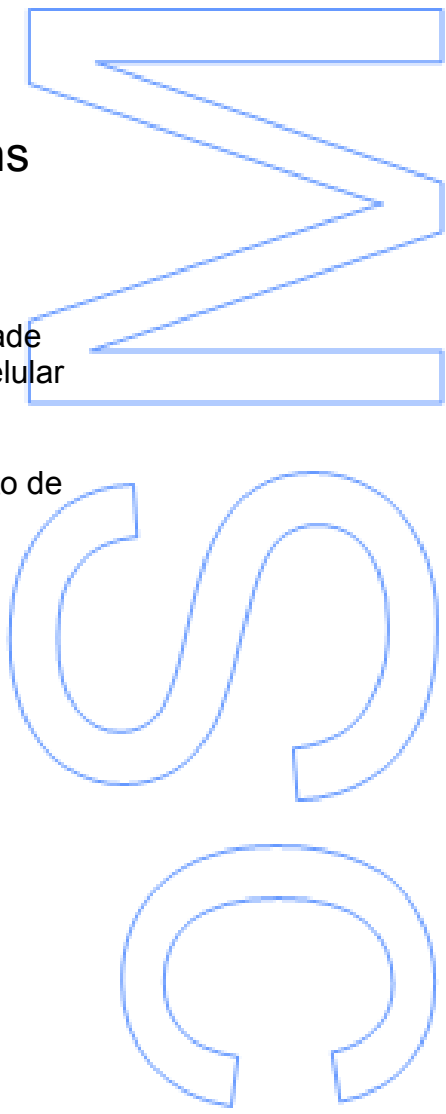
Characterization of the Exoproteome of Two Morphologically Distinct Cyanobacteria

Nuno Manuel de Resende Rego Macedo Martins
Mestrado de Biologia Celular e Molecular
Departamento de Biologia
2013
Orientador
Professora Doutora Paula Tamagnini, Professora Associada da Faculdade de Ciências do Porto; Líder de grupo em IBMC- Instituto de Biologia Celular e Molecular, Universidade do Porto
Co orientador
Doutor Paulo Oliveira, Bolseiro de Pós-doutoramento em IBMC- Instituto de Biologia Celular e Molecular, Universidade do Porto



U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Dissertação de candidatura ao grau de Mestre em Biologia Celular e Molecular submetida à Faculdade de Ciências da Universidade do Porto.

O presente trabalho foi desenvolvido sob a orientação científica da Professora Doutora Paula Tamagnini e co-orientação do Doutor Paulo Oliveira e foi realizado no Instituto de Biologia Molecular e celular da Universidade do Porto.

Dissertation for applying to a Master's Degree in Molecular and Cell Biology, submitted to the Faculty of Sciences of the University of Porto.

The present work was developed under the scientific supervision of Professor Paula Tamagnini and co-supervision of Doctor Paulo Oliveira and was done at the Institute Molecular and Cellular Biology of University of Porto.

Acknowledgements

I would like to express my deepest gratitude to my mentor Dr. Paulo Oliveira. He was a true inspiration and example in lab and taught me so much. He was truly like a Mr. Miyagi to me.

It is with immense gratitude that I acknowledge the support of my supervisor Professor Paula Tamagnini in the writing of the thesis and for opening my eyes for the wonders of the cyanobacterial world, and of course for giving me the opportunity to work in this amazing lab.

I feel compelled to acknowledge Professor Phillip Wright for the opportunity to attend to his department in the University of Sheffield and to use all the mass spectrometry equipment. In the same note I want to thank Dr. Narciso Couto and Dr. Caroline Evans for the immense help with mass spec related issues always with great dedication. And of course I want to thank Dr. Sara Pereira for the help with my protein samples. And I will not forget all the help the Sheffield PhD students gave me, especially Joseph Longworth and David Russo with lab advices; Andrew Landels with all the bioinformatic and of course a special mention to Ben Strutton without whom I would have lived in the streets for a week.

I will never forget the daily brainstorming sessions with Pedro Ferreira which always gave me great ideas which usually allowed me to overcome several setbacks. I also would like to thank everybody in our group for always making me feel welcome and giving me advice everytime it was needed.

Last but not least, I want to give a special thank for my family, they were crucial in my life and in my education, especially my mom and dad who gave me the amazing privilege of education and taught me almost everything I know (maybe not in cyanobacteria per se).

Resumo

Neste trabalho foram investigadas as proteínas extracelulares (exoproteoma) numa cianobactéria unicelular, *Synechocystis* sp. PCC 6803 e numa filamentosa, *Anabaena* sp. PCC 7120. A análise foi feita em meios com diferentes fontes de azoto; nitrato e amónia em ambas as espécies e sem nenhuma fonte de azoto combinado apenas em *Anabaena* sp. PCC 7120, uma vez que *Synechocystis* sp. PCC 6803 não é capaz de fixar azoto atmosférico.

A identificação de proteínas extracelulares foi feita utilizando uma análise de espectrometria de massa ESI-ion trap. Foram identificadas 117 proteínas extracelulares em *Anabaena* sp. PCC 7120 e 30 em *Synechocystis* sp. PCC 6803. Embora, para uma grande parte das proteínas identificadas não seja conhecida a função, foi notório a presença de proteínas relacionadas com o processamento e aquisição de nutrientes do meio. Foram também identificadas proteínas envolvidas na defesa de stresse oxidativo em culturas de *Anabaena* sp. PCC 7120. Foi, ainda, analisada a atividade destas no meio de cultura.

Para avaliar a contaminação do meio extracelular com proteínas intracelulares, foram gerados mutantes com expressão de GFP direccionada para o periplasma para, posteriormente, verificar a presença da GFP no exoproteoma. Foram também gerados, para ambas as espécies, mutantes sem o poro da membrana extracelular (TolC) da via de secreção do tipo I, posteriormente o padrão de proteínas destes mutantes será comparado com o da respetiva estirpe selvagem.

Uma análise mais detalhada das proteínas identificadas neste trabalho, conjuntamente com uma caracterização das mesmas vai proporcionar um melhor conhecimento das proteínas secretadas para o meio extracelular, por cianobactérias, e quais as suas funções na sobrevivência destas bactérias.

Palavras-chave: Proteínas extracelulares; *Synechocystis*; *Anabaena*; Secreção de proteínas

Abstract

In this work it was investigated the extracellular proteins (exoproteome) in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803 and a filamentous one, *Anabaena* sp. PCC 7120. The analysis was done in the presence of different nitrogen sources; nitrate and ammonium in both species and without combined nitrogen only in *Anabaena* sp. PCC 7120 since *Synechocystis* sp. PCC 6803 is incapable of nitrogen fixation.

The identification of the extracellular proteins was done with an ESI-ion trap mass spectrometry analysis. A total of 117 extracellular proteins were identified for *Anabaena* sp. PCC 7120 and 30 for *Synechocystis* sp. PCC 6803. Although most of the proteins identified had unknown functions it was notorious the presence of proteins related to the processing and acquisition of nutrients present in the medium. It was also identified proteins involved in the defense of oxidative stress in the *Anabaena* sp. PCC 7120 cultures. These proteins were further analyzed to confirm if indeed these proteins have activity in the extracellular milieu.

To evaluate the contamination of the extracellular milieu with intracellular proteins, mutants expressing GFP exported to the periplasm were generated to ultimately verify the presence of the GFP in the exoproteome. Deletion mutants for the outer membrane pore (TolC) of the secretion pathway Type I were also generated for both species. In the future these mutants will be used to compare their extracellular proteins pattern against their respective wild type strain.

The follow-up of the proteins identified in this work together with the characterization of the proteins identified will provide a better understanding of the mechanisms of survival of these bacteria in different environments.

Key-Words: Extracellular proteins; *Synechocystis*; *Anabaena*; Protein Secretion

Table of contents

List of tables and illustrations

General abbreviations

ABC	ATP Binding Cassette
APS	Ammonium persulphate
ATP	Adenosine triphosphate
DAB	3,3'-Diaminobenzidine
DNA	Desoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GFP	Green Fluorescent Protein
GSP	General Secretory Pathway
HSP	Horseradish peroxidase
IM	Inner membrane
LB	Lysogeny broth
LC	Liquid Chromatography
MFP	Membrane fusion protein
MS	Mass Spectrometry
NBT	Nitroblue tetrazolium
OM	Outer membrane
OMP	Outer membrane protein
PAGE	Polyacrylamide gel electrophoresis
PCC	Pasteur Culture Collection
PCR	Polymerase chain reaction
pSK+	Plasmid Bluescript SK+
PTM	Post Translation Modification
RNA	Ribonucleic acid
RND	Resistance Nodulation-cell Division
RPM	Revolutions per minute
SDS	Sodium dodecyl sulphate
SOB	Super Optimal Broth
SOD	Superoxide Dismutase
SP	Signal Peptide
TaT	Twin-arginine Transportation
TEMED	Tetramethylethylenediamine
Tris	2-Amino-2-hydroxymethylpropane

Introduction

Cyanobacteria are an ancient (over 3 billion years old) and diverse group of prokaryotes. Their ability to perform oxygenic photosynthesis together with their ability to adapt to environmental conditions, contribute to their worldwide distribution. They are extensively studied because of their variety and for their ability to convert solar power into energy and atmospheric carbon dioxide into organic molecules [1].

Within the many species of cyanobacteria, the unicellular *Synechocystis* sp. PCC 6803 has become a model of study. This strain was isolated from a fresh water lake and deposited in Pasteur culture collection in 1968 [2]. Besides being one of the most well studied cyanobacteria, it has the capacity of natural transformation with exogenous DNA, including its capacity of integration that DNA into the genome by double recombination [3]. Another important mark for its wide use was its genome sequencing in 1996, establishing itself as the first photosynthetic organism with its genome completely sequenced [4, 5]. Although it has several advantages for scientific research it lacks certain characteristics present in other cyanobacteria like multicellularity, nitrogen fixation capacity or cellular differentiation.

Another well studied cyanobacterium, the filamentous *Anabaena* sp. PCC 7120 (also known as *Nostoc* sp. PCC 7120) has also been used because of its ability to fix atmospheric nitrogen in differentiated cells, the heterocysts. Its genome has also been fully sequenced, and it became available in 2001 [6, 7].

Although both organisms are cyanobacteria, they are different. Besides all the characteristics already mentioned, one of the most important differences between them, relevant for this work, is the morphology of the cell wall. While *Synechocystis* sp. PCC 6803 is unicellular with a cell wall composed of a plasma membrane and an outer membrane, both limiting the periplasmic space, and a thin peptidoglycan layer between both membranes, *Anabaena* sp. PCC 7120 is a multicellular strain that has an individual plasma membrane for each cell and a common outer membrane for the whole filament with a dynamic continuous periplasm [8, 9]. Another very important distinction is the presence of a Surface-layer (S-layer) in *Synechocystis* sp. PCC 6803 and its absence in *Anabaena* sp. PCC 7120 [10]. The S-layer consists in repetition of structural proteins throughout the cell surface anchored to the outside of the outer membrane and it has several functions such as cell stabilization and protection against environmental factors [11]. Although *Anabaena* sp. PCC 7120 lacks the S-layer, it has an extra cellular matrix, which

consists of a dynamic mixture of polysaccharides, proteins, cell remnants and secondary metabolites [12, 13].

Protein secretion is an essential mechanism for bacterial survival; it is involved in several aspects of bacterial life such as motility, nutrient acquisition and pathogen secretion. Although there are a lot of proteomic studies for both cyanobacteria mentioned above, including total proteomics [14], periplasm proteomics [15], cell and outer membrane proteomics [16-19] , and even thylakoid membrane proteomics [20], the content of proteins accumulated in the extracellular milieu (Exoproteome) is yet to be fully identified. Only a few studies have been carried out in *Synechocystis* sp. PCC 6803 trying to explore its exoproteome [21-23], but there are still many proteins left to be identified. In *Anabaena* sp. PCC 7120 the exoproteome is still to be identified, although it has been partially characterized in a similar filamentous cyanobacterium, *Nostoc commune* [12, 24, 25]. The significant differences between both species are likely to influence which proteins belong to their respective exoproteomes. Because *Anabaena* sp. PCC 7120 is a filamentous strain it is possible to have cell-cell interaction with proteins travelling through the periplasm without being released to the growth medium [9], which can affect the composition of the exoproteome. In addition, it has been documented that the S-layer proteins present in *Synechocystis* sp. PCC 6803 are often released to the growth media, thus also influencing the composition of the exoproteome [26].

It is important to establish the clear difference between the exoproteome, the sum of all proteins present in the extracellular space, and the secretome, all the proteins transported from the interior of the cell out into the medium. Thus, it is possible for a protein to be part of the exoproteome but not to be part of the secretome, in the case when it is accumulated in the growth medium by some sort of leakage of the periplasm or cell lysis instead of being actively transported [27]. It is also important to make the distinction between the terms export, which is when a protein is transported through a membrane whether it is the plasma membrane (thus arriving the periplasm) or the outer membrane (reaching the extracellular space), and the term secretion, which is the transport of a protein all the way through the cell wall and into the growth medium [27].

Protein Secretion

Gram-negative bacteria face an extra obstacle when it comes to protein secretion, the outer membrane. While the proteins are usually exported through the inner membrane quite easily [28], they face two new obstacles when they are secreted through the outer membrane. The periplasm does not have ATP or other sources of energy so the secretion mechanisms need to

be self-energized or harness energy from the inner membrane [28]. In addition, proteins acquire tertiary structure in the periplasm, thus making the secretion even more difficult [28].

In Gram-negative bacteria proteins can be secreted in one-step, crossing the inner and outer membranes inside the same protein complex, or it can be in two-steps. In the latter secretion pathway, proteins are first exported to the periplasm through the general secretory pathway (GSP) or Sec pathway and then secreted to the extracellular space [28-32].

There are six branches of protein secretion in Gram-negative bacteria [28]: two of them consist in the secretion of the proteins after the first export into the periplasm, these are considered terminal branches of the GSP in a two-steps secretion process (Sec-dependent) [28-30]. The other four are named Type I to IV. Types I and III are considered full Sec-independent while types II and IV can be Sec-dependent or Sec-independent [28-30].

Protein export to the periplasm

The first step of the Sec-dependent pathways is the export of the proteins to the periplasm, usually done by the Sec system. The Sec system involves a first step of recognition of an N-terminal signal peptide of the protein by a ribonucleotide complex comprising of the Ffh protein (not yet identified in *Anabaena* sp. PCC 7120, slr1531 in *Synechocystis* sp. PCC 6803) and the RNA ffs (identified in both species) and by the chaperone SecB protein (not yet identified in either of the cyanobacterial strains used). The protein is then sent to the translocation unit consisting of SecA (alr4851 in *Anabaena* sp. PCC 7120, sll0616 in *Synechocystis* sp. PCC 6803), SecY (all4197, sll1814), SecE (all4197, ssl3335), SecG (not yet identified in *Anabaena* sp. PCC 7120, ssr3307) [4, 6, 33]. After this process proteins may stay in the periplasm or be further secreted to the medium by the secretion terminal branches of the Sec pathway [28, 29]. Although the Sec pathway is the most important pathway involved in secretion, there is also another cytoplasm to periplasm export pathway relevant in cyanobacteria, the twin-arginine pathway (TaT) [9, 34, 35]. This pathway is involved in the export of proteins in a similar way of the Sec pathway, but it recognizes a different N-terminal signal peptide which has two arginine's together [9, 34].

Autotransporters and single accessory pathway

The autotransporter pathway is a terminal branch of the Sec pathway. These proteins do not require accessory factors to be secreted since they have all the domains necessary to cross the outer membrane in their sequence. To trespass the outer membrane these proteins have C-terminal β-barrel domain that insert themselves in the outer membrane forming a pore structure similar to porins [28, 36]. The rest of the sequence is then secreted inside the pore and can

become anchored in the extracellular side of the outer membrane or be fully released [36]. The sequence of the protein which is secreted inside the β-barrel pore does not seem to be important, since in other works it has been substituted for different proteins which were secreted [37, 38].

Some proteins do not have the β-barrel domain necessary for secretion; they instead use single accessory factors to be secreted. These single accessory factors work in a way similar to the autotransporter, forming a β-barrel pore in the outer membrane and the soon to be secreted protein is translocated inside the pore into the medium [28]. Although both systems are very similar in function the β-barrel of the autotransporters and of the single accessory factors have very low homology [28]. Most of these single accessory factors seem to have homology with protein-translocating porin found in chloroplasts [39]. In *Synechocystis* sp. PCC 6803 it was identified and confirmed homologies of these protein-translocating porins, the synToc75 (slr1227) were homologous on other Gram-negative species are involved in the secretion of Hemolysin proteins [39]. In *Anabaena* sp. PCC 7120 there is a homologue of this transporter, the Omp85 (alr2268) [40].

Type I Secretion

The type I secretion system is a one-step, sec-independent pathway. In this secretion system the substrate protein are recognized, usually by their C-terminal signal peptide (~60 aminoacids) by an inner membrane ABC-Transporter; then they travel inside a membrane fusion protein (MFP) without going to the periplasmic space, finally they are secreted by an outer membrane protein (OMP), a protein of the TolC family [35, 41, 42]. Although each of the three parts of the complex are necessary for the secretion to occur it is the ABC-transporter that makes the selection of which proteins are going to be secreted [35]. In the same bacteria there are usually different ABC-transporters and MFP's; on the other side there is usually only one OMP in each species. Usually these different components can assembly in different combinations depending on the substrate to be secreted [35, 42]. While the ABC-transporter and the MFP are already together upon the substrate recognition, the OMP only assembles after the recognition occurs [42]. The OMP can be assembled in a different complex comprising of a MFP with an RND exporter instead of an ABC-transporter. This RND is a drug export inner membrane protein involved in processes of detoxification. This means that the OMP does not mediate the secretion of only proteins but other molecules too [41].

