U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Elucidation of the molecular mechanisms underlying the production of cyanobacterial extracellular polymeric substances (EPS)

Maksym Povkhanych

Dissertação de Mestrado em Aplicações em Biotecnologia e Biologia Sintética apresentada à Faculdade de Ciências da Universidade do Porto

Departamento de Biologia; Departamento de Química e Bioquímica

2020/2021

Elucidation of the molecular mechanisms underlying the production of cyanobacterial extracellular polymeric substances (EPS)

Maksym Povkhanych

Dissertação de Mestrado em Aplicações em Biotecnologia e Biologia Sintética apresentada à Faculdade de Ciências da Universidade do Porto

Departamento de Biologia; Departamento de Química e Bioquímica

2020/2021

Orientador

Doutora Sara Bernardes Pereira, Investigadora Júnior no grupo de Bioengenharia e Microbiologia Sintética,
i3S - Instituto de Investigação e Inovação em Saúde,
IBMC - Instituto de Biologia Molecular e Celular,
Universidade do Porto

Coorientador

Professora Doutora Paula Tamagnini, Professora Associada da Faculdade de Ciências da Universidade do Porto, Departamento de Biologia. Líder do grupo de Bioengenharia e Microbiologia Sintética,
i3S - Instituto de Investigação e Inovação em Saúde,
IBMC - Instituto de Biologia Molecular e Celular,
Universidade do Porto

Todas as correções determinadas

pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

“To attain knowledge, add things every day.

To attain wisdom, remove things

Acknowledgments

Estes últimos dois anos simbolizam um grande momento de aprendizagem para mim, tanto durante o desenvolvimento deste projeto/dissertação como a nível pessoal. Qualquer experiência da nossa vida passa por altos e baixos o que nos desafia e permite oportunidades de aprendizagem. Claro que durante estes momentos nunca estamos sozinhos e é muito importante reconhecer o papel que outras pessoas tomam no nosso percurso.

Começo por agradecer à minha orientadora Doutora Sara Pereira que sem dúvida alguma tornou este projeto e trabalho possível. Foram muitas as lições e o conhecimento partilhado durante este último ano de dissertação, sei que aprendi muito e ainda tenho muito para aprender, mas sem a tua ajuda não teria conseguido. Sempre foi mostrada disponibilidade e paciência da tua parte, principalmente para lidar com a minha escrita e sinto-me muito grato por ter tido a tua orientação. Por todo o tempo e apoio, um grande obrigado!

Agradeço também a minha coorientadora Professor Doutora Paula Tamagnini e a todos os membros do grupo *Bioengineering and Synthetic Microbiology* por qualquer esclarecimento de dúvidas e disponibilidade prestada bem como pela oportunidade de trabalhar neste projeto, espero ter contribuído. Com especial menção dos meus coleguinhas Ana, Bia, Delfim e Jéssica. Todos passámos por um ano semelhante e sei que concordariam comigo que é sempre bom ter alguém com quem se podem relacionar.

Apesar de pouco ou nada perceberem do que estudo, também agradeço os meus pais Yevgen e Iryna por sempre mostrarem apoio durante os últimos anos, sem este apoio nada seria possível. Não me posso esquecer de referir a minha irmã Eva que também esteve presente durante todo este percurso e que vai iniciar o seu brevemente. Tudo seria muito mais tedioso sem todas as piadas secas, partilha de histórias e conflitos de irmãos que temos, obrigado!

Muitas amizades foram criadas nestes dois anos e quero deixar um sincero obrigado a todos que fizeram parte do percurso. São demasiadas pessoas para poder enumerar aqui, mas também não é necessário, aqueles que me impactaram sabem disso e terão sempre um espaço na minha memória e coração! Queria, contudo, deixar uma mensagem mais especial às amizades mais antigas: Carolinas, Diogo, Mafalda, Sílvia e Vanessa agradeço-vos por todas as risadas, palhaçadas e aventuras pelas quais passámos. Estou muito grato por todos estes anos de amizade e espero que continue assim para muitos mais! Um enorme obrigado a todos!

This work was financed by FEDER- Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020- Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT- Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior in the framework of the project POCI-01-0145-FEDER-028779.

Abstract

Many bacteria produce extracellular polymeric substances (EPS), mainly composed of heteropolysaccharides, that fulfil different biological roles depending on the organism and its environment. Due to their distinctive properties, the EPS produced by cyanobacteria are a promising resource for biotechnological and biomedical applications. However, it is necessary to unveil their biosynthetic pathway(s) to optimize their production. Previously, it was shown that cyanobacteria harbour genes encoding proteins related to the three main bacterial pathways of assembly and export of EPS, but often the complete set of proteins from a single pathway is absent. Mutational analysis using the model cyanobacterium *Synechocystis* sp. PCC 6803 confirmed the involvement of proteins from different pathways in the production of EPS, including the ABC transporter component Slr0977 (KpsM) and the polysaccharide co-polymerase Sll0923 (Wzc). Together, these results suggest that the production of cyanobacterial EPS may not strictly follow one of the pathways previously characterized for other bacteria and raises the hypothesis of a crosstalk between different pathways. To clarify this hypothesis, it is important to continue elucidating the mechanisms of EPS production by identifying the functional partners of proteins related to different pathways.

The aim of this work is to generate the molecular tools (plasmids) and develop the methodology to further characterize *Synechocystis* Wzc and KpsM, including the identification of their functional partners and/or *in vivo* characterization. For that, and moving from previous results, pull-down assays were performed using purified His6-Wzc as bait and *Synechocystis* protein extracts as target. The isolated proteins are being identified by mass spectrometry. In addition, to elucidate the proteins that interact with KpsM, an *in vivo* chemical crosslink followed by pull-down assays will be performed. For that, the sequences encoding the tagged version of the protein were cloned in a replicative vector under the regulation of different promoters. For one of the plasmids, the sequence was confirmed, and the plasmid will be used to transform the *Synechocystis kpsM* knockout mutant. In addition, the generation of plasmids to complement an *Escherichia coli* acapsular mutant with *Synechocystis* Wzc or its truncated version (without the C-terminal domain where its phosphorylation/dephosphorylation occurs) and evaluate its ability to restore capsule formation was initiated, with the final goal of clarifying the role of Wzc in an *in vivo* system.

Resumo

Várias bactérias produzem substâncias poliméricas extracelulares (EPS), compostas principalmente por heteropolissacarídeos que cumprem diferentes funções biológicas dependendo do organismo e do seu meio ambiente. Devido às suas propriedades distintas, os EPS produzidos por cianobactérias são um recurso promissor para aplicações biotecnológicas e biomédicas. No entanto, é necessário elucidar a(s) sua(s) via(s) biossintética(s) para otimizar a sua produção. Foi demonstrado previamente que as cianobactérias contêm genes que codificam proteínas relacionadas com as três principais vias bacterianas de montagem e exportação de EPS; contudo, o conjunto completo de proteínas de uma única via está frequentemente ausente. A geração de mutantes utilizando a cianobactéria modelo *Synechocystis* sp. PCC 6803 confirmou o envolvimento de proteínas de diferentes vias na produção de EPS, incluindo o componente transportador ABC Slr0977 (KpsM) e a proteína da família "polysaccharide co-polimerase" Sll0923 (Wzc). Em conjunto, estes resultados sugerem que a produção de EPS em cianobactérias pode não seguir estritamente uma das vias previamente caracterizadas para outras bactérias e levanta a hipótese de haver um cruzamento (ou *crosstalk*) entre diferentes vias. Para esclarecer esta hipótese, é importante continuar a elucidar os mecanismos de produção de EPS, identificando os parceiros funcionais das proteínas relacionadas com as diferentes vias.