The ABC-transporters in the type I secretion work both in the recognition of the protein substrate and to energize the secretion process [43]. The structure of the ABC-transporters comprise two transmembrane domains and two nucleotide-binding domains which can be arranged in different

combinations [42]. The nucleotide-binding domains are more highly conserved than the transmembrane domains. This shows that the specificity of the ABC-transporter-substrate is given by the transmembrane domains while the nucleotide-binding domains have ATPase activity necessary to energize the secretion process [41]. Because the ABC-transporters are usually responsible for the recognition of one specific substrate, they are usually coded in gene clusters with their MFP and their substrate [43]. There are exceptions where the ABC-transporters and their MFP are in the same gene cluster but there is no secreted protein nearby. In these cases usually the ABC-transporters are not specific and can secrete more than one type of protein [43]. The ABC-transporters usually recognize Glycine-rich repeats (GGXGXD) present in the last 60 aminoacids, usually repeated 4-36 times. They are often present in toxins like the RTX, Hemolysin and lipases [43].

The MFP complex role in type I secretion is to make a bridge between the ABC-transporter and the OMP. It has an N-terminal fragment in the cytoplasm, a transmembrane segment and a large periplasmic domain which binds to the OMP after the substrate is recognized by the ABC-transporter [41]. In some cases the MFP N-terminal fragment is known to be essential for substrate recognition for Type I secretion [42]. Although at the moment there is no information about other functions of the MFP complex, it is postulated that it might have a more complex role; this evidence is supported by the discovery of MFP complex in Gram-positive bacteria, which do not have outer membrane [35].

The OMP function is act as a pore to secrete the substrate protein transported from the MFP. The best studied OMP is the TolC from *Escherichia coli* which is the third complement of several Type I secretion trios together with ABC-transporters and MFP's [41]. Because the TolC is ubiquitous in all Gram-negative bacteria and is important both in protein secretion and drug efflux, it has been well characterized and its structure is well known [44]. The TolC works by alternating from a closed position to an open position when it is triggered by its substrate. To do so the coiled-coils twist like an iris opening, changing its hole aperture from 3.5 Å to 16 to 20 Å [45]. This aperture is then wide enough for proteins in their secondary structure [41]. There is also the unconfirmed hypothesis that the TolC can act as a single accessory factor to secrete proteins of the periplasm independent of the ABC-Transporter-MFP complex [41].

In *Anabaena* sp. PCC 7120 and *Synechocystis* sp. PCC 6803 there is only one TolC homologue that acts as OMP for protein secretion, alr2887 and slr1270 [5, 6]. However at least in *Synechocystis* sp. PCC 6803 there are four confirmed substrates for Type I secretion, this is indication that like in most Gram-negative bacteria there is only one OMP that assembles in different ABC-transporters-MFP complexes [44].

Type II Secretion

Type II secretion also known as main terminal branch of the Sec pathway, is mainly a Sec-dependent, although not completely dependent, secretion pathway. The type II secretion system is extremely similar to the Type IV pilus biogenesis (also known as type 4 pilus biogenesis); they even share most of the proteins in some bacteria [46, 47]. Type IV pilus biogenesis pathway is also important in the motility and adhesion of bacteria, and in the uptake of DNA and the natural transformation [47, 48]. However while many proteins are identical in the type II secretion and in the type IV pilus biogenesis there are some differences described as shown on Figure 1 [28].

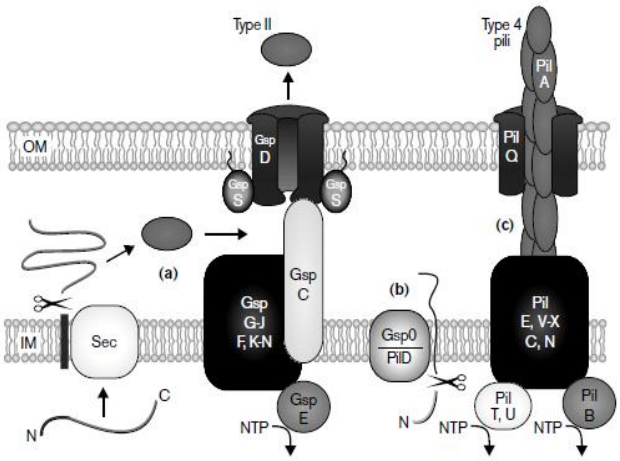


Fig 1- Models for type II secretion and type 4 pilus biogenesis. Components of the type II secreton are indicated using the Gsp nomenclature, and type 4 pilus proteins are labeled according to the Pil system. Similar shading and location indicates homologous components. (a) Type II substrates cross the IM via the Sec system. GspC may transmit energy from the IM to the OM complex. (b) GspO cleaves their amino-terminal leader sequence on the cytoplasmic face of the IM. (c) Type 4 pilus biogenesis requires the OM secretin PilQ, but no energy is necessary for secretion, which is presumably driven by the force of the PilA assembly towards the outer membrane. Adapted from [28].

Recently it has been described that both systems can work together. In that model the pilin polypeptide (PilA) is constantly being assembled in the Pil complex of the inner membrane (PilE, PilC, PilN, PilV-X) and grows from the base to the outer membrane and trespasses the outer membrane inside the secretin channel (PilQ) in a rotation motion, which will be responsible for the twitching motility of the cell. At the same time proteins present in the periplasm, exported in the Sec pathway, attach to the PilA subunits and are driven to the medium together with the PilA complex rotation motion. All of these go through the secretin pore in a way similar to the Archimedes' screw model (Figure 2). That way the protein present in the periplasm is secreted in the Type II secretion system in a Sec-dependent way, while the PilA is secreted in a Sec-independent way using the type IV pilus biogenesis [46].

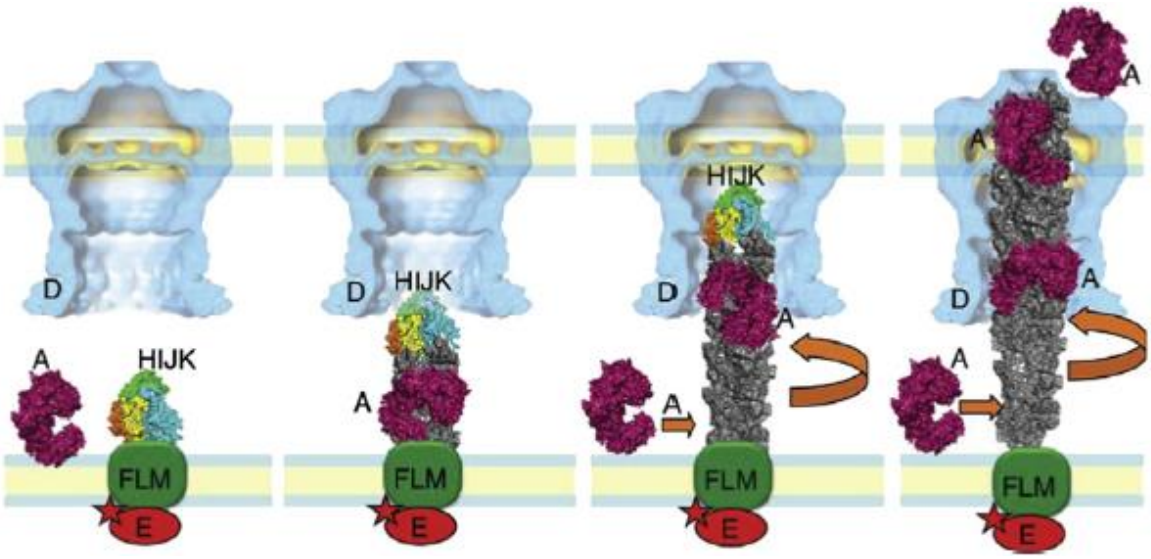


Fig 2- Models for the Archimedes' screw. The minor pseudopilin complex HIJK binds to the inner membrane complex to initiate the PilA assembly. The PilA (grey) starts to form the complex in a rotation upwards motion into the outer membrane pore. While the PilA is being assembled a periplasmic protein exported by the Sec pathway binds to the PilA and is secreted together with the PilA. Adapted from [46].

In *Synechocystis* sp. PCC 6803 it is not known the details of the Type II secretion system, whether it occurs like the Archimedes' screw model or in parallel of the type IV pilus biogenesis using only some of the same proteins. However most of the Pil proteins of the type IV pilus biogenesis have been identified in *Synechocystis* sp. PCC 6803 with the most important being the sll1694 (PilA, major pillin protein); slr1227 (PilQ; Secretin pore) and the slr1120 (PilD, leader peptidase/N-metilase of the PilA) [49]. However, a mutant secreting lichenase in the Type II pathway was already achieved in *Synechocystis* sp. PCC 6803 [22]. To do so it was made a fusion of the active lichenase protein with the N-terminal leader peptide of the PilA *Synechocystis* sp. PCC 6803 gene (sll1694) [23]. More important, that work discovered that the particular sequence of the leader peptide was not essential for the secretion but instead the relevant characteristic was the positive global charge of the N-terminal sequence [23]. This indicates that there is type II secretion in *Synechocystis* sp. PCC 6803 because not only the PilA (part of the Type IV pilus biogenesis and motility) is recognized for secretion. However there is no information that can elucidate whatever this Type II secretion occurs together with the Type IV pilus biogenesis and motility or in a parallel way.

In filamentous *Anabaena* sp. PCC 7120 there is no pillus therefor there is no Type IV pilus biogenesis. However there are homologues of the Pil main structures such as the inner membrane complex ant the outer membrane porin in the genome of *Anabaena* sp. PCC 7120 with the exception of the PilA which seems to be absent [6]. This highly suggests that although

there is no pilus formation the type II secretion still exists in *Anabaena* sp. PCC 7120 (as described in Figure 1 (a)).

Type III and Type IV secretion

Type III secretion consist in the translocation of virulence factors from the Gram-negative bacteria to the eukaryotic host. This virulence factors can be DNA or secreted proteins. The secretion is mediated by a needle-like apparatus connected to the bacteria outer membrane and inner membrane by a transperiplasmic ring and projects itself to the host membrane. The proteins and DNA then travel inside the hollow needle into the host cytoplasm [28, 50]. The Type IV secretion is a mechanism analogue to the infection with plasmid DNA and virulence factors found in *Agrobacterium tumefaciens* and it is also often used in conjugation techniques mediated by *E. coli* [51]. The type-III is only Sec-dependent while the Type IV secretion can be either Sec-independent or Sec-dependent [50, 52].

Both secretion systems occur mainly in eukaryotic pathogens, thus they are absent in cyanobacteria [50, 52].

State of the art in cyanobacterial exoproteomes

Although most studies involving extracellular proteins have the intention of studying secretion, it is very hard to identify whether a protein present in the extracellular milieu came from active secretion or periplasmic leakage or even cell lysis. Therefore many studies identify the whole exoproteome. That setback is clear in a study where the objective was to identify the exoproteome of *Synechocystis salina* [13], in which 9 of the 14 proteins identified (including the ones with highest confidence) were components of the phycobilisome, which is well established to be anchored to the thylakoid membrane [53]. Although the authors did several controls to try to prevent leakage of internal proteins to the media the results show how difficult it is to fully prevent contamination with intracellular content [13]. It is expected the same could happen in *Synechocystis* sp. PCC 6803 or in *Anabaena* sp. PCC 7120.

At the moment there are no studies about *Anabaena* sp. PCC 7120 exoproteome and only one study was done in *Synechocystis* sp. PCC 6803 [21]. This study only analyzed a fraction of the exoproteome with N-terminal sequencing techniques which leaves most of the extracellular proteins yet to be revealed. In that particular study only 7 proteins were identified from which only two had known function; the pilin polypeptide PilA1 (Sll1694) and the Slr0924 which corresponds to a plant chloroplast protein Tic22 that is involved in protein import into thylakoid membranes [54]. Nevertheless, there are at least four proteins like the Hemolysin sll1951

confirmed to be secreted by the Type I secretion method that were not identified in that study [41, 55, 56]. This indicates that the exoproteome of both species are yet to be fully unveiled.

Objectives

The first goal of this work was to identify the exoproteomes of a filamentous and a unicellular cyanobacteria, *Anabaena* sp. PCC 7120 and *Synechocystis* sp. PCC 6803 grown with different nitrogen sources. *Anabaena* sp. PCC 7120 was grown with atmospheric nitrogen, nitrate and ammonia whereas *Synechocystis* sp. PCC 6803 was grown only with nitrate and ammonia, since this strain cannot fix atmospheric nitrogen.

To assess the possibility of periplasmic leakage, mutants with periplasmic GFP were generated, and the presence of the GFP in the media will be evaluated in the future.

Once the identification of the exoproteome, which contained the secretome, was performed it is important to understand which set of proteins constitute the secretome. To identify which proteins are secreted by type I secretion mechanism, deletion mutants for the outer membrane pore (TolC) were generated for both cyanobacteria.

Material and Methods

Organisms and Growth Conditions

Synechocystis sp. PCC 6803 and *Anabaena* sp. PCC 7120, as well as their respective mutants were grown in liquid medium at 25^{0}C, with constant aeration and a regime of 16h light (40 µmol photons m^{-2} s^{-1}) /8h dark. Cyanobacteria were also grown in agar plates, at the same temperature and light conditions. They were grown in variations of $BG11_0$ [57] with different nitrogen sources. For both species it was used BG11 ($BG11_0$+1.5g/L $NaNO_3$) and $BG11_0$+NH_4Cl and $BG11_0$ only for the *Anabaena* sp. PCC 7120; due to the incapacity of *Synechocystis* sp. PCC6803 to fix atmospheric nitrogen. All the cyanobacterial mediums were supplemented with 10mM HEPES buffer.

Escherichia coli was grown in LB at 37^{0}C, either in an orbital shaker at 180RPM for liquid cultures or on plates.

To make solid medium 1.5% Bacteriological Agar or 1.5% Difco® Agar Noble was added to LB medium or to BG11, respectively.

Growth Measurements

The growth of the cyanobacteria culture was monitored by measuring the Chlorophyll *a* content. For that purpose, a sample of the culture was centrifuged, and the pellet resuspended in 100% methanol. The suspension was then left at 4 ^{0}C overnight, before being centrifuged and the supernatant's absorbance measured at 665nm in a Shimadzu® UVmini-1240 spectrophotometer. An extinction coefficient of 13.43 was used to calculate Chlorophyll *a* concentration, as described in [58].

The growth of *Synechocystis* sp. PCC 6803 was also monitored by optical density (OD) at 750nm, in a Shimadzu® UVmini-1240 spectrophotometer.

Preparing *Escherichia coli* XL1 blue competent cells

To prepare competent cells it was used a variation of the Hanahan protocol [59]. A seed stock of *Escherichia coli* XL1 blue previously prepared competent cells were used. That stock was grown in 1000mL Erlenmeyer flasks containing 250mL of SOB media at 37^{0}C to an OD600nm of 0.5. The cells were then centrifuged at 3875*g* and 4ºC for 10 minutes. The supernatant was then discarded and the pellet was resuspended in 80mL of ice cold CCMB80 buffer [60] and incubated on ice for 20 minutes. Cells were centrifuged again in the same conditions as described above. Once again the supernatant was discarded and the pellet resuspended in

10mL of ice cold CCMB80 buffer. The OD600nm of the cells suspension was measured and adjusted to a final OD of 2. The cells were then transferred to Eppendorf tubes as 100µL aliquots and stored at -80^{0}C.

Escherichia coli transformation

E. coli was transformed using a method based on a procedure described by Hanahan [59]. To make the transformation, 10µL of a given ligation was added to a 100µL aliquot of *E. coli* cells previously thawed on ice and the mixture was incubated there for 30 minutes. Then the cells were transferred to a 42^{0}C water-bath for 30 seconds, after which they were quickly transferred onto ice for 2 minutes. Once the incubation on ice was over, warm LB medium was added to the cells, which were left to recover in an orbital shaker at 37^{0}C for 2h. Finally, cells were plated in LB plates supplemented with the appropriate antibiotic.

Synechocystis sp. PCC 6803 transformation

Synechocystis was transformed using a method based on a procedure described by Williams [61]. *Synechocystis* sp. PCC 6803 cells were grown in BG11 up to an OD750nm of 0.55. Then 50mL of culture were harvested and centrifuged at 3875*g* for 10 minutes. The pellet was resuspended in fresh BG11 to a final OD750nm of 2.60. 500µL of *Synechocystis* sp. PCC 6803 suspension was transferred to 1.5mL Eppendorf tubes and incubated with 10µg of plasmid in standard growth conditions for 6h. Every 30 minutes the tubes were gently inverted to obtain a homogeneous distribution of cells and DNA in the transformation process. In the end of the incubation step, cells were plated on BG11 plates without antibiotic, which were left overnight in normal growth conditions. The following day cells were washed out of the plate with 1mL of liquid BG11. The final cell suspension was then plated in fresh BG11 plates with the appropriate antibiotic. They were plated in three dilutions with 10x dilution factor.

Triparental mating

Some cyanobacteria as it is the case of *Anabaena* sp. PCC 7120 are not naturally transformable. To circumvent this difficulty, plasmids harboring constructs of interest can be transferred into *Anabaena* sp. PCC 7120 by conjugation, making use of a well-established triparental mating technique [62]. To do so the *E. coli* strains DH5α harboring the plasmid of interest in addition to the helper plasmid pRL623, and the *E. coli* HB101, harboring the conjugative plasmid pRL443, were grown overnight in LB medium supplemented with the appropriate antibiotics.

The following day, cultures of both strains were inoculated in 25mL of fresh LB medium supplemented with the appropriate antibiotics to an initial OD600nm of 0.1. The cultures were then grown in an orbital shaker at 37ºC for 4h. At the end of this period, cultures were centrifuged at 2100*g* for 10 minutes and washed twice with LB medium. After the last wash, *E. coli* cells were centrifuged and the pellet of the HB101 cells was resuspended in 1mL of LB medium. The HB101 cell suspension was then used to resuspend the DH5α cells, resulting in a cell suspension containing both *E. coli* strains. One additional washing step was carried out and the pellet containing HB101 and DH5α was finally resuspended in 300µL of LB medium. The cell suspension was left resting quietly in a tube rack for approximately 1hour, promoting appropriate conditions for the first conjugation event between the two *E. coli* strains, while the cyanobacterial strain was being prepared for the subsequent steps of the mating. 50mL of the cyanobacterial strain were harvested in mid exponential phase by centrifugation for 10 minutes at 3875*g*. Then the cyanobacterial pellet was washed and resuspended in the least volume possible of BG11 medium. The cyanobacterial cell suspension was mixed with both *E. coli* strains and mating was let to occur for another hour. After that the mixture, now containing all three strains, was plated on BG11 plates supplemented with 5% LB medium and plates were placed under ordinary cyanobacterial growth conditions.