O objetivo deste trabalho é gerar as ferramentas moleculares (plasmídeos) e desenvolver a metodologia para posterior caracterização do Wzc e KpsM de *Synechocystis*, incluindo a identificação dos seus parceiros funcionais. Para tal, e partindo de resultados anteriores, foram realizados ensaios de "*pull-down"* utilizando a proteína His6-Wzc purificado como "*bait"* e extratos proteicos de *Synechocystis* como "*target"*. As proteínas isoladas estão a ser identificadas por espetrometria de massa. Para elucidar as proteínas que interagem com o KpsM, serão realizados ensaios de "*crosslink*" *in vivo* seguidos de "*pull-down"*. Para tal, as sequências que codificam a versão da proteína com uma cauda de poli-histidina foram clonadas num vetor replicativo sob a regulação de diferentes promotores. Para um dos plasmídeos, a sequência foi confirmada e o plasmídeo será utilizado para transformar o mutante *Synechocystis kpsM*. Além desta metodologia, foi iniciada também a geração de plasmídeos para complementar um mutante acapsular de *E. coli* com o Wzc de *Synechocystis* ou a sua versão truncada (sem o domínio C-terminal onde ocorre a sua fosforilação/desfosforilação) e avaliada a capacidade do mutante em restaurar a formação da cápsula, clarificando assim o papel do Wzc num sistema *in vivo*.

Table of Contents

List of Figures:

List of Tables:

List of Abbreviations:

Amp – Ampicillin

ATP - Adenosine triphosphate

BCA - Bicinchoninic acid assay

Cm - Chloramphenicol

CPS – Capsular polysaccharides

DNA - Deoxyribonucleic acid

EDTA - Ethylenediamine tetraacetic acid

EPS – Extracellular polymeric substances

IMAC - Immobilized metal affinity chromatography

IPTG - Isopropyl ß-D-1-thiogalactopyranoside

Km – Kanamycin

LB - Lysogeny broth

MS – Mass Spectrometry

OFR - Open reading frame

ON – Over night

OPX - Outer membrane polysaccharide export

PCP - Polysaccharide co-polymerase

PCR – Polymerase chain reaction

PHB - Poly-β-hydroxybutyrate

PMSF - Phenylmethylsulphonyl fluoride

RBS – Ribosome binding site

RPS - Released polysaccharides

RT – Room temperature

SDS-PAGE - Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

SEC – Size exclusion chromatography

TAE - Tris-acetate-EDTA

UV – Ultraviolet light

1. Introduction

1.1. Cyanobacteria

Cyanobacteria are a phylum of phototrophic gram-negative prokaryotes, the only prokaryotic organisms performing oxygenic photosynthesis. They played a critical role in Earth's history as primary producers and in the progressive oxygenation of the atmosphere (Kasting, 2013). Today they are important components of many ecosystems, as they tolerate and thrive in a wide variety of environments, including marine, fresh water, terrestrial as well as extreme hypersaline, hot spring or cold conditions (Hammer, 1986; Rai et al., 2013; Rippka, 1988; Whitton & Potts, 2000; Zakhia et al., 2008).

Cyanobacteria show a great morphological diversity, presenting unicellular, colonial, and filamentous forms, varying in dimension, shape, colour and composition of cell contents (Komárek, 2013; Komárek, 1999). In addition, some strains are also able to differentiate cells specialized in different functions, such as heterocysts (N_2-fixation) and akinetes (Singh et al., 2020). Their primary and secondary metabolism can also differ from species to species, the latter being particularly interesting as numerous secondary metabolites are promising for biotechnology (Carmichael, 1992; Dittmann et al., 2013; Shah et al., 2017; Shih et al., 2013). Besides the vast range of compounds of interest produced by cyanobacteria, their basic and low-cost growth requirements, high growth rates compared to algae and plants and ease of strain manipulation using both physiological and genetic approaches (Beyl et al., 2019; Parikh & Madamwar, 2006) are all important factors to make cyanobacteria important cell factories for the production of compounds for biotechnological applications (Lau et al., 2015). However, despite all of these advantages, the number of cyanobacterial compounds currently used in biotechnological applications is still relatively small, since the production yields are insufficient to meet market demands.

Recent advances in DNA technology allowed the sequencing of the entire genomes of many cyanobacterial species, which may be used to explore the metabolic potential of the organisms (Kumar et al., 2019). This knowledge is important to allow genetic engineering approaches envisaging the increase of the productivity of compounds of interest by, e.g., reducing carbon input into competing pathways and/or limiting storage compounds (Hagemann & Hess, 2018; Savakis & Hellingwerf, 2015).

1.2. Extracellular Polymeric Substances (EPS)

Extracellular polymeric substances (EPS), also sometimes referred to as exopolymers, are mostly composed of polysaccharides. In addition, they also may contain small amounts of lipids, nucleic acids, peptides, and sulfated sugars (Salama et al., 2016), with the sulfate group being important for the tertiary structure and polysaccharide stability (Rai et al., 2013). The production of EPS is common in many bacterial species, being mainly produced during the exponential growth phase in many strains (Rai et al., 2013). The type of microorganism (Vaningelgem et al., 2004), nutrient availability (De Vuyst & Degeest, 1999; Myszka & Czaczyk, 2009; Petit et al., 2004), growth phase, and environmental conditions (Bahat-Samet et al., 2004; Kannenberg & Brewin, 1994) all influence the composition of bacterial EPS.

Bacterial EPS may serve several important roles, including defence against desiccation, scavenging of trace metals, protection from toxic contaminants (particularly metals)(De Philippis et al., 2011; García-Meza et al., 2005; Pereira et al., 2009), high temperature, UV damage and calcification of their environment (Li et al., 2017; Rai et al., 2013; Suresh Kumar et al., 2007).

1.3. Cyanobacterial EPS

The EPS of cyanobacteria can remain attached to the cell surface, being referred to as sheaths, capsules, or slimes depending on their thickness, consistency, and appearance (sometimes referred to as capsular polysaccharides - CPS) or be released to the extracellular medium (referred as released polysaccharides - RPS) (Pereira et al., 2009; Rossi & De Philippis, 2015). They are often composed of a variety of hexoses (fructose, galactose, glucose, and mannose), pentoses (arabinose, ribose, and xylose), and deoxyhexoses (fucose and rhamnose) (Kehr & Dittmann, 2015; Rossi & De Philippis, 2015). Around 75% of the cyanobacterial EPS are composed of six or more distinct monosaccharides, up to thirteen (Pereira et al., 2009; Rossi & De Philippis, 2015; Sarma, 2013), which contrasts with the polymers synthesized by other bacteria, which usually contain less than four distinct monomers (Pereira, Sousa, et al., 2019). This complexity results in a variety of linkage forms, which may explain the existence of complex repeating units and thus, the wide variety of possible structures and architectures of the cyanobacterial polymers (Kehr & Dittmann, 2015; Pereira et al., 2009).

Numerous authors have uncovered other distinctive features of cyanobacterial EPS such as the presence of sulphate groups, which are rare in bacterial EPS but a characteristic shared with archaea and eukaryotes (Pereira et al., 2009). Additionally, they may contain the simultaneous presence of more than one uronic acid (most often glucuronic and galacturonic acid). Both the uronic acid and sulphate groups contribute to their anionic nature, conferring an overall negative charge to the macromolecule (Okajima et al., 2018; Sutherland, 2001).

Methylation, acetylation, presence of peptide moieties and amphiphilic behaviour have also been described (De Philippis et al., 2011). The amphiphilic nature is the result of the presence of the hydrophobic ester-linked acetyl group, peptidic moieties, and deoxysugars and the hydrophilic sulfated sugars, uronic acids and ketal-linked pyruvyl groups (Pereira, Sousa, et al., 2019). All of these characteristics make cyanobacterial EPS very attractive for biotechnological and biomedical applications. However, it is necessary to understand the cyanobacterial EPS biosynthetic pathways, not only to optimize production yields but also to engineer polymer variants tailored for specific applications. On the other hand, when the main purpose is to potentiate the use of cyanobacteria as “cell factories” for other carbon-based products, the highly energy-consuming process of EPS production can strongly impair productivity, being necessary to redirect the carbon flux toward the production of those compounds (Pereira, Sousa, et al., 2019; Santos et al., 2021).

1.4. EPS biosynthesis in Cyanobacteria

Pereira et al. conducted an *in silico* analysis of available cyanobacterial genome sequences and identified the presence of genes encoding proteins possessing the conserved domains involved in the last steps of bacterial EPS production. The results demonstrated that most cyanobacterial strains harbour genes encoding proteins related to the three main pathways of EPS assembly and export, but often not the complete set defining one pathway. The data obtained also showed that, in cyanobacteria, these genes are frequently found in several copies distributed across the genome, either isolated or in small clusters, suggesting that, in cyanobacteria, EPS production may not follow exactly the same pathways described for other bacteria. Interestingly, as the cyanobacterial strain complexity/genome size increases, more copies of the EPS-related genes are present (Pereira et al., 2015).

At the moment, the EPS production in cyanobacteria is being mainly studied by generating and characterizing deletion mutants on putative EPS-related genes using the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*). This strain is a freshwater, unicellular, non-toxic, and non-diazotrophic cyanobacterium (Ikeuchi & Tabata, 2001; Stanier et al., 1971). It was the first cyanobacterial strain to have its genome sequenced and as a result, the genome is well characterized and annotated (Kaneko et al., 1996). It has as relatively small genome with approximately 3.5 megabases, consisting on one multi-copy circular chromosome and seven plasmids, encoding for a total of 3600 open reading frames (OFRs), of which approximately 3200 are in the main chromosome and 400 are in the plasmids (Mitschke et al., 2011). *Synechocystis* is a moderate EPS producer and is naturally transformable, being often used as a model cyanobacterial organism (Kufryk et al., 2002). This strain has genes encoding proteins putatively involved in the Wzy- or ABC transporter dependent pathways, while lacking most of the genes encoding proteins related to the Synthase dependent pathway (Pereira et al., 2015). Several of these EPS-related genes have

been targeted for deletion and, interestingly, none of the mutants obtained so far displayed a complete abolishment in EPS production, even if some mutants showed significant changes in the amount and/or composition of EPS compared to the wild-type while others showed no significant. This raises the hypothesis of functional redundancy due to the existence of multiple copies for some of the targeted genes. In addition, the involvement of proteins putatively related to the Wzy- or the ABC-dependent pathways in *Synechocystis* EPS production was confirmed (Pereira et al., 2015; Pereira, Santos, et al., 2019). Regarding the genes/proteins putatively related to the Wzy-dependent mechanism, the deletion of *wzc* (*sll0923*) resulted in a decrease in the amount of RPS and CPS, whereas Δ*wzb (slr0328)* exhibited only a reduction in RPS (Jittawuttipoka et al., 2013; Pereira, Santos, et al., 2019). Furthermore, a Δ*wzc*Δ*wzb* mutant exhibited a decrease in CPS and an increase of RPS, suggesting that in the absence of both Wzc and Wzb the RPS biosynthesis is diverted to another route (Pereira, Santos, et al., 2019). Wzc and Wzb were further characterized *in vitro.* The phosphatase activity of Wzb was verified, and Wzc was shown to possess the ATPase and auto-kinase activities of bacterial tyrosine kinases. Most importantly, Wzc was a substrate of Wzb (at least *in vitro*), providing the first evidence that tyrosine phosphorylation and dephosphorylation may play a role in the regulation of cyanobacterial EPS production (Pereira, Santos, et al., 2019). The OPX protein Wza (Sll1581) was also shown to participate in EPS production (Allen et al., 2019; Jittawuttipoka et al., 2013). Regarding putative components of the ABC transporter-dependent pathway, the involvement of KpsM (Slr0977 and Sll0574) and KpsT (Slr0982 and Sll0575) homologues in EPS production was also confirmed (Fisher et al., 2013; Santos et al., 2021) The total carbohydrate content of a *kpsM* (*slr0977*) deletion mutant was comparable to that of the wild-type, but the mutant produced significantly less capsular polysaccharides (CPS) and released polysaccharides (RPS). Additionally, the mutant accumulated significantly more polyhydroxybutyrate (PHB) (a carbon storage compound) than the wild-type/complemented mutant, which suggests redirection of the carbon flow from EPS biosynthesis to the production of storage molecules (Santos et al., 2021). Inactivation of *kpsM* also showed an impact on the transcript levels of genes putatively involved in the Wzy-dependent pathway of EPS assembly and export, with an upregulation of *wza* and *wzc* and downregulation of *wzb*. Based on these results it was hypothesized that the mutant attempts to compensate the absence of KpsM either by using a distinct export route or by using other components, as the two canonical bacterial pathways might not function as separate entities in *Synechocystis* (Santos et al., 2021). Furthermore, the monosaccharidic composition of the EPS produced by the *kpsM* knockout mutant (*slr0977*) was different than that of the polymer produced by the wild-type. The authors proposed that in the absence of the native pore Slr0977, there is an increase in the activity of another homologue that led to the improper transport of extracellular carbohydrates (Fisher et al., 2013).

1.5. EPS biosynthesis

In general, the EPS biosynthesis in bacteria begins with the conversion of monosaccharides into nucleotide sugars in the cytoplasm. Then, glycosyltransferases transfer these sugars to specific acceptors in the plasma membrane. These initial steps are strain dependent processes (Nouha et al., 2018; Pereira et al., 2015; Pereira et al., 2009). On the other hand, the polymerization, assembly, and export of the polymer that follow these steps appear to be somewhat conserved in bacteria, usually following one of three main pathways (Schmid et al., 2015; Whitfield & Larue, 2008) schematically represented in Figure 1, the ABC transporter-, the wzy- or the synthase-dependent pathways. In the ABC transporter-dependent pathway, the polymer is fully polymerized in the cytoplasmic side of the plasma membrane before being transported to the periplasmic side by an ABC transporter composed by the permease KpsM and the ATP-binding element KpsT (Willis & Whitfield, 2013). Several polymers are produced following these pathways, such as the *Escherichia coli* group 2 capsular polysaccharides (Whitfield, 2006). In the Wzy-dependent pathway, oligosaccharide lipid-linked repeating units are formed in the cytoplasmic side of the plasma membrane before being flipped to the periplasmic side by the flippase Wzx and polymerized by Wzy. Subsequently, the polymer is exported through the transenvelope complex formed by the polysaccharide co-polymerase (PCP) Wzc and the outer membrane polysaccharide export (OPX) protein Wza. The conformation of this complex appears to be important to regulate EPS transport, being dependent on the phosphorylation state of the C-terminal domain of Wzc that, in turn, is controlled by the function of the low molecular weight phosphatase Wzb (Pereira, Sousa, et al., 2019; Schmid et al., 2015). This pathway is involved, e.g., in the production of *E. coli* group 1 capsular polysaccharides (Whitfield, 2006). The third is the Synthase-dependent pathway, where the polymer is simultaneously polymerized and exported by a single synthase protein that, in some cases, is a part of a multiprotein complex (Schmid et al., 2015).