Approximately 30h later plates were washed with 1mL of BG11. The cells washed out from the plate were then transferred and plated onto fresh BG11 plates supplemented with 25µg/mL Kanamycin (for the *Synechocystis* sp. PCC 6803 conjugation) or 40µg/mL Neomycin (for the *Anabaena* sp. PCC 7120 conjugation), in a dilution series of 1%, 10% and undiluted.

Exoproteome extraction and concentration

To analyze the exoproteome, 200mL of each culture were harvested by day 6 of growth, except for *Anabaena* sp. PCC 7120 grown without combined nitrogen which was taken at day 14 due to its slower growth rate in $BG11_0$. The samples were then centrifuged to separate cells from the respective medium. The supernatant containing the exoproteome was then filtered through 0.2 µm pore size filters (Whatman®) to further remove any cell contaminant. The proteins present in the filtered supernatant were then concentrated with Amicon Ultra 15 mL Centrifugal Filter (Millipore®) with 3kDa Molecular Weight Cut-Off by centrifugation at 4000*g*. Concentrated exoproteome samples were either used immediately for further analysis or saved at -20ºC.

Proteome extraction

To obtain cell-free extracts cyanobacterial cells were grown in previous described conditions. Then 50mL of culture was centrifuged for 10 minutes at 3875*g* and 4ºC. The pellet was

resuspended with 2mL of Protein extraction buffer [10 mM HEPES, 0.5% triton X-100, 10mM EDTA, 2 mM DTT, pH 8.0, supplemented with protease inhibitor cocktail (Roche Diagnostics GmbH)], followed by a centrifugation step in a refrigerated microcentrifuge for 5min at 11000*g* and 4ºC. After this washing step, the pellet was finally resuspended in 1.5mL of the same buffer. Cells were disrupted by sonication (Branson Sonifier 250) with cycles of 30 seconds sonication followed by incubation on ice for 2min; a total of 5 cycles were performed for each sample. Each cycle had a duty cycle of 50%. Cell debris and unbroken cells were separated from the extracts by centrifugation 10min at 11000*g* at 4ºC. Cell free extracts were either used immediately for analysis or stored at -20ºC.

The determination of the protein content was performed using in-plate BCA™ Protein Assay (Pierce), using bovine serum albumin (BSA) as standard.

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Polyacrylamide gels were made accordingly to [63] with either 10 or 12% (w/v) acrylamide/Bis-acrylamide (29:1) concentration in the separating gel (Bio-Rad; 1.5mm thick).

Samples for electrophoresis were prepared by mixing the respective sample with loading buffer in a 4:1 proportion. Then samples were heated at 95ºC for 5min, cooled to room temperature and loaded in the wells, as well as the protein marker Precision Plus Protein™ All Blue Standards, (Bio-Rad®). The electrophoresis was then performed with constant electric current of 20 mA per gel and stopped when the bromophenol blue reached the bottom of the gel.

The gel was then incubated with fixation buffer (40% (v/v) methanol and 7% (v/v) acetic acid) for one hour and later transferred on to a staining solution containing Coomassie blue (Brilliant Blue G - Colloidal Concentrate, Sigma-Aldrich®). Staining was carried out in an orbital shaker (60 RPM) overnight, after which the gel was destained with deionized water with several washes.

After the gel was thoroughly destained it was scanned in a Molecular Imager GS800 calibrated densitometer (Bio-Rad).

Superoxide Dismutase (SOD) and Catalase activity in-gel assay

Native-gel electrophoresis

The gels for SOD and Catalase activity assays were made the same way as described for “SDS-PAGE” gels but without the adition of SDS to any component, including the separating and stacking gel, sample buffer and running buffer. If the gel was meant to assay the SOD activity a

gel containing a final concentration of 10% polyacrylamide was chosen, while catalase activity was assayed in 7.5% polyacrylamide gels.

Samples were prepared by mixing the respective extract with native sample buffer (125mM Tris-HCl pH 6,8; 80% (v/v) glycerol; 0.02% (v/v) bromophenolblue) in a 3:1 proportion. For SOD and catalase activity assays 100µg and 60µg of cell free extracts were loaded, respectively. For the exoproteome samples one tenth of the concentrated extracellular medium, corresponding to a total culture volume of 20mL, was used for both activity gels.
Electrophoresis was carried out with a constant voltage of 100V until the bromophenol blue reached the bottom of the gel.

In-gel Zymography

To assay the catalase activity the gel was incubated for 45 min in a 20mL solution of 50µg/mL of HSP (Horseradish Peroxidase) in 50mM KPi. Then 10.25µL of hydrogen peroxide was added to the incubation mixture and incubated for another 10 min in constant agitation. After that period the gel was washed twice with water and developed in a solution containing 0.5mg/mL of DAB (3,3'-diaminobenzidine tetrahydrochloride) in 50mM KH_2PO_4 until the gel got a brown coloration.

The SOD activity was done accordingly to [64], the gel was incubated in 20mL of a 2.5mM NBT (nitroblue tetrazolium) solution for 20 min in darkness with agitation. Then 30mL of SOD development solution (86µM Riboflavin; 28mM TEMED dissolved in 36mM KH_2PO_4) [65] was added to the incubation mixture and it was left with agitation and obscurity for another 15min. The gel was finally exposed to the light of a 60W lamp until the bands appeared.

2D-Gel Electrophoresis

The isoelectric focusing step was performed using 7cm pH 3-10 nonlinear strips (ReadyStrip™ IPG strips, Bio-Rad). The strips were first hydrated overnight (at least 16h) with 100µL of hydration buffer (Bio-Rad) with 25µL of each exoproteome sample. To perform the hydration step, strips were placed upside down on top of the previously described mixture, which was lying on a Rehydration tray (GE Healthcare) and then covered with mineral oil.

The isoelectric focusing step was carried out by putting the hydrated strips into an IEF running unit (Bio-Rad). An electric field with a gradient of 0-2000V for a total of 8000Vh and maximum intensity of 50mA per strip was applied.

After the electrofocusing is done, the strips were incubated in hydration buffer supplemented with 10mg/mL of DTT for 15 min, and then they are incubated for another 15 min in hydration buffer supplemented with 25mg/mL iodoacetamide. The strips are then loaded on top of a 10%

SDS-polyacrylamide gel prepared as described above with the exception that the well comb was not used. A mixture of 1% Agarose with 0.02% bromophenol blue was added on top of the strip, immobilizing it. After the Agarose was gelified the gel was transferred to the electrophoresis tank. Electrophoresis settings and gel fixation and staining were performed as previously described in the "SDS-PAGE" section.

In-Gel Trypsin Digestion

The SDS-polyacrylamide gel pieces were cut out of the gel using a clean scalpel with the least amount of gel possible extracting the whole Coomassie stained band. The gel piece was then put on a siliconized tube (LoBind Tube, Eppendorf®) and covered with 200µL of 200mM ammonium bicarbonate in 40% (v/v) acetonitrile and incubated at 37ºC for 30min. The solution was then discarded and the step was repeated once more. Then the gel piece was dried in a vacuum concentrator. To perform the reduction and alkylation steps the dried gel piece was incubated in 200µL of 10mM DTT for 1 hour at 56ºC and then centrifuged at 13000*g* for 10 seconds. The liquid was discarded. Then 200µL of 55mM Iodoacetamide was added to the gel piece of interest and incubated for 30 min in darkness at room temperature. At the end of this step all liquid was discarded. The gel piece under treatment was washed twice with 200µL of 50mM ammonium bicarbonate and once with 200µL of 50mM ammonium bicarbonate in 40% (v/v) acetonitrile. The gel pieces were centrifuged again for 10 seconds at 13000*g* and all liquid was discarded, before being dried in a vacuum concentrator.
To digest the peptides present in each gel piece, 0.4µg of Trypsin (Promega) dissolved in 40mM ammonium bicarbonate in 9% (v/v) acetonitrile, was added to the gel pieces and incubated overnight at 37ºC.

The following day the liquid was collected to a new siliconized tube. 20µL of 100% acetonitrile were added to the gel piece and incubated at 37ºC for 15min. Then an additional 50µL of 5% formic acid was added to the gel piece, which was still immersed in 100% acetonitrile, and incubated at 37ºC for another 15min. At the end of the incubation step the supernatant was recovered to the new siliconized tube, finally containing the extracted peptides. These were completely dried in the vacuum dryer and stored at -20ºC until further analysis by the Liquid Chromatographer copulated with tandem Mass Spectrometer (LC-MS-MS).

Mass Spectrometry Readings

The mass spectrometry readings were done in an amaZon ETD (Bruker®) mass spectrometer with electrospray ionization and ion trap mass analyzer (ESI-ion trap) copulated with a liquid chromatographer (LC) UltiMate 3000 rapid separation (Dionex®).

The dried digested peptides from the gel piece were dissolved in 10µL Buffer C (3% Acetonitrile, 0.1% Trifluoroacetic acid). Then they were centrifuged at 17000*g* for 10 minutes and transferred to the LC vial, from which the LC was programmed to inject 7µL of sample.

The LC was configured to perform a 90 minute binary gradient with increasing concentration of the organic Buffer B (97% Acetonitrile, 0.1% Formic Acid) relative to the aqueous Buffer A (3% Acetonitrile, 0.1% Formic Acid) with a constant pump flow of 0.300µL/min. The gradient started with 0 % buffer B for 5 minutes, followed by a linear ramp from 5 to 40% buffer B for 70min, then a isocratic wash with 90% buffer B for 7 min, and column re-equilibration at 5% B for 8 min. Between every two sample injections, the LC-MS-MS run a 40 minute isocratic wash with 100% Buffer A to act as a blank and purge the system.

The MS ionization source was calibrated for a capillarity voltage of 3600V and an end plate offset of 500V. The trap scan range was between 300 to 1500m/z and 4 precursor ions were automatically chosen for the MS/MS analysis.

Mass Spectrometry Bioinformatic Data Analysis

After the completion of the mass spectrometry analysis data, .mgf files were created using an in-house algorithm. The peptides, present in each sample were then identified using the software EasyProt (developed between the Biomedical Proteomics Research Group of the Human Protein Science department at the University of Geneva, the Swiss Center for Applied Human Toxicology and the Swiss Institute of Bioinformatics), and the UniProt databases, from July 2013, of each organism used. The parameters used allowed for peptide and MS/MS tolerance up to 0.6 Da and two defined missed cleavages. A set of protein level modifications were defined, methyl-thiol of cysteins (+46Da; Fixed) and oxidation of methionine (+16Da; Variable). Data quality filtering was performed allowing only peptide identification with a Z-score > 5 and a minimum length of 6 aminoacid. To determine which proteins were present in the sample it was chosen a false discovery rate of no more than 1%, calculated using a reverse target-decoy [66] a further stringency of a minimum of two peptides identified per protein was applied. For the results shown in this work a further restriction of 50 minimum protein Z-score was applied in order to increase the results accuracy.

DNA isolation

Genomic DNA was extracted according to a previously described method [67]. In summary, 50mL of cell culture was harvested and centrifuged for 10 minutes at 3875*g*. The cell pellet was washed twice with storage buffer (150mM NaCl; 1mM EDTA; 10mM Tris HCl, pH 8,0). After the

last centrifugation the pellet was resuspended in 250µL of Resuspension Buffer (10mM EDTA; 50mM Tris HCl, pH8,0). To the cell suspension it was added: 350mg of 0.6mm diameter glass beads (Sigma®), 25 µL of 10% SDS and 500 µL of phenol/chloroform 1:1 (pH 7.5). The mixture was thoroughly vortexed for 30 seconds, followed by incubation on ice for another 30 seconds. These 1min cycles were repeated in a total of seven times. In the end the mixture was centrifuged at 11000*g* for 10 minutes at 4ºC to separate the aqueous and organic phases. The aqueous phase was then transferred to a new Eppendorf tube and mixed with 500µL of chloroform, to perform another extraction as previously described. The DNA present in the aqueous phase was precipitated with one tenth of its volume of 3M sodium acetate (pH 5.2) and 2.5 volumes of 100% ethanol. The mixture was incubated at -20 ^{0}C for 30 minutes for the DNA to precipitate, and a centrifugation at 11000*g* and 4ºC for 30min was carried out to pellet the genomic DNA. The supernatant was discarded and the pellet was washed with 70% ethanol. The dried pellet was dissolved in 50µL of deionized water. DNA concentration was determined using a Nanodrop ND-1000 (Thermo Scientific®).

DNA electrophoresis

DNA electrophoresis was carried out in 0.8% agarose gels containing sodium borate buffer (5mM $Na_2B_4O_7{\cdot}10H_2O$) and stained with ethidium bromide as described in [68].

PCR – Polymerase Chain Reaction

The PCR's were carried out with GoTaq® DNA polymerase (Promega) kit, following the manufacturer's instructions. The final magnesium chloride and dNTP concentrations in each reaction were 2mM and 0.2mM, respectively, while 1u of the enzyme was used. Oligonucleotides were included at a final concentration of 1µM. The PCR reactions were carried out in a thermocycler (Bio-Rad) using the following profile: 2 minutes denaturation step at 95ºC; then 35 cycles of 30 seconds at 95ºC, 30 seconds of annealing temperature (table 1) and then 72ºC for 1 minute for every kb of the target DNA; in the end of the cycles another 7 minutes of extension step at 72ºC.

Oligonucleotides Used

Table 1- Oligonucleotides used in this work.

Primer pair/ Target	Sequence 5'-> 3'	Annealing Temp (ºC)	Target size (bp)	Restriction Site
Synechocystis sp. PCC 6803 *rbcL* promoter				
SrbcLF	aaactcgagggatccTCACCATTTGGACAAAACATCAG	65	301	XhoI; BamHI
SrbcLR	aaaaagcttCTAGGTCAGTCCTCCATAAACATTG			HindIII
Synechocystis sp. PCC 6803 *hoxE* promoter				
ShoxEF	aatctcgagggatccATAATTTATTTCTCGTAGGTCTAAC	64	967	XhoI; BamHI
ShoxER	aataagcttGGAAAAATCCTCAAAAAGGAGC			HindIII
slr0513 signal peptide				
Ssp1F	aataagcttTCGATGACAACTAAGATTTCCCG	66	120	HindIII
Ssp1R	atactgcagGGACTGGGCAGAGGCCCG			PstI
Anabaena sp. PCC 7120 *rbcL* promoter				
ArbcLF	aaactcgagggatccAGCGAAACTGACCAAAGGTTAATC	65	838	XhoI; BamHI
ArbcLR	ataaagcttATCTATCCTTCCAAGATGTCACTCTTTTTAC			HindIII
Anabaena sp. PCC 7120 *patS* promoter				
ApatsF	ttactcgagggatccGTCAGTATTGTTTCGGTGATCAGTAG	63	912	XhoI; BamHI
ApatsR	ttaaagcttATTACTTTTCAACAGAACATTTGGTTAC			HindIII
all3333 signal peptide				
Asp1F	taaaagcttAATATGCCAAAGTTAAACAGGCG	66	108	HindIII
Asp1R	ttactgcagATTTAACTGAGTTTTGGCAATAG			PstI
GFP				
GFPF	aaactgcagAAAGGAGAAGAACTTTTC	59	729	PstI
GFPR	attggatccTTATTTGTATAGTTCATCC			BamHI
5´flanking region of *tolC* (slr1270)				
5'FWD	aaactcgagAAGTGTTTCTCCTCCGGACA	63	553	XhoI
5'RV	aaactgcagGGTTTGTTCTGCCACCACTT			PstI
3´flanking region of *tolC* (slr1270)				
3'FWD	aaactgcagCTTACCCCTGGATGAAAGCA	60	517	PstI
3'RV	aaaggatccGCCAATAATCGCCTGTAGGA			BamHI
Kanamycin resistance cassette (*kan*^{r})				
Kan FWD	aaactgcagTGAGGTCTGCCTCGTGAAGAA	59	1219	PstI
Kan RV	aaactgcagAAAGCCACGTTGTGTCTCAAA			PstI

DNA Digestion

All DNA digestions performed in this work were made accordingly with the protocol supplied by the manufacturer (Fermentas®). In case of double restriction reactions the protocol followed was according to the DoubleDigest™ Fermentas® tool (http://www.thermoscientificbio.com/webtools/doubledigest/).

DNA Ligation

All DNA ligations were made with 1u T4 DNA ligase (Thermo Scientific®) according to the instructions of the manufacturer (Thermo Scientific®). In every ligation 100ng of vector were used in 20µL reactions while the amount of insert used was 3 times the molar concentration of the vector.

The ligations were performed at 22^0C for at least 2h, followed by an inactivation step at 65^0C for 10 minutes before transformation in *E. coli*.

DNA Purification and quantification

DNA purification from gel or from enzymatic reactions was carried out with a NZYGelpure (Nzytech, Lda.) kit following the manufacturer instructions. DNA concentration was determined using a Nanodrop ND-1000 (Thermo Scientific®)

Plasmid Preparation

Plasmid preparation was done with a GenElute™ plasmid miniprep kit (SIGMA®) following the manufacturer instructions and using 5mL of an overnight grown culture. DNA concentration was determined using a Nanodrop ND-1000 (Thermo Scientific®).

DNA Sequencing

The DNA sequencing was performed by Stab Vida Company.