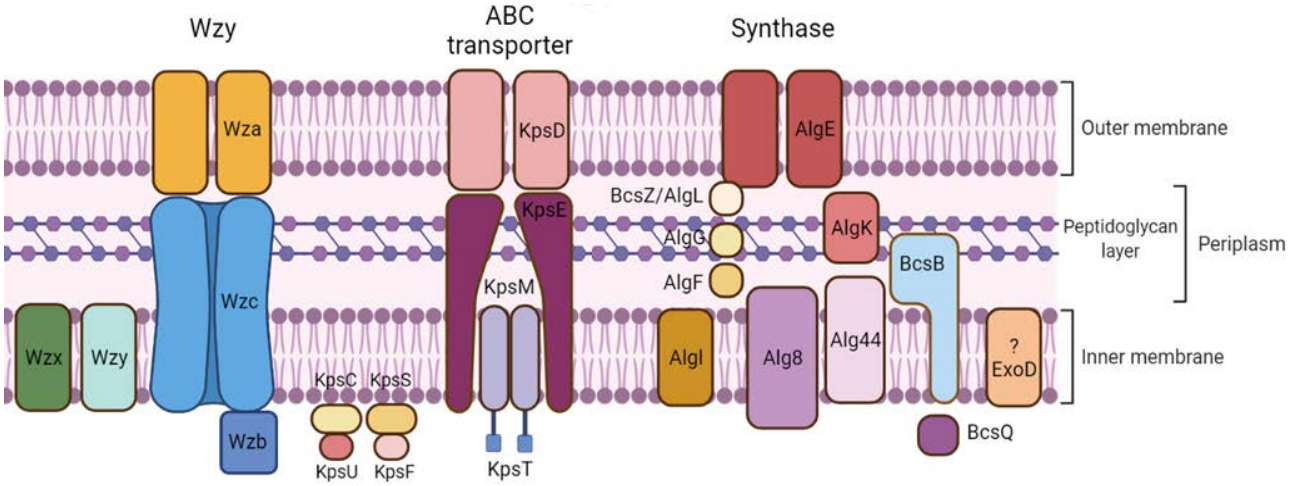

Figure 1 - Schematic representation of the main pathways of bacterial EPS assembly and export. Adapted from Pereira et al., 2015.

1.6. Aims

The objective of this work was to generate the molecular tools (plasmids) and develop the methodology to further characterize of *Synechocystis* Wzc (Sll0923) and KpsM (Slr0977), including the identification of their functional partners. This knowledge is crucial to elucidate the molecular mechanisms of EPS production and to clarify the hypothesis of crosstalk between proteins associated to different pathways. To achieve that, a tagged version of *Synechocystis* Wzc was overexpressed in *E. coli*, purified, and used as bait in pull-down assays using S*ynechocystis* protein extracts as a target (*in vitro* studies). In addition, plasmids containing the sequence encoding a tagged version of KpsM were generated. These plasmids will be used to complement the *kpsM* deletion mutant and, subsequently, the proteins interacting with KpsM will be identified by *in vivo* chemical crosslink treatment followed by pull-down assays (*in vivo* studies). Furthermore, the generation of plasmids to complement *E. coli* acapsular mutants with *Synechocystis* Wzc or its truncated version (without the C-terminal domain were its phosphorylation/dephosphorylation occurs) and evaluate its ability to restore capsule formation was initiated. This approach allows the characterization of Wzc *in vivo*, using of a well-characterized Wzy-dependent EPS production system while avoiding *Synechocystis* genomic redundancy.

2. Materials and Methods

2.1. Organisms and growth conditions

The *Escherichia coli* strains DH5α, XL1-Blue, M15(pREP4) and E69 were cultured at 37ºC in LB medium (Miller, 1972) or medium supplemented with Ampicillin (100 ug/mL), Kanamycin (25 ug/mL) and/or Chloramphenicol (25 ug/mL). For solid medium, liquid LB was supplemented with agar at 1.5%.

Synechocystis sp. PCC 6803 substrain Kazusa (Pasteur Culture Collection) was used in this study. *Synechocystis* wild-type and Δ*wzc* were grown in BG11 media at 30°C with orbital agitation (150 rpm) under a 12h light (50 mmol photons $m^{-1}s^{-2}$)/12h dark cycle. BG11 medium was supplemented with kanamycin (700 mg/mL) for the mutant's maintenance.

2.2. Growth assessments

Growth measurements were performed by monitoring the Optical Density (OD) at 600 nm for *E. coli* or 730 nm for *Synechocystis* (Anderson & McIntosh, 1991) using a Shimadzu UVmini-1240 spectrophotometer (Shimadzu Corporation). In addition, the chlorophyll *a* content of *Synechocystis* was determined as previously described (Meeks & Castenholz, 1971).

2.3. DNA Extractions

The *Synechocystis* and *E. coli* genomic DNA were extracted using Maxwell® 16 System (Promega). Isolation of plasmid DNA was performed using the NZYMiniprep kit (NZYTech) following the manufacturer's instructions.

2.4. Agarose gel electrophoresis

Nucleic acids electrophoresis was performed using standard techniques (Sambrook, 2001). For that, 0.8 or 1% (w/v) agarose gels were prepared with 1x TAE buffer supplemented with GreenSafe (NZYTech). A Gel DocTM XR+ Imager was used to visualize the bands under UV light (Bio-Rad). The GeneRuler™ DNA Ladder Mix (ThermoScientific) was used as a molecular weight marker.

2.5. DNA purification and quantification

DNA fragments extracted from gels, enzymatic, or PCR reactions were purified using the NZYGelpure purification kit (NZYTech). All purified DNA was quantified by spectrophotometry using a Nanodrop ND-1000.

2.6. Plasmid generation

Generation of pSEVA351*::*$P_{Trc.x.tetO2}$/P_{psbA2*}*::*B0030*::his6-kpsM*

The plasmid pSEVA351 (plasmid map shown in Annexes; Figure A) was selected as replicative vector for the complementation of the *Synechocystis kpsM* knockout mutant with the respective tagged gene derivate. To achieve that, a DNA fragment containing the P_{psbA2*} or $P_{Trc.x.tetO2}$ promoter and the synthetic RBS B0030 were obtained from, digestion of the pSB1A2::P_{psbA2*}*::*B0030 or pSB1A2::$P_{trc.x.tetO2}$*::*B0030 previously generated in the host research group, with *Eco*RI and *Spe*I. The gene encoding the his-tagged version of *Synechocystis kpsM* (hereafter *his6-kpsM*) was obtained by PCR using pGEM-Teasy::*his6-kpsM* as template (Annexes; Figure B). Each PCR reaction mixture (50 µL) contained 1U of High-Fidelity Phusion DNA polymerase (Thermo Scientific), 1x Phusion HF buffer (Thermo Scientific), 0.2 µM of each deoxyribonucleotide triphosphate, 500 µM of each primer (Table 1) and 1-10 ng of the previously generated pGEM-Teasy::*his6-kpsM* as template. The PCR profile was: 30 sec at 98°C followed by 35 cycles of 10s at 98°C, 30s at 70°C and 2 min at 72°C, with a final extension at 72°C for 10 min. Primers were designed to introduce a *Xba*I and *Pst*I restriction sites. The PCR product obtained was purified, digested with *Xba*I and *Pst*I and purified again. The digested DNA fragments containing the P_{psbA2*} and RBS B0030 or the *his6-kpsM* were cloned into pSEVA351 previously digested with *EcoR*I and *Pst*I, originating plasmid pSEVA351*::*$P_{Trc.x.tetO2}$/P_{psbA2*}::B0030*::his6-kpsM* (theoretical sequences shown in Annexes; Figure C and D). The constructs were verified by DNA sequencing (Stab Vida).

Generation of pBAD18-cm*::e69wza::e69wzb* and amplification of *e69wzc*, *wzc* and wzc_{Trunc}

The plasmid pBAD18-cm (Guzman et al., 1995; plasmid map shown in Annexes; Figure E) was selected for the functional complementation of the acapsular *E. coli* mutant CWG655 (Reid & Whitfield, 2005). The *E. coli* K30 E69 *wza* (*e69wza)*, *wzb* (*e69wzb*) and *wzc* (*e69wzc)*, as well as the *Synechocystis wzc* (*wzc*; *sll0923*) and its truncated version (wzc_{Trunc}*;* $sll0923_{Trunc}$) were amplified by PCR using genomic DNA extracted from *E. coli* K30 E69 or *Synechocystis*, respectively. The reaction mixtures (50 µL) contained 1u of High-Fidelity

Phusion DNA polymerase (Thermo Scientific), 1x Phusion HF buffer (Thermo Scientific), 50 µM/µL of each deoxyribonucleotide triphosphate, 500 µM/µL of each primer (Table 1) and 1-10 ng of *E.coli* K30 E69 or *Synechocystis* genomic DNA. Primer pairs were designed to introduce an RBS and enzyme restriction sites. The PCR profiles were: 30 sec at 98°C followed by 35 cycles of 10s at 98°C, 30s at 54°C (*e69wza*), 62°C (*e69wzb)*, 65°C (*e69wzc*) or 59ºC (*wzc and* wzc_{Trunc}) and 2 min at 72°C, with a final extension at 72°C for 10 min. The introduced restriction sites were *Xba*I and *Sal*I for *e69wza*, *Sal*I and *Sph*I for *e69wzb* and *Sph*I for *e69wzc*, *wzc and* wzc_{Trunc}. After purification, amplicons were digested with the respective enzymes, purified, and *e69wza* and *e69wzb* were cloned into pBAD18-cm previously digested with *Xba*I and *Sal*I, generating plasmid pBAD18-cm*::e69wza::e69wzb* (partial sequence shown in Annexes; Figure F). The construct was confirmed by DNA sequencing (Stab Vida).

Table 1 - List of primers used in this work.

Designation	Sequence	Function	Reference
pBAD18-cm_5'	CTGTTTCTCCATACCCGTT	Amplification of pBAD18-cm cloning site	(Guzman et al., 1995)
pBAD18-cm_3'	CTCATCCGCCAAAACAG	Amplification of pBAD18-cm cloning site	
FC2e69wzaF	GCTCTAGACATTAAAGAGGAGAAATACTAGATGAAGAAAAAACTTGTTAG	Amplification of *E. coli wza*, introducing *Xba*I site and RBS B0030	This work
FC2e69wzaR	CGCGTCGACTTAGTTTGGCCATCTCTTAATG	Amplification of *E. coli wza*, introducing *Sal*I site	
FC2e69wzbF	CGCGTCGACATTAAAGAGGAGAAATACTAGATGGCCAAACTAATGTTTG	Amplification of *E. coli wzb*, introducing *Sal*I site and RBS B0030	
FC2e69wzbR	GCGCATGCTTATTCACCTAATTTCTCAGCCCA	Amplification of *E. coli wzb*, introducing *Sph*I site	
FC2e69wzcF	GCGCATGCCATTAAAGAGGAGAAATACTAGATGACTTCAGTCACATCTAAGC	Amplification of *E. coli wzc*, introducing *Sph*I site and RBS B0030	
FC3e69wzcR	GCGCATGCAGAGTCTATGGTTTTGTCAGTC	Amplification of *E. coli wzc*, introducing *Sph*I site	
FC2sll0923F	GCGCATGCATTAAAGAGGAGAAATACTAGATGACCTACAGTTACCCAGAA	Amplification of *Synechocystis wzc*, introducing *Sph*I site and RBS B0030	
FC3sll0923R_2	GTGGCATGCGTTCTGTGCTCCCGATCCTAATTG	Amplification of *Synechocystis wzc*, introducing *Sph*I site	
FC3sll0923 TruncR	CGCATGCTTATTAGCCGGACGTAGAAGTGATAGCA	Amplification of *Synechocystis* wzc_{Trunc}, introducing *Sph*I site	
Histag_F	GCTCTAGAATGAGAGGATCGCATCACCATC	Amplification of *Synechocystis his6-kpsM*	This work
slr0977Rev_compR	GTTTCTTCCTGCAGCGGCCGCTACTAGTATTAAATCACATCAGCGAAGGTGC	Amplification of *Synechocystis his6-kpsM*	
sll0923_compR2	GTTTCTTCCTGCAGCGGCCGCTACTAGTACTAATTGTCTGCCTGGGCAAC	Amplification of *Synechocystis his6-wzc*	
PS1	AGGGCGGCGGATTTGTCC	Amplification of pSEVA351 cloning site	(Silva-Rocha et al., 2013)
PS2	GCGGCAACCGAGCGTTC	Amplification of pSEVA351 cloning site	

2.7. Transformation and colony PCR

E. coli DH5α or XL1-Blue competent cells were thawed on ice for 10-15 min, 10 µL of ligation reactions were added and the mixtures were kept on ice for 30 min. The heat shock was done by placing tubes at 42ºC for 45 or 90 sec followed by incubation on ice for 2 min before adding 890 µL of LB medium. Cells were then incubated for 1h or 1h30 min at 37ºC with orbital shaking (approximately 150 rpm) and plated in selective medium.

Colony PCR was used to confirm the presence of transformants. Each PCR reaction mixture (20 µL) contained 0.5U of GoTaq DNA polymerase (Promega), 1x GoTaq Flexi buffer (Promega), 250 µM of each deoxyribonucleotide triphosphate, 500 µM of each primer (Table 1) and 5µL of bacterial colony suspension. The colony PCR profile was: 5 min at 95°C followed by 35 cycles of 45s at 95°C, 45s at 52°C (pBAD18-cm) or 55°C (pSEVA351) and 3 min at 72°C, with a final extension at 72°C for 10 min.

2.8. Overexpression and purification of His6-Wzc

The overexpression and purification of his-tagged Wzc (hereby His6-Wzc) was performed as previously described (Pereira, Santos, et al., 2019). Briefly, *E. coli* M15(pREP4; plasmid map shown in Annexes; Figure G) transformed with pQE-30::*his6-wzc* (plasmid map shown in Figure H, partial sequence shown in Figure J;) were grown ON in LB medium supplemented with ampicillin and kanamycin at 37ºC and 150 rpm. After reaching an OD600 of 0.7-0.8, the cultures were incubated for 30 min on ice with occasional agitation. After the cold shock, IPTG was added to a final concentration of 0.5 mM and cultures were incubated for 18h at 20ºC and orbital shaking (150 rpm). Subsequently, the cultures were centrifuged at 3 000g for 25 min at 4ºC, the cell pellets were weighted and stored at -80ºC. To purify His6-Wzc, cell pellets were thawed on ice and resuspended in Buffer A (50mM Hepes, 100 mM NaCl, pH 8.0) at a volume of 5 mL per gram of cell. $MgCl_2$, DNAse and PMSF were then added to a final concentration of 1 mM, 10 ug/mL, and 1 mM, respectively. Cells were broken using a chilled French press at a pressure 30 kpsi. Unbroken cells and debris were pelleted at 30 000g for 30 min at 4ºC and the supernatant was further centrifuged at 200 000g for 1 hour at 4ºC. The pellet (membrane fraction) was resuspended in Buffer A. The membrane proteins were solubilized by adding Triton X-100 to a final concentration of 1% and samples were incubated at 4ºC ON with mild agitation to maximize protein solubilization. Subsequently, the detergent insoluble fraction was pelleted at 40 000g, for 45 min at 4ºC. Samples were stored at -80ºC until further use.