Confocal Microscopy

Anabaena sp. PCC 7120 harboring pRL25C, a self-replicating plasmid expressing a fusion protein with the N-terminal signal peptide from the periplasmic *nrtA* (all3333*)* together with GFP, under the control of either the *patS* or the *rbcL* promoters, were studied using a Leica SP2 AOBS SE laser scanning confocal microscope. Cells grown either in non-nitrogen fixing or nitrogen-fixing conditions were loaded on a 1% low-melting-point agarose bed, dissolved in BG11 or $BG11_0$, respectively, and covered with a clover slip. The GFP emission, collected between 500 and 540 nm, was observed when cells were exposed to an Ar laser beam of 488 nm, while cyanobacterial autofluorescence, acquired between 640 and 700 nm, was visualized after excitation at 633 nm, using a HeNe laser. Wild-type cells were used to define the basal autofluorescence signal in the GFP channel, and the same acquisition settings were used throughout the various experiments.

Strategy for the generation of periplasmic GFP mutants

Various pRL25C based plasmids were constructed, containing different well characterized native promoters to drive expression of the signal peptide fused with the GFP encoding gene, and were transferred by conjugation into the corresponding cyanobacterium. For the *Anabaena* sp. PCC 7120 mutants the promoters chosen were the *rbcL* (alr1524 – encodes the large subunit of RuBisCO) promoter, which is a strong constitutive promoter [69], and the *patS* (asl2301) promoter, which is activated in cells that initiate differentiation upon deprivation of fixed nitrogen [70], in addition, the signal peptide chosen was the one from NrtA (all3333), a well described periplasmic protein involved in nitrate import [9]. For the *Synechocystis* mutants the promoters chosen were the *rbcL* (slr0009 – encodes the large subunit of RuBisCO) promoter, which is a strong constitutive promoter [69] and the *hoxE* (sll1220 – encodes one of the subunits of the bidirectional hydrogenase) promoter, which is a weak constitutive promoter; additionally, the signal peptide chosen was from FutA2 (slr0513), which is a periplasmic iron binding protein [71]. In the end two constructs for each strain were obtained as shown in figure 3.

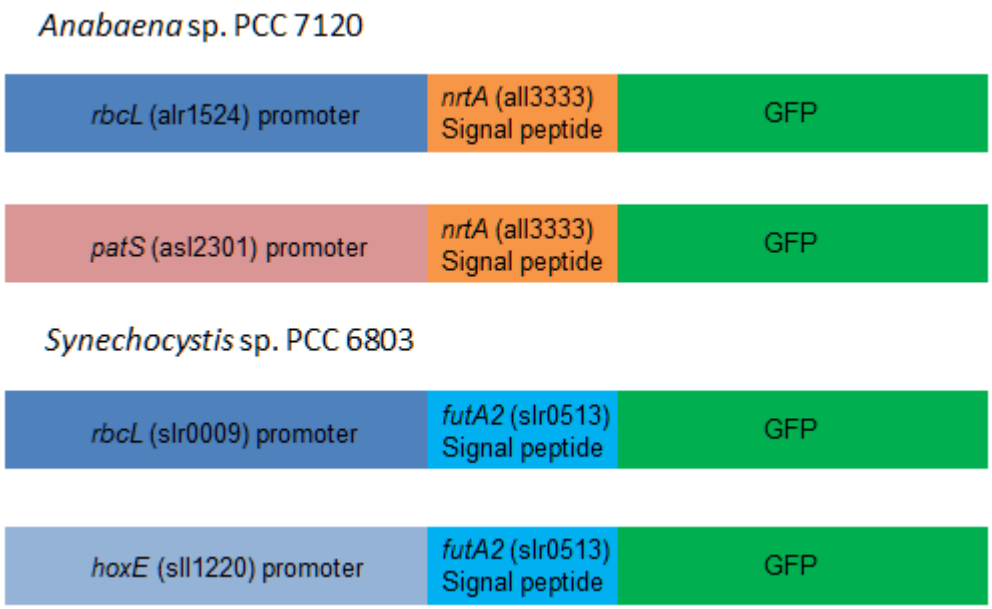


Fig 3 – Scheme of the four constructs cloned in the pRL25C plasmid.

To isolate the *Anabaena* sp. PCC 7120 fragments *rbcL; patS and nrtA* signal peptide it was made a PCR using the *Anabaena* sp. PCC 7120 genome and the oligonucleotides from table 1. To isolate the *Synechocystis* fragments *rbcL; hoxE and futA2* signal peptide it was made a PCR using the *Anabaena* sp. PCC 7120 genome and the oligonucleotides from table 1. To isolate the

GFP it was made a PCR using as template an in-house plasmid containing GFP gene and the oligonucleotides from table 1.

The constructs were cloned inside a Bluescript SK+ plasmid and transformed in *Escherichia coli* XL1 blue. The plasmid was then sent to sequencing to verify there were no errors in the PCR amplification. After the confirmation the constructs were excised of the Bluescript SK+ and cloned in the final plasmid (pRL25C) and transformed into *E. coli* DH5α already with the helper plasmid pRL623. The mutants were then ready to conjugate the pRL25C with the constructs into the appropriate cyanobacteria using the Triparental mating technique.

Strategy for the generation of Δ*tolC* mutants

The *Anabaena* sp. PCC 7120 *ΔtolC* mutant was already being made in our lab, so it was only necessary to make the *Synechocystis* sp. PCC 6803 one. For that purpose, it was constructed a pSK+ plasmid with a 5´ region of the ORF of the slr1270 gene (region between 122-675bp of the ORF) followed by a Kanamycin resistance cassette and a 3´ region of the same ORF (region between 1020-1503 bp of the ORF). This construct was made with the objective of making double homologous recombination with the gene present in the chromosomes that way disrupting the ORF and preventing the production of TolC.

To make the mutant the ORF regions were amplified with the primer 5´FWD and 5´REV for the 5´fragment and 3´FWD and 3´REV for the 3´Fragment (table 1). The inserts 5´and 3´were cloned inside the plasmid pSK+ as schemed on figure 4.

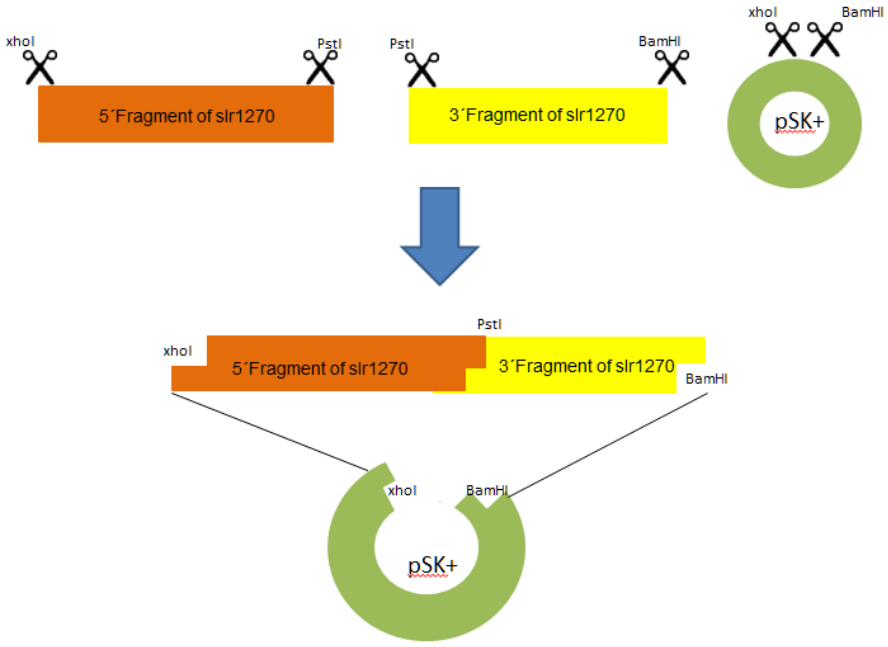


Fig 4- Scheme of the cloning strategy used for the pSK+ with the flanking regions of slr1270

After the cloning the plasmid was sent to sequencing to verify the correct sequence of the flanking regions of the slr1270.

To make the selection marker, the Kanamycin cassette was amplified using the Kan FWD and the Kan REV primers (table 1). After the amplification it was digested with PstI and cloned between the 5´fragment and the 3´fragment previously cloned in the pSK+. This new plasmid was then transformed in *E. coli* XL1 blue which was plated on LB with kanamycin. Because only mutants with functional Kanamycin cassette would grow in the plates, there is no need to sequence the cassette. The mutant growth on LB with Kanamycin is proof enough the cassette is working.

The final plasmid inside the *E. coli* mutants was then extracted and used to transform *Synechocystis* sp. PCC 6803. The *Synechocystis* sp. PCC 6803 mutants were then plated in BG11 with increasing concentration of Kanamycin to promote selective pressure of the cassette and thus promoting the copy segregation on all the chromosomes.

Results and Discussion

Exoproteome Identification

In order to identify the peptides present in the different exoproteome samples the first step is to determine the complexity of the exoproteomes to decide the strategy for the mass spectrometry identification. To have an idea of the protein pattern for each exoproteome sample SDS-polyacrylamide gels were used to separate the various samples

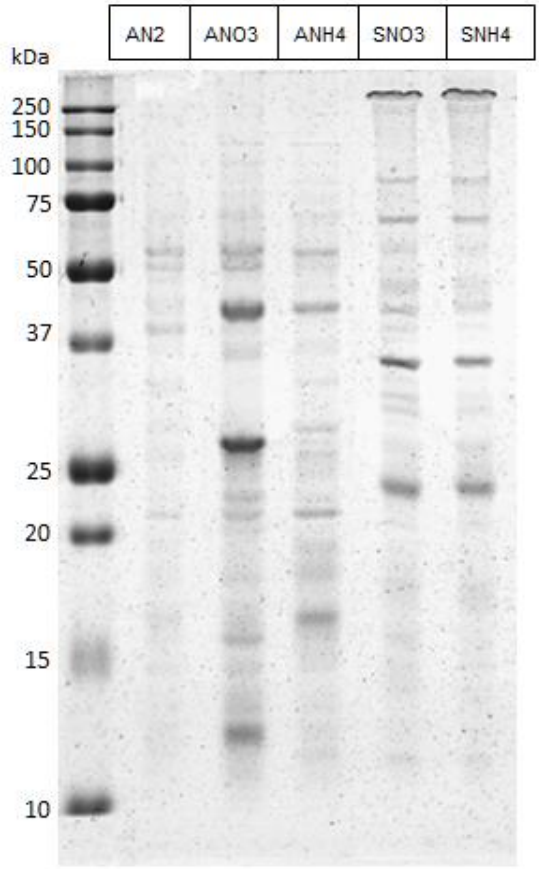


Fig 5- SDS-PAGE separation of the exoproteome samples. First lane is the molecular marker, the lanes two to six are *Anabaena* sp. PCC 7120 grown in $BG11_0$ (AN2); *Anabaena* sp. PCC 7120 grown in BG11 (ANO3); *Anabaena* sp. PCC 7120 grown in $BG11_0$ supplemented with NH_4Cl (ANH4); *Synechocystis* sp. PCC 6803 grown in BG11 (SNO3) and *Synechocystis* sp. PCC 6803 grown in $BG11_0$ supplemented with NH_4Cl (SNH4) respectively.

The SDS-PAGE results showed that each sample have few bands and thus a small array of different proteins. For that reason it was decided that the strategy for the exoproteome identification was to separate the exoproteome samples on a regular SDS-polyacrylamide gel, followed by cutting out gel bands and finally to proceed with the identification of the peptides present in each band by MS. This approach has the advantage of giving second validation to the MS technique because excised bands contain proteins of a given molecular mass that can be

compared to the theoretical mass of the proteins identified. If the molecular mass is similar, the confidence of that correct protein identification is increased. However, it has the disadvantage of analyzing only fractions of the gel which possess proteins that can be detected by the staining method in use (in this case, Coomassie Blue staining). That can make some of the low abundant or less stained proteins to be left out of the analysis.

It was also noticed that although the band pattern changes noticeably in the different *Anabaena* sp. PCC 7120 samples grown in different nitrogen source they are similar in both *Synechocystis* sp. PCC 6803 samples grown with nitrate or ammonia (figure 5). For that reason it was decided that *Synechocystis* sp. PCC 6803 grown in ammonia would not be chosen for further analysis, to prioritize the resources for other conditions.

To validate this strategy for MS analysis it was important to know how complex the exoproteomes were. If each sample was not enriched in multiple proteins and peptides, simple enough, that most proteins could be separated and resolved only according to their mass in an SDS-polyacrylamide gel, then each band may be cut out, processed and analyzed directly in the LC-MS-MS. Alternatively, if the sample was too complex and each band on an SDS-polyacrylamide gel represented several different proteins, then the sample should be submitted to further separation prior to the injection. To confirm that the samples are in fact of low complexity it is important to know if each band represents only a few different proteins instead of being the sum of many proteins with the same mass. 2D-gel electrophoresis was performed for all exoproteome samples.

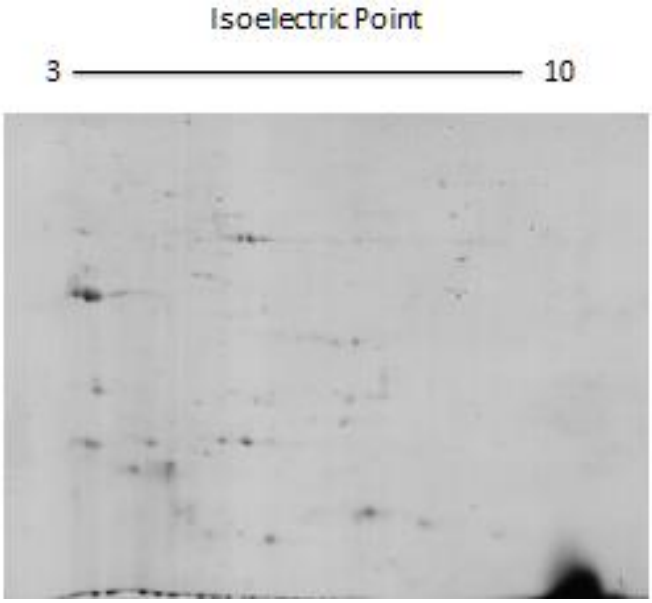


Fig 6- 2D-gel separation of the *Anabaena* sp. PCC 7120 exoproteome. *Anabaena* sp. PCC 7120 was grown for 10 days in $BG11_0$. The isoelectric focusing strips have a pH range of 3 (left) to 10 (right) non-linear, and the gel has 10% polyacrylamide.

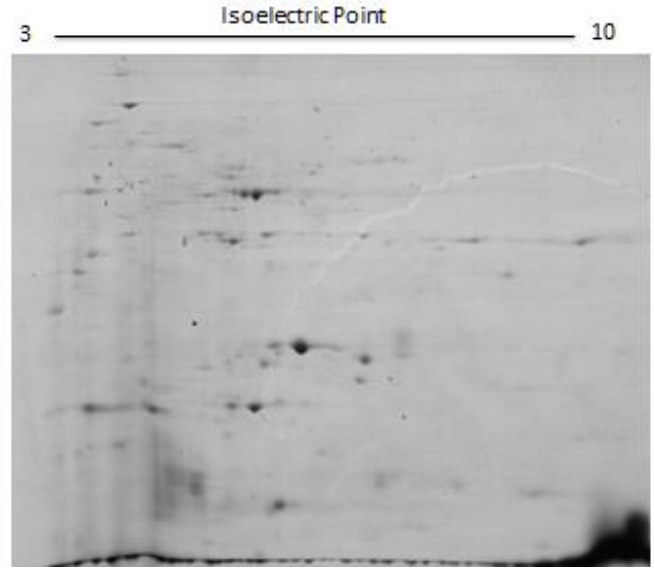


Fig 7- 2D-gel separation of the *Anabaena* sp. PCC 7120 exoproteome. *Anabaena* sp. PCC 7120 was grown for 6 days in BG11. The isoelectric focusing strips have a pH range of 3 (left) to 10 (right) non-linear, and the gel has 10% polyacrylamide.

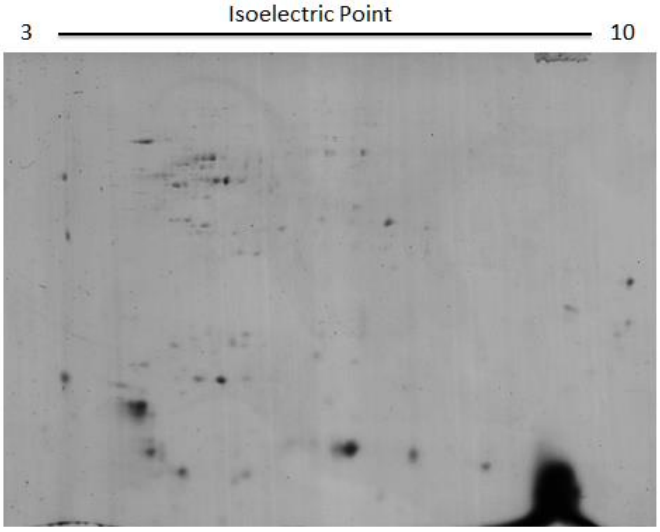


Fig 8- 2D-gel separation of the *Anabaena* sp. PCC 7120 exoproteome. *Anabaena* sp. PCC 7120 was grown for 6 days in $BG11_0$ supplemented with NH4Cl. The isoelectric focusing strips have a pH range of 3 (left) to 10 (right) non-linear, and the gel has 10% polyacrylamide.

After the 2D-Gel analysis (figures 6-8) was completed, it was possible to conclude that all exoproteome samples of *Anabaena* sp. PCC 7120 are of relative low complexity. The same was concluded for the *Synechocystis* sp. PCC 6803 samples (data not shown). In fact, each band detected on the SDS-polyacrylamide gels (see below) is composed either by a single peptide or, in the worst case, by a relatively low number of peptides (see figures 6-8). Therefore, it was decided that all the exoproteome samples should be analyzed in the LC-MS-MS after a simple SDS-polyacrylamide gel electrophoresis separation, after which bands were excised and processed for identification.

To separate the proteins for MS analysis the samples were separated on a new SDS-PAGE this time loading the maximum amount of protein possible in each well. The most visible bands of each condition, with the exception of the SNH4 for reasons previously referred, were then chosen for analysis.