The purification of His6-Wzc was performed by immobilized metal affinity chromatography (IMAC). For that, the protein extracts were loaded in 1 mL His Trap Columns (GE Healthcare) in Wzc-binding buffer (50 mM HEPES pH 8.0, 100 mM NaCl, 20 mM imidazole, 0.2% Triton X-100), and bound proteins were eluted using a step gradient in which

the imidazole concentrations increased from 50 mM for 15 min, 200 mM for 20 min and 500 mM for 15 min. Pooled fractions containing His6-Wzc were concentrated using Amicon Ultra-15 Centrifugal Filter units with 30 kDa MW cut-off (Merck Millipore) and the buffer was exchanged to pull-down binding/washing buffer (50 mM HEPES, 300 mM NaCl, 0.2% Triton-X100, pH 8.0) by diafiltration. When necessary, His6-Wzc was further purified by size exclusion chromatography (SEC) using a SuperoseTM 6 10/300 GL (GE healthcare). The concentrations of the purified his-tagged protein solutions were determined by BCA colorimetric assay (Thermo Scientific) using bovine serum albumin as standard. Protein samples were separated by electrophoresis either on 10% or on pre-cast gradient 4-15% (Bio-Rad) SDS-polyacrylamide gels and visualized with Roti-Blue (Carl Roth).

2.9. Pull-down assays

Pull-down assays were performed using purified His6-Wzc (bait) and protein extracts of *Synechocystis* wild-type or *Δwzc* (target). For that, *Synechocystis* was grown until OD730nm reached approximately 1.5 and cultures were harvest by centrifugation at 4000 g for 5-10 min, at RT. Subsequently, cells were resuspended in pull-down buffer (10 mM HEPES; 70 mM NaCl; 0.2% Triton-X100; pH 7.4) supplemented with Complete EDTA-free Protease Inhibitor (Roche) and broken by sonication using a Branson sonifier operating with the following settings: output 3, duty cycle 50%, 6 x 15 sec. Cell debris were removed by centrifugation at 5000 g for 5 min at 4°C to obtain the total protein extract. When necessary, the membrane fraction was separated from the soluble proteins by ultracentrifugation at 100 000 g, for 1h at 4°C and resuspended in pull-down buffer. Protein concentration was measured as described above, and pull-down assays were performed using Dynabeads His-Tag Isolation & Pulldown (Novex Life technologies) according to the manufacturer's instructions. Briefly, 125 ug of His6-Wzc (IMAC) or 50 ug of His6-Wzc (IMAC and SEC) in Pull-down binding/washing buffer supplemented with 50 mM imidazole were incubated with the beads in a roller at 4°C for 1h. After washings, 2, 3 or 1.5 mg of *Synechocystis* total, soluble or membrane protein extract respectively, were added to tubes and samples were incubated in a roller, at 4°C ON. After washings, the proteins/protein complexes were recovered by incubation in elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM Imidazole, 0.2% Triton-X100, pH 8 0) in a roller for 15-30 min at RT. Controls included samples without bait or target. The proteins eluted were analysed by SDS-PAGE and visualized using Roti-Blue (Carl Roth) or by staining with silver nitrate (Gromova & Celis, 2006).

3. Results

3.1. Identification of Wzc and KpsM protein interactions

In order to clarify mechanisms involved in EPS biosynthesis, and moving from the previously established methodology for His6-Wzc overexpression and purification, the recombinant protein was used as bait in pull-down assays to isolate its potential functional partners from *Synechocystis* protein extracts. In addition, since no recombinant *Synechocystis* KpsM is yet available, plasmids containing the sequence encoding a tagged version of this protein were generated in order to complement the *Synechocystis kpsM* knockout mutant. This will allow the identification of its functional partners by *in vivo* chemical crosslink treatment followed by pull-down assays.

3.1.1. Identification of Wzc protein interactions using purified His6-Wzc

To perform the pull-down assays using the recombinant His6-Wzc, we began by overexpressing this protein in *E. coli* M15(pREP4) (Figure 1). This strain contains the plasmid pREP4 that constitutively encodes the *lac* repressor protein (LacI). Therefore, the genes cloned in pQE-30 vectors are not transcribed until IPTG is added to the cultures. As observable in Figure 1, the His6-Wzc protein (theoretical molecular weigh 84.5 kDa; Pereira *et al.* 2019) was fully overexpressed. It is also possible to observe that in the absence of IPTG there is residual expression of His6-Wzc since a band with the expected size is also visible in the non-induced samples, even if with lower intensity.

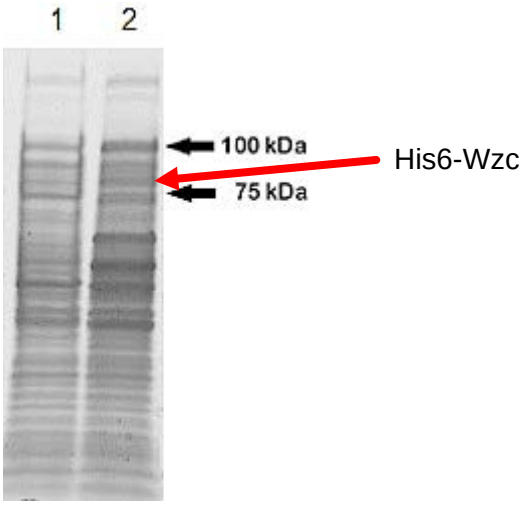

Figure 2 – SDS-PAGE analysis of the overexpression of His6-Wzc in *E. coli* M15(pREP4). Lane 1 - Non-induced sample. Lane 2 – Sample induced with 0.5mM of IPTG. Gel stained with Roti-Blue.

Subsequently, the His6-Wzc was isolated from extracts of the induced cells by IMAC (Figure 2). When the concentration of imidazole reached approximately 200 mM (38% buffer B), it was possible to observe a peak (fractions 19 to 29) indicating the elution of high amount of protein (Figure 2A). The SDS-PAGE analysis of the pool of these fractions revealed the presence of a band with the expected size for His6-Wzc, together with other bands (Figure 2B). Since the pull-down assays involve a further step of isolation based on the affinity of the His-tag for the metals coating the beads, the sample obtained was used for the subsequent assays.