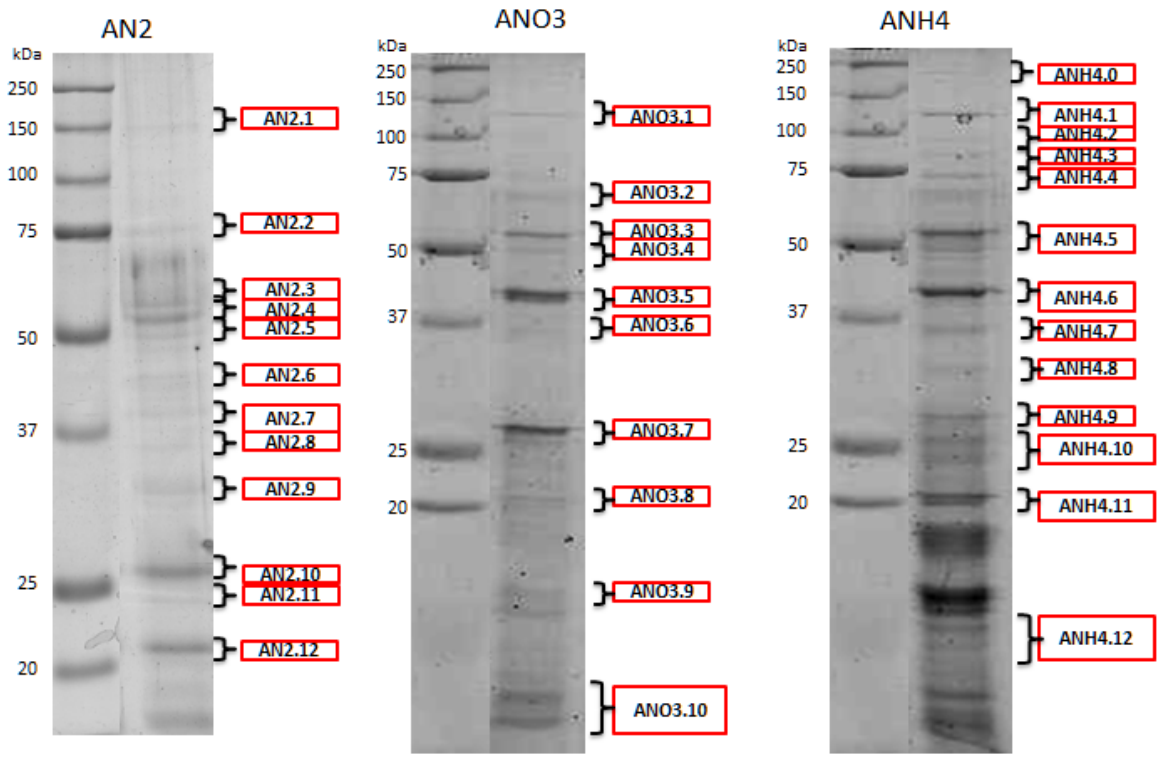


Fig 9- SDS-PAGE separation of the *Anabaena* sp. PCC 7120 exoproteome. The gel on left is the exoproteome of *Anabaena* sp. PCC 7120 grown 14 days in $BG11_0$ (AN2). The gel on the middle is the exoproteome of *Anabaena* sp. PCC 7120 grown 6 days in BG11 (ANO3). The gel on the right is the exoproteome of *Anabaena* sp. PCC 7120 grown 6 days in $BG11_0$ supplemented with NH_4Cl (ANH4). First lane of each gel is the molecular marker with the corresponding weights marked in kDa, on the left. Second lane is the exoproteome with the red rectangles pointing to the gel locations where the gel was cut for further process and MS identification. Inside the rectangles it was written the code of each fraction.

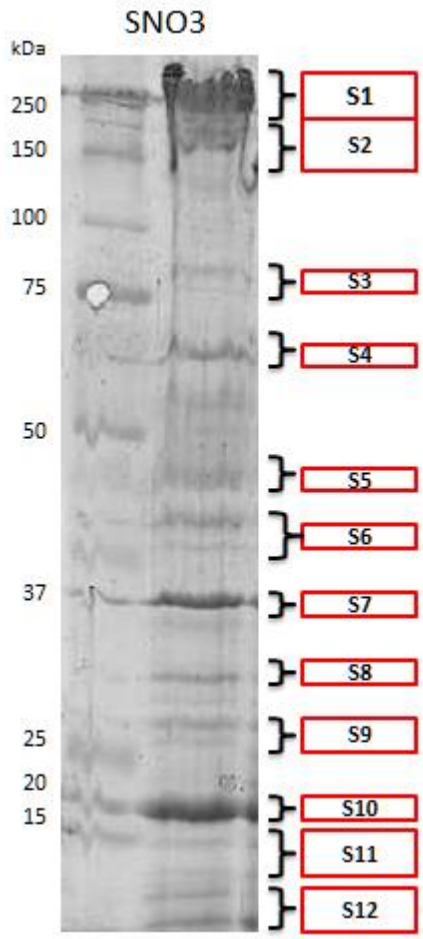


Fig 10- SDS-PAGE separation of the *Synechocystis* sp. PCC 6803 exoproteome. The gel on the left is the exoproteome of *Synechocystis* sp. PCC 6803 grown 6 days in BG11 (SNO3). First lane of each gel is the molecular marker with the corresponding weights marked in kDa, on the left. Second lane is the exoproteome with the red rectangles pointing to the gel locations where the gel was cut for further process and MS identification. Inside the rectangles it was written the code of each fraction.

In figures 9 and 10 it is represented the bands excised for MS analysis and their respective code. A total of 48 bands were chosen, processed and the peptides therein further identified.

After the mass spectrometer analyzed all the samples they were run on easyprot to identify which proteins were present in each piece of the gel. The results for *Anabaena* sp. PCC 7120 are shown on table 2 and the results for *Synechocystis* sp. PCC 6803 are shown on table 3. For each list it is displayed the accession number, the protein description, the protein score, the percentage of coverage, the number of total and unique peptides identified, the theoretical protein mass and the nitrogen source in which the protein was identified together with the number of the band where it was found. If a protein is identified in more than one band it was only shown the MS parameters for the one with the top score, although the others are mentioned in the band section in order of descending protein score. The list is ordered with descending protein mass to correlate with the bands of the SDS-PAGE separation.

Table 2- List of proteins identified in *Anabaena* sp. PCC 7120 exoproteome.

AC	Description	Protein Score	% Covr.	#Pept	#Uniq Pept.	Protein Mass (Da)	Nitrogen source	Bands
All2655	All2655 protein	159,402	3,31	11	10	339468	NH_4^+	0
Alr4238	Alr4238 protein	404,401	11,59	32	18	193292	NH_4^+	0; 3
All0275	Glycerophosphoryl diester phosphodiesterase	418,499	20,74	29	18	109861	NO_3^-	1
		511,361	25,12	36	20	109861	NH_4^+	1; 3; 2
All0167	Maltooligosyltrehalose synthase	95,705	8,46	7	6	105858	NH_4^+	4
Alr3588	Alr3588 protein	312,695	18,28	25	12	96326	NH_4^+	2
Alr1310	Alr1310 protein	148,834	12,1	7	7	85772	NH_4^+	3
Alr2887	Alr2887 protein	64,69	6,87	4	4	80838	N_2	4
Alr2887	Alr2887 protein	116,781	11,19	6	6	80838	NO_3^-	2
Alr4794	Alr4794 protein	456,943	26,51	36	19	79486	NO_3^-	2
Alr0880	Oligopeptidase A	221,097	18,09	15	10	78181	NH_4^+	4
All2533	Prolyl endopeptidase	152,433	13,21	8	7	77917	N_2	2
		173,232	14,95	13	8	77917	NH_4^+	4
Alr0169	Cyclomaltodextrin glucanotransferase	60,473	7,32	4	4	74287	NH_4^+	4
Alr3344	Transketolase	215,334	18,06	10	9	72071	N_2	2
		464,207	28,96	26	19	72071	NH_4^+	4; 6
Alr1381	Calcium-dependent protease	171,603	13,14	10	8	71527	NO_3^-	4
All0168	Alpha-amylase	65,048	8,71	4	4	70895	NO_3^-	2
Alr1742	Chaperone protein dnaK2	362,633	28,28	27	16	67908	NH_4^+	8
All2843	Alkaline phosphatase	372,079	32,06	23	15	67221	NO_3^-	2
		82,421	14,04	5	5	67221	NH_4^+	4
Alr0474	Alr0474 protein	153,683	13,42	10	7	66225	NO_3^-	6; 8
All1342	All1342 protein	77,334	8,86	4	4	65734	NO_3^-	2
All0875	All0875 protein	165,873	13,77	9	7	63834	N_2	2
		219,467	18,3	14	10	63834	NH_4^+	4
Alr3659	Alr3659 protein	57,56	6,62	3	3	61699	N_2	2
Alr0140	Periplasmic oligopept.-bind protein of ABC trans.	288,687	26,09	15	11	60659	N_2	4; 3
		241,036	23,73	11	10	60659	NO_3^-	3
		157,206	17,21	7	7	60659	NH_4^+	5
Alr4550	Uncharacterized protein alr4550	264,913	25,57	22	11	60581	N_2	5; 6; 4
		275,097	25,57	23	11	60581	NO_3^-	4
All0207	All0207 protein	79,689	9,59	4	4	59997	NO_3^-	4
Alr4976	Phosphodiesterase/alkaline phosphatase D	204,085	18,7	13	9	59168	NO_3^-	3
Alr2771	Dihydroxy-acid dehydratase	130,558	10,83	7	6	59130	NO_3^-	2
All4499	All4499 protein	251,768	24,73	12	10	58564	N_2	5; 4
		317,611	33,33	25	12	58564	NO_3^-	4; 3
All7614	All7614 protein	63,572	6,22	5	4	58354	N_2	4
		72,567	8,23	5	4	58354	NO_3^-	4
Alr1050	Glucose-6-phosphate isomerase	507,82	35,23	35	19	57819	N_2	3; 4
		317,196	25,76	18	12	57819	NO_3^-	3
		343,058	27,08	20	14	57819	NH_4^+	5
Alr2190	Alpha-amylase	62,656	7,52	5	4	57063	N_2	5
Alr0834	Porin major outer membrane protein	138,507	18	9	7	54439	N_2	5; 6
		123,31	14,68	8	6	54439	NO_3^-	4
All0005	ATP synthase subunit alpha	99,212	13,04	5	5	54422	N_2	6; 3
		140,127	15,81	9	6	54422	NH_4^+	5
All3325	All3325 protein	144,782	19,83	7	7	53193	NH_4^+	5
Alr2328	Glutamine synthetase	276,37	26,16	21	12	53136	N_2	5; 4; 3; 6
		287,235	29,96	25	12	53136	NO_3^-	3; 2; 4; 5
		249,44	23,63	19	10	53136	NH_4^+	5;6
Alr1524	Ribulose bisphosphate carboxylase large chain	51,758	9,45	3	3	53045	N_2	5
		151,829	18,49	7	7	53045	NO_3^-	4

Alr1329	Alr1329 protein	68,209	10,21	4	4	52720	N_2	7
All5039	ATP synthase subunit beta	196,109	23,44	13	8	52016	N_2	4; 3; 5
		118,827	17,84	7	6	52016	NO_3^-	4
		98,254	15,35	8	5	52016	NH_4^+	7; 5
Alr0237	Probable cytosol aminopeptidase	124,023	12,63	6	5	51918	NH_4^+	5
All3964	Phosphoglucomutase/phosphomannomutase	336,806	30,95	14	13	51820	N_2	3; 4
		253,579	28,63	14	11	51820	NO_3^-	3
		388,981	30,74	23	14	51820	NH_4^+	5;8;10
All4388	All4388 protein	96,658	12,98	5	5	51580	N_2	5
Alr0996	Protease	148,059	17,21	8	6	51461	N_2	8
Alr4745	Dihydrolipoyl dehydrogenase	68,118	9,26	4	4	50451	NO_3^-	4
Alr4072	Alr4072 protein	81,345	11,36	4	4	50373	N_2	1; 3
		94,098	11,36	8	5	50373	NO_3^-	3; 4
		66,903	8,72	6	4	50373	NH_4^+	7; 2; 5
Alr2877	Bicarbonate transport bicarbonate-binding protein	338,257	28,38	43	14	50292	NO_3^-	5; 7
		338,208	24,67	54	14	50292	NH_4^+	6; 10; 9
All4968	Glutathione reductase	153,796	15,9	8	7	49478	N_2	5; 4
		207,023	21,79	11	9	49478	NO_3^-	4
All4539	L-sorbosone dehydrogenase	185,747	19,87	8	8	49217	N_2	6
		174,29	18,53	10	8	49217	NO_3^-	5
		221,889	22,54	9	9	49217	NH_4^+	6
All4121	Ferredoxin--NADP reductase	347,839	35,45	26	14	48838	N_2	9; 7; 8; 10
		87,869	11,36	6	5	48838	NO_3^-	10
		148,413	17,05	8	7	48838	NH_4^+	7
All3556	Succinate-semialdehyde dehydrogenase	86,477	12,53	6	4	48820	N_2	5
		128,477	16,48	6	5	48820	NO_3^-	4
Alr0608	Nitrate transport protein NrtA	361,162	31,36	31	14	48476	NO_3^-	5; 8; 4; 9
		288,499	25,23	29	12	48476	NO_3^-	7
		186,252	19,32	11	8	48476	NH_4^+	6; 10
All1951	Substrate-binding protein of ABC transporter	191,508	27,4	10	9	47683	N_2	6
		316,316	34,47	18	13	47683	NO_3^-	5
		233,87	29,68	16	10	47683	NH_4^+	6
Alr0132	Alr0132 protein	88,996	15,77	4	4	46552	NH_4^+	6
All3538	Enolase	139,08	16,55	10	6	45965	N_2	6
		151,209	16,55	8	6	45965	NH_4^+	6
Alr5103	LL-diaminopimelate aminotransferase 1	241,106	23,6	14	10	44987	N_2	5
		126,055	15,57	9	6	44987	NO_3^-	4
Alr1834	Alr1834 protein	201,243	26,76	10	9	44801	N_2	6
		179,25	26,76	9	8	44801	NO_3^-	5
		174,122	24,49	10	8	44801	NH_4^+	6
Alr0267	Alr0267 protein	323,961	27,45	26	13	44342	N_2	7; 9; 10; 11; 6; 8
		132,87	15,2	13	6	44342	NO_3^-	10; 7
Alr4448	Endo-1,4-beta-xylanase	65,152	8,9	5	4	44046	N_2	7
All1683	Phosphoserine aminotransferase	58,769	10,97	3	3	43141	N_2	7
Alr1299	Phosphoribosylglycinamide formyltransferase 2	60,178	9,72	4	3	42898	N_2	6
All4131	Phosphoglycerate kinase	353,365	39,25	15	13	42441	N_2	6; 7
		288,49	33,75	12	11	42441	NO_3^-	5
		208,843	24,25	9	8	42441	NH_4^+	6; 10; 12; 9
Alr1004	Alanine--glyoxylate aminotransferase	73,455	15,75	4	4	41869	N_2	7
All4575	Phosphate-binding periplasmic protein of ABC transp.	241,371	26,79	10	8	41343	N_2	8; 7
		285,911	32,14	18	10	41343	NO_3^-	6
		307,389	32,14	23	11	41343	NH_4^+	7; 10
Alr0051	IMP dehydrogenase	76,753	14,47	4	4	40447	NO_3^-	5
Alr1364	Alr1364 protein	162,624	20,06	12	6	39762	NH_4^+	7

All3909	Uroporphyrinogen decarboxylase	96,988	12,86	7	4	38820	NH_4^+	7
All4563	Fructose-bisphosphate aldolase	63,812	11,14	5	3	38617	NH_4^+	7
Alr4123	Phosphoribulokinase	72,645	10,78	4	3	38455	NH_4^+	12
All4050	All4050 protein	84,542	14,81	5	4	37259	NH_4^+	12
All5062	Glyceraldehyde-3-phosphate dehydrogenase 2	109,119	15,43	6	5	36910	NH_4^+	7
All2563	Transaldolase	247,379	30,72	12	9	36178	N_2	8
		279,784	34,94	15	11	36178	NO_3^-	6
		135,117	19,88	7	6	36178	NH_4^+	7
All2315	Ketol-acid reductoisomerase	201,451	22,05	12	8	36011	NO_3^-	6
		188,55	19,94	14	8	36011	NH_4^+	7
Alr2948	Alr2948 protein	119,707	15,22	7	5	35883	NH_4^+	7
All0004	ATP synthase gamma chain	84,993	17,78	4	4	35278	NH_4^+	10; 8
Alr1548	Alr1548 protein	164,826	23,08	7	7	34470	NO_3^-	7
		109,963	16,35	6	5	34470	NH_4^+	10
Alr4907	Ornithine carbamoyltransferase	51,543	9,8	4	2	33401	NO_3^-	6
		50,688	9,48	4	2	33401	NH_4^+	7
Alr0530	Phycobilisome 32.1 kDa linker polypeptide, rod	240,816	38,81	12	10	32209	N_2	9
		53,609	10,84	3	3	32209	NO_3^-	8
		387,439	47,2	29	15	32209	NH_4^+	7; 9; 11; 10
Alr0534	Phycobilisome rod-core linker polypeptide CpcG1	66,886	12,9	3	3	31933	NO_3^-	9
		59,293	12,9	3	3	31933	NH_4^+	12
Alr7261	Alr7261 protein	129,563	22,92	8	5	31514	NH_4^+	10
All2316	Aldo/keto reductase	71,642	14,79	5	3	31453	NH_4^+	7
Alr1362	Alr1362 protein	79,361	15,44	5	4	31369	NO_3^-	8
Alr0525	Phycobilisome 34.5 kDa linker polypeptide, phycoerythrocyanin-associated, rod	53,12	12,23	3	3	31295	NO_3^-	9
Alr3539	Alr3539 protein	117,739	17,14	5	5	30959	N_2	10
		118,424	17,14	5	5	30959	NO_3^-	10; 7
		336,731	37,5	21	12	30959	NH_4^+	10
Alr5182	Oxidoreductase	170,411	28,07	8	7	30915	NH_4^+	10; 8
Alr2313	Alr2313 protein	55,653	9,59	3	3	29543	NH_4^+	11
Alr7346	Alr7346 protein	100,149	24,81	6	4	29430	NH_4^+	7
Alr2535	Branched-chain AA ABC transport, periplasmic protein	155,614	25,37	9	6	29196	N_2	7
All2425	All2425 protein	115,718	18,43	6	5	27812	NH_4^+	9
All4464	Phosphoadenosine phosphosulfate reductase	51,983	14,52	5	3	27556	NO_3^-	7
All0268	All0268 protein	68,415	14,12	5	4	27330	NH_4^+	10
Alr0069	Ribonuclease PH	92,786	21,86	4	4	27152	NH_4^+	10
All7633	All7633 protein	80,292	19,42	4	4	27082	NO_3^-	8
		54,047	13,64	3	3	27082	NH_4^+	11
Alr4385	Triosephosphate isomerase	136,047	22,41	7	6	26552	NH_4^+	9
Alr3607	Alr3607 protein	86,1	14,66	6	4	25991	NO_3^-	9
		70,547	14,66	3	3	25991	NH_4^+	12
Alr3090	Alr3090 protein	170,917	25,65	10	7	25665	NH_4^+	10
All3797	Beta-Ig-H3/fasciclin (Fragment)	72,285	14,59	3	3	24903	NH_4^+	12
Alr4979	Alr4979 protein	255,703	40,09	19	9	24200	NH_4^+	10
All4214	50S ribosomal protein L4	76,51	18,57	4	4	23691	NH_4^+	10
All3791	Ribonuclease D	108,59	20,1	6	5	23598	NH_4^+	10
Alr1965	ATP phosphoribosyltransferase	97,129	18,22	5	4	23325	NO_3^-	8
		62,603	14,02	3	3	23325	NH_4^+	11
All2105	FMN-dependent NADH-azoreductase	93,902	22,71	6	5	22785	NH_4^+	10
Alr4641	Peroxiredoxin	109,796	25,62	6	5	22631	NH_4^+	11
Alr2938	Superoxide dismutase	98,326	21,5	7	4	22460	N_2	12
		183,84	33,5	11	7	22460	NO_3^-	8

		233,166	37,5	31	9	22460	NH_4^+	11
All2108	All2108 protein	66,864	16,92	3	3	22283	NO_3^-	8
Alr7524	Alr7524 protein	90,319	22,8	5	4	21723	NO_3^-	8
		116,263	24,87	9	5	21723	NH_4^+	11
All1380	All1380 protein	162,475	39,06	9	6	21223	NO_3^-	8
		131,837	32,29	9	5	21223	NH_4^+	11
Alr3808	Nutrient stress-induced DNA-binding protein	141,914	29,35	9	5	20687	N_2	1
Alr0528	C-phycocyanin subunit beta	87,848	33,53	5	4	18387	N_2	6; 2; 9; 11; 7
		121,291	32,95	7	5	18387	NO_3^-	9; 8
		130,47	32,95	11	5	18387	NH_4^+	12; 11; 10
Alr0523	Phycoerythrocyanin subunit beta	138,712	36,63	7	5	18284	NH_4^+	12
All4038	All4038 protein	111,254	30,86	6	4	17884	NH_4^+	12
All3653	Allophycocyanin subunit alpha-B	128,488	27,95	9	5	17811	NH_4^+	12
Alr0529	C-phycocyanin alpha chain	224,364	44,79	48	7	17458	NO_3^-	9
Alr0021	Allophycocyanin subunit alpha 1	87,128	26,09	4	4	17346	N_2	9; 11; 8; 7
		171,815	45,34	7	6	17346	NO_3^-	9
		142,106	31,68	10	6	17346	NH_4^+	12
Alr0022	Allophycocyanin subunit beta	71,019	22,22	3	3	17305	N_2	9; 8; 6; 7; 10; 11
		56,835	20,99	4	3	17305	NO_3^-	10
		199,149	48,15	27	7	17305	NH_4^+	12; 11
Alr3402	Nucleoside diphosphate kinase	182,223	47,65	13	7	16604	NH_4^+	12
All2375	All2375 protein	90,166	30,34	3	3	15341	NH_4^+	12
All4287	Peptidyl-prolyl cis-trans isomerase B	77,934	28,47	6	3	14874	NH_4^+	10
All4749	All4749 protein	104,989	41,79	4	4	14695	NH_4^+	12
Alr0529	Phycocyanin A subunit (Fragment)	82,921	43,96	4	3	9757	N_2	9; 7; 6; 8
		143,747	60,44	10	4	9757	NH_4^+	12; 11; 10

It was identified a total of 46 proteins in the 12 gel pieces analyzed for Anabaena sp. PCC 7120 grown without any combined nitrogen source, 59 proteins in the 10, *Anabaena* sp. PCC 7120 grown in nitrate, gel pieces and 80 proteins in the 13 gel pieces from *Anabaena* sp. PCC 7120 grown in ammonia. The total different proteins identified between the three conditions were 117 (Table 2). The great increase in exoproteins when grown in ammonia can be caused by toxicity of the ammonia itself which could lead to intracellular protein contamination, or it could simply mean that the cells secrete more proteins in that condition.

Table 3- List of proteins identified in *Synechocystis* sp. PCC 6803 exoproteome.

AC	Description	Protein Score	% Covr.	#Pept.	#Uniq Pept.	Protein Mass (Da)	Nitrogen Source	Band
Sll0724	Sll0723 protein	403,702	14,12	29	21	192167	NO_3^-	2; 1
Sll1951	Hemolysin	130,748	4,6	30	8	178261	NO_3^-	1; 2
Slr1227	Chloroplast import-associated channel IAP75	173,207	12,31	11	9	92307	NO_3^-	3
Slr1855	Slr1855 protein	57,895	8,57	7	4	70004	NO_3^-	4
Slr1841	Slr1841 protein	112,117	10,95	36	6	67601	NO_3^-	4;2
Slr0442	Slr0442 protein	71,156	9,82	4	4	62908	NO_3^-	3
Sll1009	Iron-regulated protein	54,453	5,58	13	3	60484	NO_3^-	3
Slr2004	Slr2004 protein	80,964	10,55	8	5	53482	NO_3^-	5
Slr0040	Bicarbonate-binding protein CmpA	178,137	23,01	15	9	49461	NO_3^-	5
Sll1450	Nitrate transport protein NrtA	132,658	19,73	13	7	48967	NO_3^-	5
Slr1751	Carboxyl-terminal protease	401,046	40,43	80	17	46833	NO_3^-	6
Slr1165	Sulfate adenylyltransferase	111,73	18,72	11	7	43696	NO_3^-	5
Slr1722	IMP dehydrogenase subunit	64,237	10,08	5	4	40235	NO_3^-	6
Sll1306	Sll1306 protein	129,754	17,31	15	7	38271	NO_3^-	8
Slr0513	Iron uptake protein A2	335,87	39,88	90	14	38156	NO_3^-	7; 10; 9
Sll1491	Uncharacterized WD repeat-containing protein sll1491	80,147	10,06	12	5	37347	NO_3^-	8
Slr1410	Uncharacterized WD repeat-containing protein slr1410	74,258	13,77	6	4	35990	NO_3^-	8
Sll0314	Sll0314 protein	77,889	14,01	5	4	35444	NO_3^-	9
Sll0837	Sll0837 protein	66,629	13,95	4	4	32907	NO_3^-	8
Sll0319	Sll0319 protein	169,897	23,57	16	9	32316	NO_3^-	5
Sll1579	Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated, rod 2	73,67	16,48	5	4	30797	NO_3^-	8
Sll1785	Sll1785 protein	70,879	16,79	5	4	30011	NO_3^-	9
Sll1298	Putative carboxymethylenebutenolidase	60,252	15,85	4	3	26942	NO_3^-	9
Sll1810	50S ribosomal protein L6	65,586	20,67	6	4	19667	NO_3^-	11
Slr1622	Inorganic pyrophosphatase	55,134	10,65	3	3	19088	NO_3^-	10
Sll1577	C-phycocyanin beta chain	99,699	26,16	16	5	18126	NO_3^-	11; 12
Sll1578	C-phycocyanin alpha chain	64,889	32,1	4	3	17587	NO_3^-	9; 12; 11
Sll1694	General secretion pathway protein G	113,079	34,52	42	5	17574	NO_3^-	10; 12; 11; 2; 9; 1; 8; 7; 5; 3
Slr1986	Allophycocyanin beta chain	86,855	26,71	4	4	17216	NO_3^-	9; 12; 11

It was identified a total of 29 proteins in the 12 gel pieces from *Synechocystis* sp. PCC 6803 grown in nitrate (Table 3).

These tables list the first comprehensive survey ever performed of the exoproteome of two morphologically and metabolically different cyanobacteria. Despite the large number of proteins identified, one cannot assume that all exoproteins have been identified, not only due to the disadvantages presented by the technical approach and method in use, but also because a limited number of growth conditions were tested. Eventually, under other growth conditions, one can expect to identify a few extra exoproteins. Nevertheless, since the conditions tested here

are the main standard growth conditions to cultivate *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120, one can confidently assume that the identified proteins represent the exoproteome core. In spite of the fact that some exoproteome proteins could have been missed it is also possible that some of the proteins identified here, especially the ones with lower score, can be contaminants from the cells due to the incapacity of the technique. The total number of proteins identified in *Anabaena* sp. PCC 7120 was very high when compared to other studies of cyanobacterial exoproteome [12, 21, 25]. This is probably caused by the combination of three factors; that there was never a deep analysis of the exoproteome before, that *Anabaena* sp. PCC 7120 is more prone to membrane shear and periplasmic or cytosol leakage and the fact that *Anabaena* sp. PCC 7120 can itself secrete more proteins that the species studied before.

In *Anabaena* sp. PCC 7120 ‘s exoproteome it was identified several proteins homologous to the ones identified in *Nostoc commune* like the case of SOD(alr2938); phycobilisome linker (alr0530) and Catalase (alr3090) [12]. Some of the proteins identified with the highest score like phytase (Alr4238), Glycerophosphoryl diester phosphodiesterase (All0275), Alkaline phosphatase (All2843) Oligopeptidase A (Alr0880) and Proteases (Alr0996; Alr1381) were related to the digestion of nutrients this suggest that these proteins can be secreted with the function of breaking down nutrients present in the medium to later absorb them. It was also found proteins involved in the binding and transportation of nutrients like the Nitrate transport protein NrtA (Alr0608) and Bicarbonate-binding protein (Alr2877).

In this work it was only identified the proteins PilA2 (Sll1694) and the Slr1855 (with low score) in common with previous study in *Synechocystis* sp. PCC 6803 [21]. However it was identified several new exoproteins including the Hemolysins (Slr1855; Sll1951) which were predicted to be secreted. Similarly to what happened in *Anabaena* sp. PCC 7120 it was identified proteins related to the digestion of nutrients, like the Carboxyl-terminal protease (Slr1751), a putative Polysaccharide deacetylase (sll1306) and Inorganic pyrophosphatase (Slr1622) as well as proteins related to binding and transportation of nutrients like the Iron uptake protein A2 Futa2 (Slr0513), the Nitrate transport protein NrtA (Sll1450) and the Bicarbonate-binding protein CmpA (Slr0040). From the 8 proteins that contain the SLH (S-layer homology) domain only the Slr1841 was identified. This is probably due to the fact that although all the other proteins of the S-layer are extracellular they are still bound to the cell and are not harvested in our methods with the exception of Slr1841 [72].

One way to increase the validation of the results is to compare each protein theoretical mass with the estimated mass of the gel band they was identified in. This analysis has the limitation of losing proteins that suffer protein cleavage or PTM’s that change their mass.

One protein that is a good example of this limitation is the Nitrate transport protein NrtA (alr0608 ~48kDa), the protein appears in the exoproteome of Anabaena sp. PCC 7120 grown both in nitrate and in ammonia. But while in the presence of ammonia and absence of nitrate the protein only is identified with a high score in a band close to its theoretical mass (band ANH4.6 ~42kDa, figure 9; Table 2), in the nitrate medium the protein was identified with high scores in two bands with different sizes; in band ANO3.5 (~40kDa) which is close to NrtA's expected size, and also in band ANO3.7 (~28kDa) which is smaller than the expected for this protein. In both bands the protein had very high score, 361.162 for band ANO3.5 and 288.499 in band ANO3.7 (table 2); both are higher than the score of NrtA identified in ammonia: 186.252. After analyzing the individual peptides identified in each band it was noticed that in the "lighter" band 7, the NrtA only had peptides identified after the aminoacid number 184. While the NrtA present in the "heavier" band 5 had a total of 15 sequences identified in the first 184 aminoacids plus the same peptides identified after the 184 aminoacid. The absence of a single peptide identified in the N-terminal region of the NrtA protein present in the band 7 suggests that it occurs some sort of processing with the excision of the N-terminal region when the cells grow in nitrate, since in the other two conditions there is no high score NrtA identified in this region (~28kDa). The NrtA is a nitrate transporter which is described to be anchored by the N-terminal region to the inner membrane facing the periplasm. Its function is suggested to be to scavenge the nitrate present in the media and transport it through a nitrate channel to the cytoplasm [73]. It is already interesting to think this protein is accumulated in the media when grown in nitrate and ammonia. But more intriguing is why, in nitrate, there is an accumulation of a truncated form of the protein, and if this version is important to scavenge nitrate.

Superoxide Dismutase and Catalase activities:

The highest score protein of the pieces AN2.12, ANO3.8 and ANH4.11 (table 2) was the Superoxide dismutase (FeSOD; alr2938). Although this protein has no signal peptide and the localization prediction tools [74], predict it as a periplasmic protein, it is clearly present in the extracellular milieu. Extracellular SOD has been described in *Nostoc commune* another filamentous cyanobacteria [12], and it has been postulated that its function is to prevent the oxidative stress caused by the production of superoxide by the heterocysts glycan's when exposed to the UV light [12]. This is supported by our results which identify SOD in all *Anabaena* sp. PCC 7120 samples, a cyanobacterium capable of forming heterocysts, but is not identified in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803's exoproteome.

The Mass Spectrometry results only state that the protein is present outside the cell, but they do not state whether it has activity or not. To test the extracellular protein activity, a SOD activity in-

gel assay was made with exoproteome samples. It was also made a control using Proteome extracts of the same cells from which the exoproteome was concentrated. And it was also added proteome and exoproteome samples from *Synechocystis* sp. PCC 6803 to validate the absence of SOD in the exoproteome MS analysis of this specie.

It was made 2 gels, one with normal SOD activity staining treatment and one incubated with 5mM H_2O_2 during the gel revelation. The H_2O_2 is used as a selective inhibitor of the FeSOD.

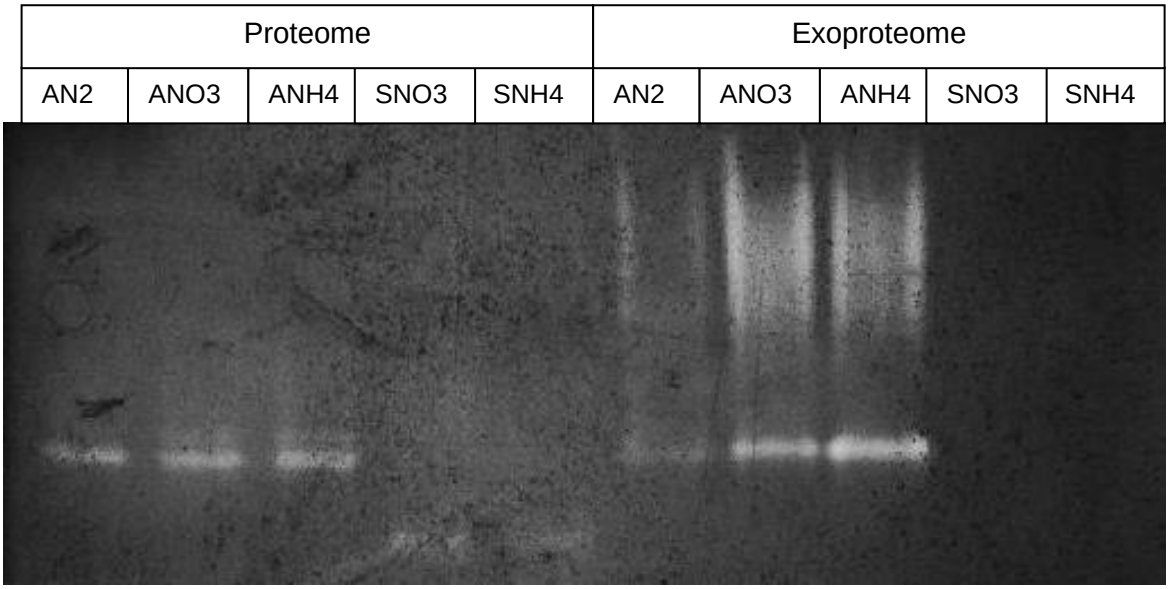


Fig 11- SOD activity gel. First five lanes are proteome extracts, *Anabaena* sp. PCC 7120 grown in $BG11_0$ (AN2); in BG11 (ANO3) and $BG11_0$+NH_4Cl (ANH4) and *Synechocystis* sp. PCC 6803 grown in BG11 (SNO3) and $BG11_0$+NH_4Cl (SNH4) in that order. Second 5 lanes are exoproteomes extracts of the same samples in the same order.