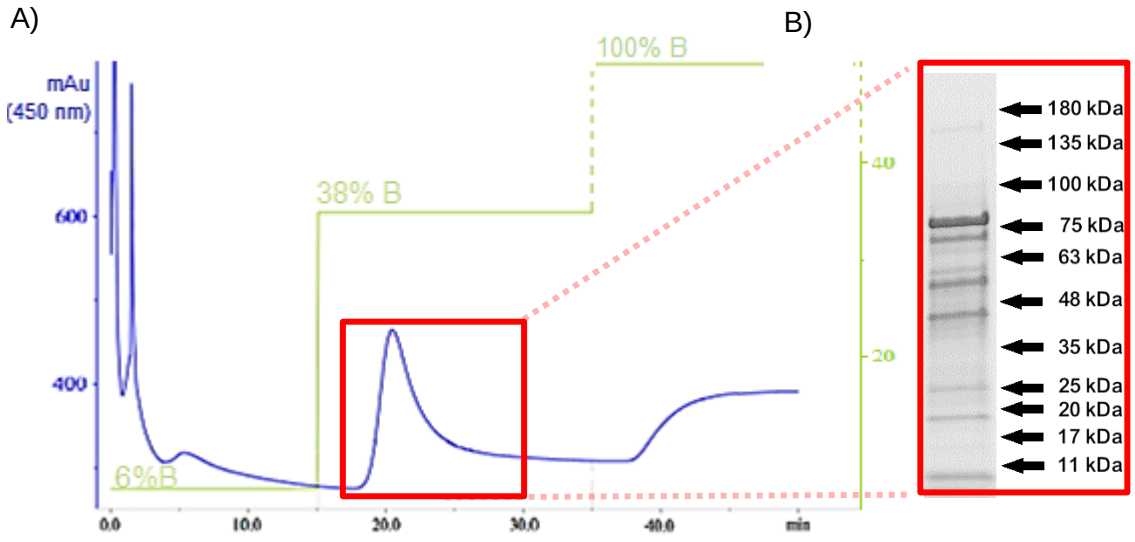

Figure 3 - IMAC purification of His6-Wzc. A) IMAC chromatogram. His6-Wzc was eluted using a step gradient (% of elution buffer B in green). The peak highlighted in red was selected for further analysis. B) SDS-PAGE analysis of the pool IMAC fractions selected. The stronger band (approximately 85 kDa) is consistent with His6-Wzc. Gel stained with Roti-Blue. NzyColour Protein Marker II (3µl) used.

After purification, the binding of His6-Wzc to the metal coated beads was tested using different amounts of purified protein, followed by analysis by SDS-PAGE (Figure 3). His6-Wzc was partially bound to the beads, being the amount of protein bound higher when 250 ng of purified protein was used compared to 125 ng. However, it is important to take into consideration that the amount of other proteins that bind unspecifically to the beads is also increased when loading a higher amount of protein, which may increase the number of false positives in the pull-down assays. Therefore, the subsequent assays were performed using 125 ng of purified protein. Loading the same initial protein concentration but varying the elution time and/or imidazole concentration did not lead to higher elution of His6-Wzc (data not shown).

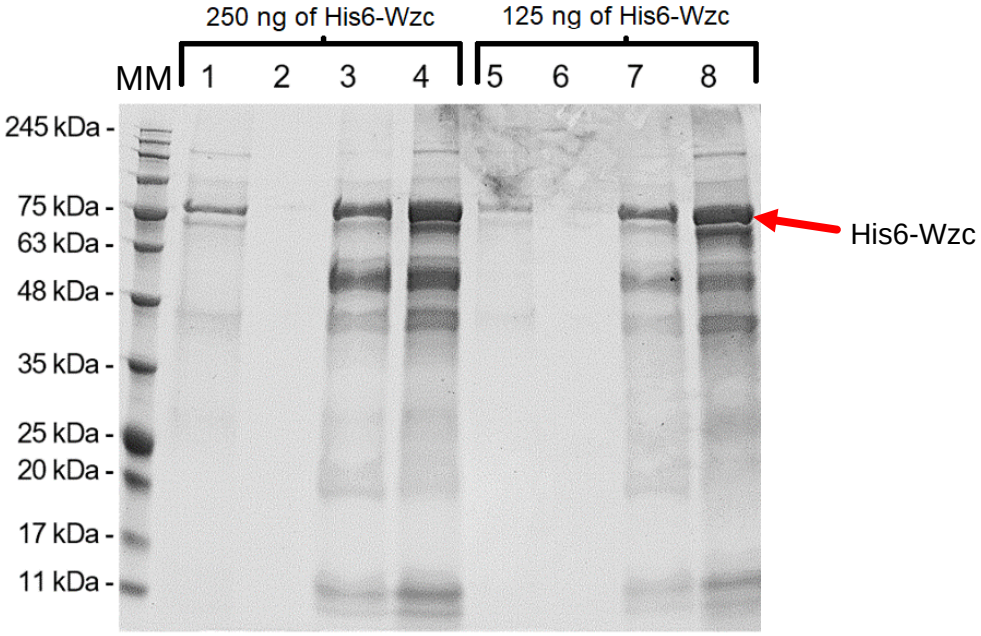

Figure 4 - SDS-PAGE analysis of His6-Wzc binding to IMAC beads. Lane 1 to 4 - 250 ng of purified His6-Wzc; Lane 5 to 8 - 125 ng of purified His6-Wzc. Lanes 1 and 5 – Protein extract recovered after binding of His6-Wzc onto the beads. Lanes 2 and 6 - Washes of the beads after loading the His6-Wzc. Lanes 3 and 7 – Supernatant containing the eluted His6-Wzc/His6-Wzc complexes. Lanes 4 and 8 - Beads after elution of His6-Wzc. Gel stained with Roti-Blue. MM: NzyColour Protein Marker II (3µl).

After partial isolation of the His6-Wzc on IMAC beads, the pull-downs were performed using total protein extracts of *Synechocystis* wild-type and Δ*wzc* as target (Figure 4).

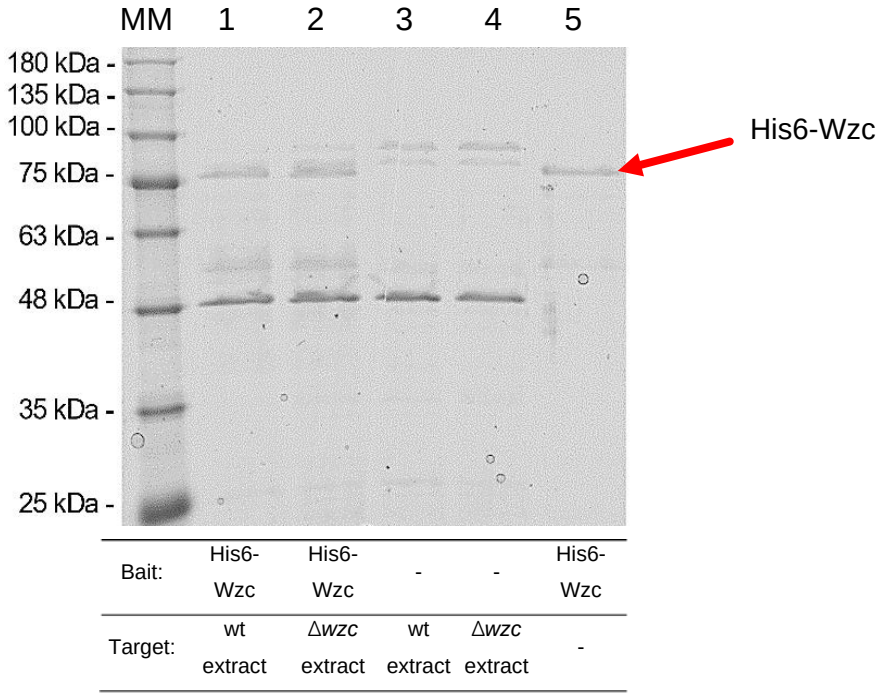

Figure 5 - SDS-Page analysis of pull-down assays using His6-Wzc purified by IMAC as bait and total protein extracts of *Synechocystis* wild-type (wt) or Δ*wzc* as target. Pull-downs were performed using 125 ng of bait and 2 mg of target. Lanes 1 – Purified His6-Wzc used as bait and total protein extract of the wt used as target. Lane 2 - Purified His6-Wzc used as bait and total protein extract of the Δ*wzc* used as target. Lane 3 – Control with no bait and total protein extract of the wt used as target. Lane 4 – Control with no bait and total protein extract of the Δ*wzc* used as target. Lane 5 – Control with purified His6-Wzc used as bait and no target. Gel stained with Roti-Blue. M: NzyColour Protein Marker II (3µl).