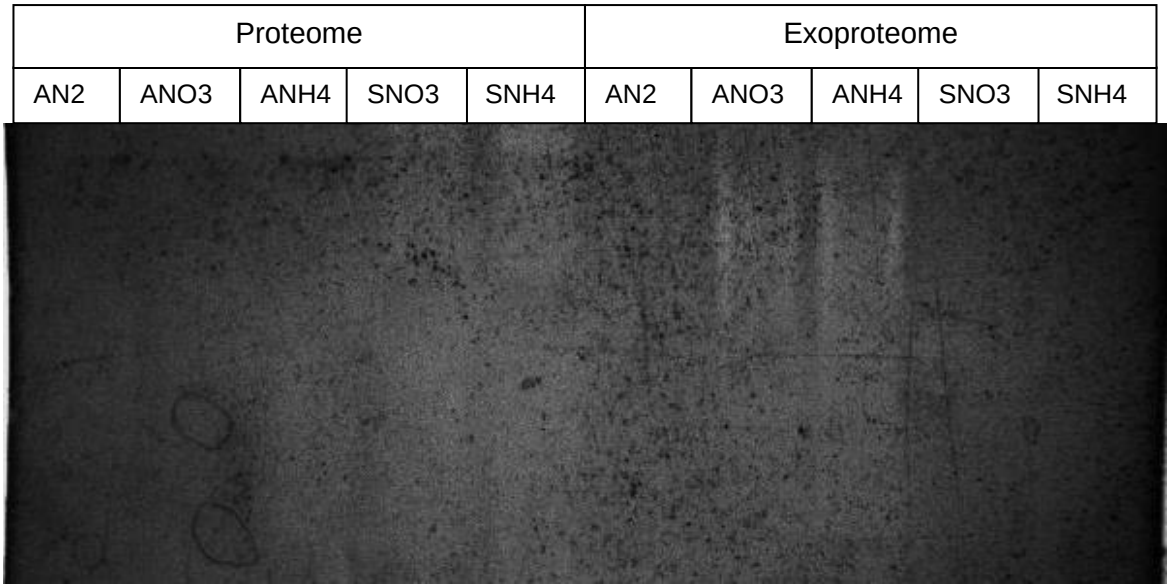


Fig 12- SOD activity gel treated with 5mM H_2O_2. First five lanes are proteome extracts, *Anabaena* sp. PCC 7120 grown in $BG11_0$ (AN2); in BG11 (ANO3) and $BG11_0$+NH_4Cl (ANH4) and *Synechocystis* sp. PCC 6803 grown in BG11 (SNO3) and $BG11_0$+NH_4Cl (SNH4) in that order. Second 5 lanes are exoproteomes extracts of the same samples in the same order.

In Figure 11 we can clearly see the presence of SOD activity in every *Anabaena* sp. PCC 7120 sample whether it is a proteome or an exoproteome extract. The exoproteome samples differ from the proteome ones, with the exoproteome having a blurred smear with SOD activity not present in the Proteome samples. In *Synechocystis* sp. PCC 6803 there was only detected some small activity in the proteome samples.

In Figure 12 we can see the bottom bands of the *Anabaena* sp. PCC 7120 samples disappear while the blurred smear persist although in a fainter way. The *Synechocystis* sp. PCC 6803 samples seem to have lost the SOD activity band present in Figure 11.

In *Anabaena* sp. PCC 7120 it was described that it has two SOD, a Fe one and a Mn one. Previous assays had come to the conclusion that the bottom band is the FeSOD, that in the top of the blur there is a MnSOD and the rest of the blur it's heterodimers with polipeptides from both FeSOD and MnSOD [75]. That explains how the bottom band disappears with the incubation of 5mM H_2O_2 and the blur remains, although appearing to have lost some activity.

Comparing the proteome and the exoproteome of the *Anabaena* sp. PCC 7120 samples it is quite surprising that only the FeSOD band appears on the proteome and both the FeSOD and the blur comprising of heterodimers of FeSOD and MnSOD appear on the exoproteome, when the mass spectrometry analysis only found the FeSOD on the exoproteome.

The reason why the MnSOD does not appear on the proteome can be related to the way the proteins are extracted, because the MnSOD is mainly membrane associated [75] it's possible that it is lost during the extraction. The reason why the blur comprising of heterodimers of FeSOD and MnSOD appears in the gel while the MnSOD was not identified on the MS analysis could be that although the MnSOD peptides are present in the media they could have characteristics that would make them hard to identify on the MS, like hydrophobicity or size of the peptides after the trypsin digestion. Also it was only analyzed sections of the SDS-PAGE gel with a strong coomassie stain. It is possible that the Mn-SOD is present in the gels but wasn't excised in the gel thus not being identified in the MS analysis while still being present in the extracellular media.

As for the *Synechocystis* sp. PCC 6803 results there was no SOD activity in the exoproteome samples, which correlates with the absence of SOD peptides identified in the MS analysis (Table 3) and they had SOD activity in the control proteome samples (figures 11). The SOD activity in the proteome samples is inhibited by the H_2O_2 which is expected since *Synechocystis* sp. PCC 6803 only has a FeSOD [76].

In *Anabaena* sp. PCC 7120 it was confirmed that although the FeSOD (alr2938) is predicted to be periplasmic it accumulates in the media and has activity. Although extracellular SOD has been described on *Nostoc commune* [12] it is the first time that it has been identified in *Anabaena* sp. PCC 7120 *exoproteome.*

In the ANH4.10 band (table 2) one of the proteins with a higher score is a catalase (Alr3090). As it was previously mentioned other enzymes involved in oxidative stress have been confirmed to be part of the exoproteome although they are not predicted to be exported. Like the FeSOD (alr2938) described above, the catalase (Alr3090) has no signal peptide and it is predicted to be cytoplasmic instead of extracellular [74]. In spite of this prediction it was found in the media of *Anabaena* sp. PCC 7120 grown in ammonia.

Similar to what was done for the FeSOD (alr2938), an activity gel for catalase was made to test if the protein is not only present in the media but has activity.

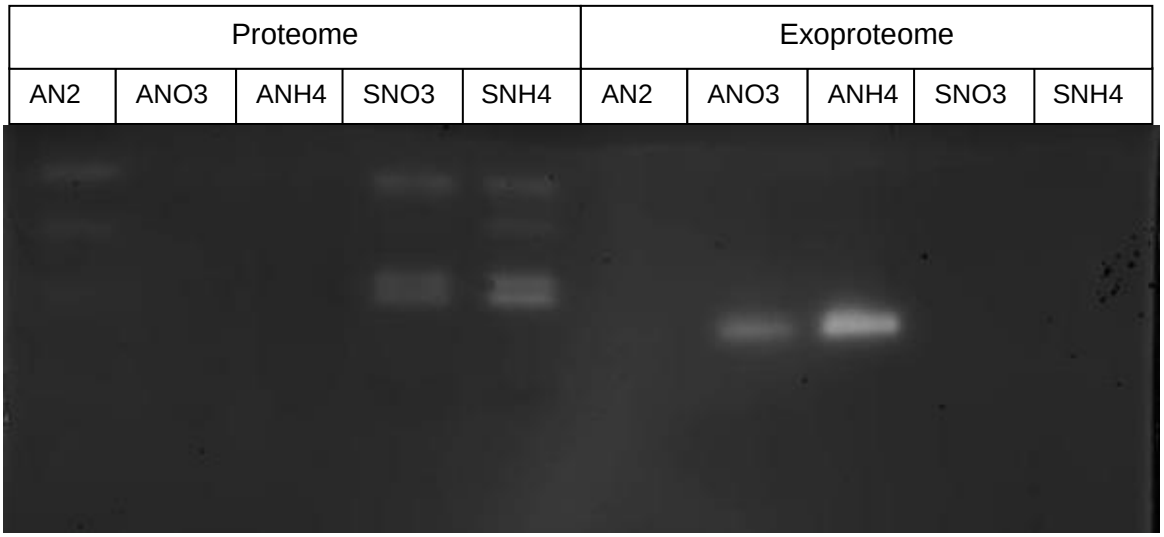


Fig 13- Catalase activity gel- . First five lanes are proteome extracts, *Anabaena* sp. PCC 7120 grown in $BG11_0$ (AN2); in BG11 (ANO3) and $BG11_0$+NH_4Cl (ANH4) and *Synechocystis* sp. PCC 6803 grown in BG11 (SNO3) and $BG11_0$+NH_4Cl (SNH4) in that order. Second 5 lanes are exoproteomes extracts of the same samples in the same order.

In figure 13 it can be noticed that in fact not only the catalase is present in the sample from *Anabaena* sp. PCC 7120 grown in ammonia, it has activity. Catalase activity is also observed in *Anabaena* sp. PCC 7120 grown in nitrate although it was not identified in MS analysis. This can be explained by the fact that the corresponding gel piece chosen in the sample grown in ammonia, the ANH4.10, was not chosen in the sample from *Anabaena* sp. PCC 7120 grown in nitrate (figure 9). So the protein might be there but it wasn't excised and analyzed in MS scan. There is no noticeable activity in the exoproteome from *Anabaena* sp. PCC 7120 without any combined nitrogen source. On top of that, the corresponding bands with the same mass as ANH4.10 were excised from the nitrogen fixing samples, the AN2.10 and AN2.11 bands (figure 9), and Catalase was not identified in either gel pieces. Both of these results suggest that Catalase is not being accumulated in the extracellular media when grown in $BG11_0$.

In the proteome samples it is interesting to notice that the samples which had catalase activity outside the cell (Anabaena sp. PCC 7120 grown in nitrate and ammonia) did not had Catalase activity in their proteome extracts while all the other proteome samples had activity. It is also interesting how the samples from *Anabaena* sp. PCC 7120 in nitrogen fixing condition had several Catalase bands all in different position when compared to the one found in the exoproteome samples of *Anabaena* sp. PCC 7120 grown in both nitrate and ammonia.

The *Synechocystis* sp. PCC 6803 exoproteome samples did not have any noticeable Catalase activity nor was any Catalase found in the MS analysis. The *Synechocystis* sp. PCC 6803

proteome samples had Catalase activity which was expected although they had several activity bands while only one Catalase-Peroxidase was identified in *Synechocystis* sp. PCC 6803 (sll1987). This several bands can be other catalase isoforms.

Cyanobacterial periplasmic containing GFP mutants

After the identification of the proteins that accumulate outside the cyanobacterial cell it was important to determine if the identified proteins are exclusively secreted by active mechanisms or if they could be accumulating outside the cell because of some sort of periplasmic leakage or cell lysis. For that purpose an experiment was designed based on the construction of cyanobacterial mutants, in each of which the reporter GFP is synthesized in the cytoplasm, but is sorted to the periplasm, where it accumulates as a soluble protein. This experiment will allow verifying if the GFP would be confined in the periplasm or if it would leak into the growth. Therefore, mutants expressing GFP fused with a periplasmic signal peptide were made for each cyanobacterium in study (for details see Figure 3).

Unfortunately it was only possible to obtain conjugants in *Anabaena* sp. PCC 7120, possibly because the pRL25C plasmid does not replicate in *Synechocystis* sp. PCC 6803.

After the identification of the exoproteome of *Synechocystis* sp. PCC 6803 it was noticed that one of the most abundant is the FutA2 (slr0513), which is the one the signal peptide was isolated from. This would invalidate the experience because the original protein itself is found in the media in contrary of what was initially thought.

To confirm if the mutants were expressing GFP and exporting it to the periplasm the cells were observed under the confocal microscope.

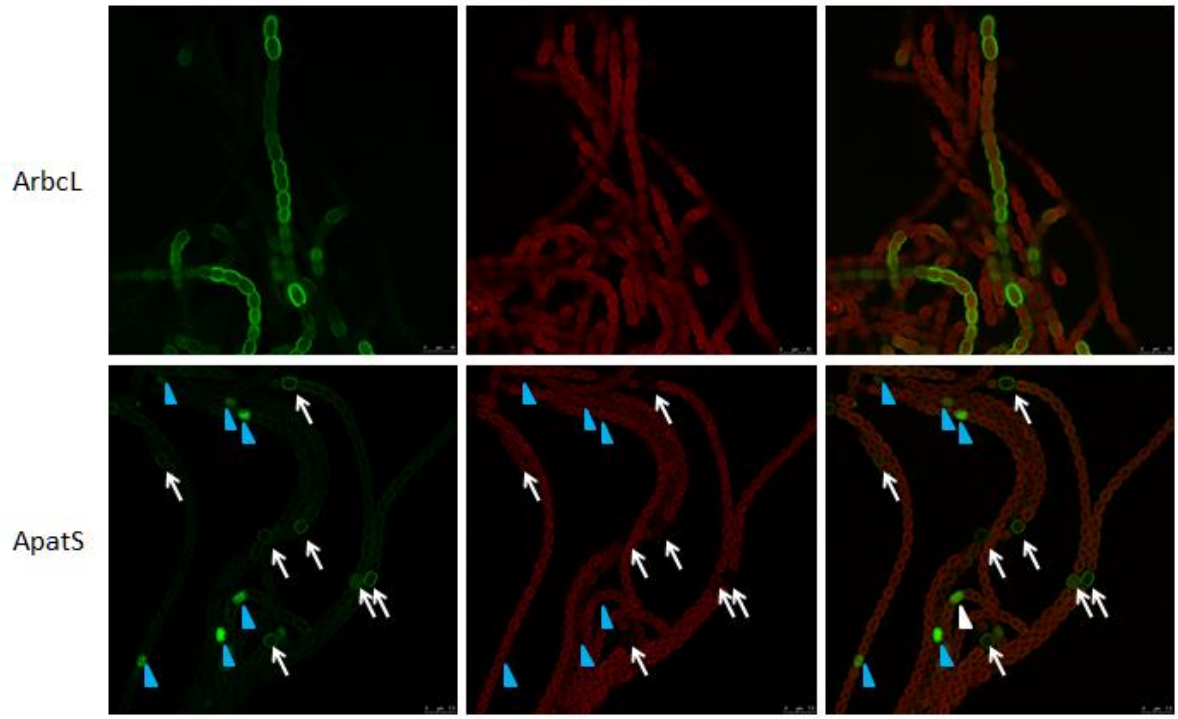


Fig 14- Confocal fluorescent microscopy of the *Anabaena* sp. PCC 7120 mutants. The left images are the green channel, the middle are the red channel and the right are the merge of the two channels. On top panel there is the ArbcL::ASP1::GFP (A) mutant grown in BG11 . On the bottom panel there is the ApatS::ASP1::GFP (A) mutant grown in $BG11_0$. Arrows point the heterocysts and blue arrowheads point to cells differentiating into heterocysts.

It is possible to confirm that both *Anabaena* sp. PCC 7120 mutants were expressing the GFP and it is localized in the periplasm. In the mutants with the rbcL promoter the GFP is expressed in vegetative cells (figure 14) and the mutants with patS promoter are expressing GFP only in heterocysts as it was expected (arrows, figure 14). The mutants with patS promoter have GFP in the periplasm of vegetative cells nearby because the periplasm is continuous and dynamic throughout the filament as described in [9]. Some cells (blue arrowheads, figure 14) appear to have GFP in their cytoplasm that is probably because they are being differentiated into heterocysts and the GFP is still being expressed in the cytosol and exported to the periplasm.

ΔTolC mutant generation

As it was previously mentioned the TolC-like proteins are outer membrane pores of a three protein efflux pump complex. They are important in the secretion of proteins to the medium. To identify which proteins were actively secreted by this complex transporter in *Anabaena* sp. PCC 7120 and *Synechocystis* sp. PCC 6803 TolC deletion mutants were generated (Δslr1270 for *Synechocystis* sp. PCC 6803 and Δalr2887 for *Anabaena* sp. PCC 7120). It is expected, that these mutants exihbit some differences in the exoproteome content when compared to the wild-

type. These differences should be the result of the lack of secretion of the proteins which are, in normal conditions, secreted by the type I secretion system.

When a high concentration of Kanamycin was reached, it was made a colony PCR, using the 5´FWD and 3´REV primers (table 1), to verify the presence of the wild-type copy of the slr1270 in the chromosomes.

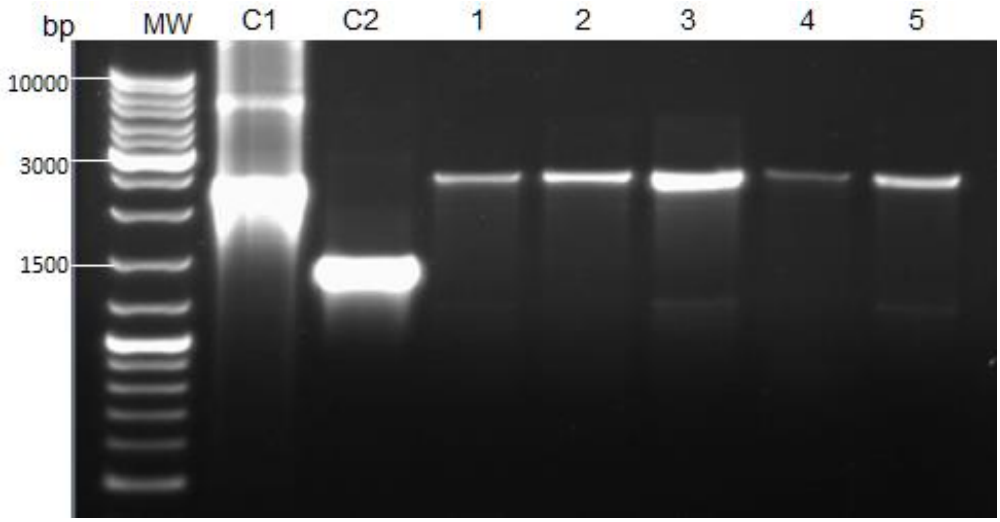


Fig 15- DNA electrophoresis with the results from the colony PCR of the *Synechocystis* sp. PCC 6803 Δslr1270 mutants. MW- GeneRuler™ 100-10000bp. C1 is the result of the colony PCR using as template the pSK+ with 5´ fragment+Kan^r+3´Fragment. C2 is the result of the colony PCR using as template the DNA of wild-type *Synechocystis* sp. PCC 6803. 1-5 are the result of the colony PCR for the 5 *Synechocystis* sp. PCC 6803 Δslr1270 mutant colonies selected.

Each colony tested had an amplicon with a size of approximately 2300bp, which is the expected size of the 5´fragment+Kan^r+3´Fragment, and the same size as the C1 control although the colonies 3 and 5 appear to have a faint band with the C2 control size (Figure 15).

This suggest the colonies 1,2 and 4 could have the mutant contruct totally segregated in all the chromosomes and there is no copy left of the wild-type slr1270 while the mutants 3 and 5 still have wild-type copies of the slr1270.

Conclusions

- It was identified a total of 46 extracellular proteins in Anabaena sp. PCC 7120 grown without any combined nitrogen source, 59 exoproteins when grown in nitrate and 80 exoproteins when grown in ammonia.
- It was identified 29 proteins in the extracellular milieu of *Synechocystis* sp. PCC6803
- Two proteins related with oxidative stress that were identified for *Anabaena* sp. PCC 7120 grown in nitrate and ammonia were further tested to verify if they were active in the extracellular medium. Concluding that not only both proteins were active in nitrate and ammonia exoproteomes, SOD was also active in the exoproteome of Anabaena sp. PCC 7120 grown without any combined nitrogen source.
- An *Anabaena* sp. PCC 7120 mutant expressing GFP directed to the periplasm was generated.
- A *Synechocystis* sp. PCC 6803 *tolC* deletion mutant was also generated

Future Perspectives

For better knowledge of the secretion system of both cyanobacteria it is important to understand more about periplasmic and cytosol leakage, thus understanding the dynamics of the membranes integrity in both species. Although the GFP expressing mutants in *Anabaena* sp. PCC 7120 are already generated, they have not yet been tested for GFP leakage. For that purpose it is important to growth the mutants in the same conditions as the wild-type and determine the presence of GFP in the exoproteome with a Western-blot technique. The next step would be to make *Synechocystis* sp. PCC 6803 mutants similarly to what was made in *Anabaena* sp. PCC 7120 but using a different signal peptide from the one chosen previously since the signal peptide chosen is from a protein (FutA2) identified in the exoproteome. This would make the results of membrane integrity inconclusive since the signal peptide could be responsible for GFP secretion. Because it is common for secreted proteins to first be exported in the Sec system it would be advisable to make mutants exporting the GFP to the periplasm through other pathways, like the twin-arginine. A control mutant with GFP without any signal peptide should also be made in both cyanobacteria to determine if there is cell lysis and contamination of the medium with cytosolic proteins. With the results of this new set of mutants it would be possible to determine if there are non-secreted proteins being leaked to the medium and if so are they exclusively periplasmic or cytosol proteins as well.

After the creation of the exoproteome lists it is now important to make bioinformatic analysis of the proteins to try and identify which are in fact programed for secretion and which are not expected to be secreted. It is also important to do a search in N-terminal and C-terminal sequences of the proteins to search for common patterns and determine the leader peptides that address the proteins for secretion. This can be very challenging since there are plenty of different secretion systems which should each have their own leader peptide. It also would be very helpful to categorize the different proteins identified according to their function, this is especially challenging because many of the proteins have unknown function.

It is necessary to confirm the full segregation of the *Synechocystis* sp. PCC 6803 ΔtolC mutant by Southern blot.

After confirming the *tolC* deletion the mutants will be grown in conditions similar to the wild-type and their exoproteome will be compared to the wild-type. It is expected some bands to absent in the mutant exoproteome, possibly because the main proteins in those bands are secreted in type-I secretion mechanism. Because the wild-type bands were identified in MS-MS it will be possible to identify type-I secretion substrates.

Later on the project it would be very interesting to do new deletion mutants of key proteins of other secretion systems identified in both cyanobacteria to determine the substrates of every secretion method in a way similar to what was described for the type-I secretion.

References

1. Schopf, J.W., *The Fossil Record: Tracing the Roots of the Cyanobacterial Lineage*, in *The Ecology of Cyanobacteria*, B. Whitton and M. Potts, Editors. 2002, Springer Netherlands. p. 13-35.
2. Stanier, R.R.J.D.J.B.W.M.H.a.R.Y., *Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria.* Journal of General Microbiology, 1979. **111**: p. 1-61.
3. Grigorieva, G. and S. Shestakov, *TRANSFORMATION IN THE CYANOBACTERIUM SYNECHOCYSTIS SP 6803.* FEMS Microbiol Lett, 1982. **13**(4): p. 367-370.
4. Ikeuchi, M., *Complete genome sequence of a cyanobacterium Synechocystis sp. PCC 6803, the oxygenic photosynthetic prokaryote.* Tanpakushitsu Kakusan Koso, 1996. **41**(16): p. 2579-83.
5. Kaneko, T. and S. Tabata, *Complete genome structure of the unicellular cyanobacterium Synechocystis sp. PCC6803.* Plant and Cell Physiology, 1997. **38**(11): p. 1171-1176.
6. Kaneko, T., et al., *Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium anabaena sp strain PCC 7120.* DNA Research, 2001. **8**(5): p. 205-213.
7. Ohmori, M., et al., *Characterization of genes encoding multi-domain proteins in the genome of the filamentous nitrogen-fixing Cyanobacterium anabaena sp strain PCC 7120.* DNA Research, 2001. **8**(6): p. 271-284.
8. Mariscal, V. and E. Flores, *Multicellularity in a heterocyst-forming cyanobacterium: pathways for intercellular communication.* Adv Exp Med Biol, 2010. **675**: p. 123-35.
9. Mariscal, V., A. Herrero, and E. Flores, *Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium.* Mol Microbiol, 2007. **65**(4): p. 1139-45.
10. Šmarda, J., et al., *S-layers on cell walls of cyanobacteria.* Micron, 2002. **33**(3): p. 257-277.
11. Engelhardt, H., *Are S-layers exoskeletons? The basic function of protein surface layers revisited.* Journal of Structural Biology, 2007. **160**(2): p. 115-124.
12. Helm, R.P., M., *Ecology of Cyanobacteria II: Their Diversity in Space and Time (B.A. Whitton, Ed.)*, S.S.B.M. B.V., Editor 2012. p. 461-480.
13. Rakleova, G., et al., *DIFFERENTIALLY SECRETED PROTEINS OF ANTARCTIC AND MESOPHILIC STRAINS OF SYNECHOCYSTIS SALINA AND CHLORELLA VULGARIS AFTER UV-B AND TEMPERATURE STRESS TREATMENT.* Biotechnology & Biotechnological Equipment, 2013. **27**(2): p. 3669-3680.
14. Wegener, K.M., et al., *Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations.* Mol Cell Proteomics, 2010. **9**(12): p. 2678-89.
15. Fulda, S., et al., *Proteomics of Synechocystis sp. strain PCC 6803. Identification of periplasmic proteins in cells grown at low and high salt concentrations.* Eur J Biochem, 2000. **267**(19): p. 5900-7.
16. Huang, F., et al., *Proteomics of Synechocystis sp. strain PCC 6803: identification of plasma membrane proteins.* Mol Cell Proteomics, 2002. **1**(12): p. 956-66.
17. Moslavac, S., et al., *Proteomic analysis of the outer membrane of Anabaena sp. strain PCC 7120.* J Proteome Res, 2005. **4**(4): p. 1330-8.
18. Pisareva, T., et al., *Proteomics of Synechocystis sp. PCC 6803. Identification of novel integral plasma membrane proteins.* FEBS J, 2007. **274**(3): p. 791-804.
19. Rajalahti, T., et al., *Proteins in different Synechocystis compartments have distinguishing N-terminal features: a combined proteomics and multivariate sequence analysis.* J Proteome Res, 2007. **6**(7): p. 2420-34.
20. Srivastava, R., T. Pisareva, and B. Norling, *Proteomic studies of the thylakoid membrane of Synechocystis sp. PCC 6803.* Proteomics, 2005. **5**(18): p. 4905-16.
21. Sergeyenko, T.V. and D.A. Los, *Identification of secreted proteins of the cyanobacterium Synechocystis sp strain PCC 6803.* FEMS Microbiol Lett, 2000. **193**(2): p. 213-216.

22. Sergeyenko, T.V. and D.A. Los, *The effect of various stresses on the expression of genes encoding the secreted proteins of the cyanobacterium Synechocystis sp PCC 6803.* Russian Journal of Plant Physiology, 2002. **49**(5): p. 650-656.
23. Sergeyenko, T.V. and D.A. Los, *Cyanobacterial leader peptides for protein secretion.* FEMS Microbiol Lett, 2003. **218**(2): p. 351-357.
24. Wright, D.J., et al., *UV irradiation and desiccation modulate the three-dimensional extracellular matrix of Nostoc commune (Cyanobacteria).* J Biol Chem, 2005. **280**(48): p. 40271-81.
25. Hill, D.R., et al., *Water stress proteins of Nostoc commune (Cyanobacteria) are secreted with UV-A/B-absorbing pigments and associate with 1,4-beta-D-xylanxylanohydrolase activity.* J Biol Chem, 1994. **269**(10): p. 7726-34.
26. Sakiyama, T., et al., *Functions of a hemolysin-like protein in the cyanobacterium Synechocystis sp PCC 6803.* Arch Microbiol, 2011. **193**(8): p. 565-571.
27. Desvaux, M., et al., *Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue.* Trends Microbiol, 2009. **17**(4): p. 139-45.
28. Thanassi, D.G. and S.J. Hultgren, *Multiple pathways allow protein secretion across the bacterial outer membrane.* Curr Opin Cell Biol, 2000. **12**(4): p. 420-30.
29. Koster, M., W. Bitter, and J. Tommassen, *Protein secretion mechanisms in Gram-negative bacteria.* International Journal of Medical Microbiology, 2000. **290**(4–5): p. 325-331.
30. Lory, S., *Secretion of proteins and assembly of bacterial surface organelles: shared pathways of extracellular protein targeting.* Curr Opin Microbiol, 1998. **1**(1): p. 27-35.
31. Salmond, G.P.C. and P.J. Reeves, *Membrance traffic wardens and protein secretion in Gram-negative bacteria.* Trends Biochem Sci, 1993. **18**(1): p. 7-12.
32. Pugsley, A.P., *The complete general secretory pathway in gram-negative bacteria.* Microbiol Rev, 1993. **57**(1): p. 50-108.
33. Economou, A., *Following the leader: bacterial protein export through the Sec pathway.* Trends Microbiol, 1999. **7**(8): p. 315-320.
34. Barnett, J.P., et al., *The Tat protein export pathway and its role in cyanobacterial metalloprotein biosynthesis.* FEMS Microbiol Lett, 2011. **325**(1): p. 1-9.
35. Saier, M.H., *Protein secretion and membrane insertion systems in gram-negative bacteria.* Journal of Membrane Biology, 2006. **214**(1-2): p. 75-90.
36. Henderson, I.R., F. Navarro-Garcia, and J.P. Nataro, *The great escape: structure and function of the autotransporter proteins.* Trends Microbiol, 1998. **6**(9): p. 370-378.
37. Klauser, T., J. Pohlner, and T.F. Meyer, *Extracellular transport of cholera toxin B subunit using Neisseria IgA protease beta-domain: conformation-dependent outer membrane translocation.* EMBO J, 1990. **9**(6): p. 1991-9.
38. Suzuki, T., M.C. Lett, and C. Sasakawa, *Extracellular transport of VirG protein in Shigella.* J Biol Chem, 1995. **270**(52): p. 30874-80.
39. Reumann, S., J. Davila-Aponte, and K. Keegstra, *The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog.* Proc Natl Acad Sci U S A, 1999. **96**(2): p. 784-9.
40. Ertel, F., et al., *The evolutionarily related beta-barrel polypeptide transporters from Pisum sativum and Nostoc PCC7120 contain two distinct functional domains.* J Biol Chem, 2005. **280**(31): p. 28281-9.
41. Delepelaire, P., *Type I secretion in gram-negative bacteria.* Biochimica Et Biophysica Acta-Molecular Cell Research, 2004. **1694**(1-3): p. 149-161.
42. Kanonenberg, K., C.K.W. Schwarz, and L. Schmitt, *Type I secretion systems - a story of appendices.* Res Microbiol, 2013. **164**(6): p. 596-604.
43. Binet, R., et al., *Protein secretion by Gram-negative bacterial ABC exporters – a review.* Gene, 1997. **192**(1): p. 7-11.

44. Koronakis, V., J. Eswaran, and C. Hughes, *Structure and function of TolC: the bacterial exit duct for proteins and drugs.* Annu Rev Biochem, 2004. **73**: p. 467-89.
45. Andersen, C., et al., *Transition to the open state of the TolC periplasmic tunnel entrance.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11103-8.
46. Campos, M., et al., *The type II secretion system - a dynamic fiber assembly nanomachine.* Res Microbiol, 2013. **164**(6): p. 545-555.
47. Nunn, D., *Bacterial Type II protein export and pilus biogenesis: more than just homologies?* Trends Cell Biol, 1999. **9**(10): p. 402-408.
48. Bakkali, M., *Could DNA uptake be a side effect of bacterial adhesion and twitching motility?* Arch Microbiol, 2013. **195**(4): p. 279-289.
49. Bhaya, D., et al., *Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp. PCC6803.* Mol Microbiol, 2000. **37**(4): p. 941-51.
50. Coburn, B., I. Sekirov, and B.B. Finlay, *Type III secretion systems and disease.* Clin Microbiol Rev, 2007. **20**(4): p. 535-49.
51. Fronzes, R., P.J. Christie, and G. Waksman, *The structural biology of type IV secretion systems.* Nat Rev Microbiol, 2009. **7**(10): p. 703-14.
52. Alvarez-Martinez, C.E. and P.J. Christie, *Biological diversity of prokaryotic type IV secretion systems.* Microbiol Mol Biol Rev, 2009. **73**(4): p. 775-808.
53. Glazer, A.N., *Light harvesting by phycobilisomes.* Annu Rev Biophys Biophys Chem, 1985. **14**: p. 47-77.
54. Sergeyenko, T.V. and D.A. Los, *Identification of secreted proteins of the cyanobacterium Synechocystis sp. strain PCC 6803.* FEMS Microbiol Lett, 2000. **193**(2): p. 213-6.
55. Sakiyama, T., et al., *Purification and characterization of a hemolysin-like protein, Sll1951, a nontoxic member of the RTX protein family from the Cyanobacterium Synechocystis sp. strain PCC 6803.* J Bacteriol, 2006. **188**(10): p. 3535-42.
56. Sakiyama, T., et al., *Functions of a hemolysin-like protein in the cyanobacterium Synechocystis sp. PCC 6803.* Arch Microbiol, 2011. **193**(8): p. 565-71.
57. Stanier, R.Y., et al., *Purification and properties of unicellular blue-green algae (order Chroococcales).* Bacteriol Rev, 1971. **35**(2): p. 171-205.
58. Moslavac, S., *Outer membrane proteins of Anabaena sp. strain PCC 7120*

in *Fakultät für Biologie*2007, Ludwig-Maximilians-Universität München: Munich.
59. Hanahan, D., *Studies on transformation of Escherichia coli with plasmids.* J Mol Biol, 1983. **166**(4): p. 557-80.
60. Addison, C.J., S.H. Chu, and R.N. Reusch, *Polyhydroxybutyrate-enhanced transformation of log-phase Escherichia coli.* Biotechniques, 2004. **37**(3): p. 376-8, 380, 382.
61. Williams, J.G.K., *CONSTRUCTION OF SPECIFIC MUTATIONS IN PHOTOSYSTEM-II PHOTOSYNTHETIC REACTION CENTER BY GENETIC-ENGINEERING METHODS IN SYNECHOCYSTIS-6803.* Methods Enzymol, 1988. **167**: p. 766-778.
62. Elhai, J. and C.P. Wolk, *Conjugal transfer of DNA to cyanobacteria.* Methods Enzymol, 1988. **167**: p. 747-54.
63. Raymond, S. and L. Weintraub, *Acrylamide gel as a supporting medium for zone electrophoresis.* Science, 1959. **130**(3377): p. 711.
64. Raghavan, P.S., H. Rajaram, and S.K. Apte, *Nitrogen status dependent oxidative stress tolerance conferred by overexpression of MnSOD and FeSOD proteins in Anabaena sp. strain PCC7120.* Plant Mol Biol, 2011. **77**(4-5): p. 407-17.
65. Beauchamp, C. and I. Fridovich, *Superoxide dismutase: improved assays and an assay applicable to acrylamide gels.* Anal Biochem, 1971. **44**(1): p. 276-87.
66. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry.* Nat Methods, 2007. **4**(3): p. 207-14.

67. Tamagnini, P., et al., *Hydrogenases in Nostoc sp. Strain PCC 73102, a Strain Lacking a Bidirectional Enzyme.* Appl Environ Microbiol, 1997. **63**(5): p. 1801-7.
68. Brody, J.R. and S.E. Kern, *Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis.* Biotechniques, 2004. **36**(2): p. 214-6.
69. Huang, H.H., et al., *Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology.* Nucleic Acids Res, 2010. **38**(8): p. 2577-93.
70. Yoon, H.S. and J.W. Golden, *Heterocyst pattern formation controlled by a diffusible peptide.* Science, 1998. **282**(5390): p. 935-8.
71. Badarau, A., et al., *FutA2 is a ferric binding protein from Synechocystis PCC 6803.* J Biol Chem, 2008. **283**(18): p. 12520-7.
72. Smarda, J., et al., *S-layers on cell walls of cyanobacteria.* Micron, 2002. **33**(3): p. 257-77.
73. Frias, J.E., E. Flores, and A. Herrero, *Nitrate assimilation gene cluster from the heterocyst-forming cyanobacterium Anabaena sp. strain PCC 7120.* J Bacteriol, 1997. **179**(2): p. 477-86.
74. Yu, N.Y., et al., *PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes.* Bioinformatics, 2010. **26**(13): p. 1608-15.
75. Raghavan, P.S., H. Rajaram, and S.K. Apte, *N-terminal processing of membrane-targeted MnSOD and formation of multiple active superoxide dismutase dimers in the nitrogen-fixing cyanobacterium Anabaena sp. strain PCC7120.* FEBS J, 2013. **280**(19): p. 4827-38.
76. Bhattacharya, J., et al., *Synechocystis Fe superoxide dismutase gene confers oxidative stress tolerance to Escherichia coli.* Biochem Biophys Res Commun, 2004. **316**(2): p. 540-4.