It was possible to observe a band with the expected size for His6-Wzc in all of the conditions where this protein was used as bait, confirming its successful isolation on the beads. However, there was also presence of other bands in these conditions highlighting the need to further purify His6-Wzc. Regarding samples containing *Synechocystis* wt or Δ*wzc* protein extracts, several bands are visible, these bands however were also present in the controls without the bait, suggesting that they result from unspecific binding of *Synechocystis* proteins to the beads.

Moving from these results, His6-Wzc was further purified by size exclusion chromatography (SEC) and the pull downs were performed using either *Synechocystis* wild-type membrane or soluble extracts. This approach was used since it is well established that membrane proteins are underrepresented in a total extract and Wzc is expected to interact with either membrane or soluble proteins (Sato et al., 2007; Xu et al., 2021). The SDS-PAGE results of the pull-down assay are shown in Figure 5.

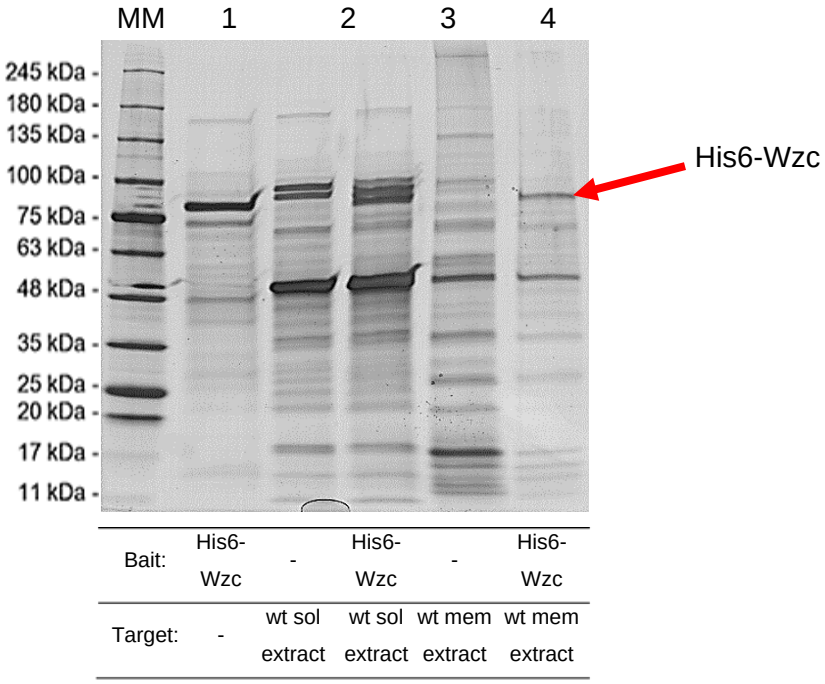

Figure 6 – SDS-Page analysis of pull-down assays using His6-Wzc purified by IMAC and SEC. Lanes 1 - Control with purified His6-Wzc used as bait and no target. Pull-downs were performed using 50 µg of bait and 1.5 or 3 mg of target, for *Synechocystis* membrane protein or soluble proteins extracts, respectively. Lane 2 - Control with no bait and soluble protein extracts as target. Lane 3 - Purified His6-Wzc used as bait and soluble protein extracts as target. Lane 4 - Control with no bait and membrane protein extracts as target. Lane 5 - Purified His6-Wzc used as bait and membrane protein extract as target. Gel stained with silver staining. MM - NzyColour Protein Marker II (3µl).

As it is possible to observe (Figure 5, lane 1) the further step of His6-Wzc purification was not sufficient to remove all unspecific proteins in all of the conditions containing His6-Wzc as bait. However, it is necessary to take into consideration that this gel was stained with silver nitrate, which is a methodology with higher sensitivity compared to the stain with Rote-Blue (Gromova & Celis, 2006). In addition, many of the bands that are observed in the samples containing the *Synechocystis* membrane or soluble extracts are also present in the controls with no target. Despite that, some differences are observed, particularly for bands showing lower intensity in the gel. In the case of the pull-downs performed with the membrane extracts, a higher number of bands is observed in the control without bait, suggesting that the absence of His6-Wzc leads to more unspecific binding of membrane proteins to the beads. At the moment, the proteins present in the samples are being identified by Mass Spectrometry (MS) Once the results are obtained, the data will be analysed.

3.1.2. Generation of plasmids to identify KpsM protein interactions

To identify the proteins that interact with KpsM a previously generated *Synechocystis kpsM* knockout mutant will be complemented with a replicative plasmid (pSEVA351) containing the sequence encoding the his-tagged version of the native KpsM. This cloning step is necessary, since at the moment, no antibody is available for *Synechocystis* KpsM. Once the complementation of the mutant is achieved, chemical crosslink will be performed to permanently link all the interacting partners in the cell. After that, the protein complexes containing His6-KpsM will be isolated by affinity (IMAC) and its putative functional partners will be identified by MS. This methodology was selected since it has the potential to more easily identify transient interactions compared to the use of purified recombinant protein used as baits in pull-down assays.

In this work, we began by amplifying the gene encoding His6-KpsM from the previously generated pGEM-Teasy::*his6-kpsM*. Subsequently, a DNA sequence containing the promoter $P_{trc.x.tetO2}$ (Ferreira et al., 2018) and a synthetic RBS were obtained from digestion of plasmid pSB1A2:: $P_{trc.x.tetO2}$::B0030 available in the host group. After cloning of the DNA fragment in the replicative plasmid pSEVA351, the resulting plasmids were isolated from *E. coli* and the presence of the inserts was evaluated by restriction analysis (Figure 6).

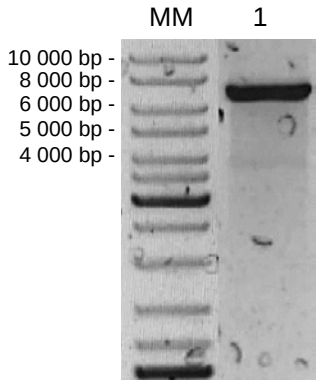

Figure 7 – Restriction analysis of pSEVA351::$P_{trc.x.tetO2}$::B0030::*his6-kpsM*, isolated from *E. coli* colony. Plasmids were digested with *Xba*I. The expected band size is 6176 bp. MM - GeneRulerTM DNA Ladder Mix molecular marker.

The pSEVA351::$P_{trc.x.tetO2}$::B0030::*his6-kpsM* restriction profile obtained was similar to what was expected. Therefore, the construct was sequenced.

The sequencing of the plasmid revealed an insertion of a cytosine 15 bp after the His-tag, leading to the disruption of the reading frame (frameshift mutation) as illustrated in Figure 7.

A

atg	aga	gga	tcg	cat	cac	cat	cac	cat	cac	gga	tcc	aaa	act	tcc	ccc	cca	gaa	ctg	att
M	R	G	S	H	H	H	H	H	H	G	S	K	T	S	P	P	E	L	I
att	gaa	gca	gga	cgc	acg	gag	cgt	cag	tat	tgg	caa	gac	cta	tgg	cgt	tac	cgg	gaa	ttg
I	E	A	G	R	T	E	R	Q	Y	W	Q	D	L	W	R	Y	R	E	L

B

atg	aga	gga	tcg	cat	cac	cat	cac	cat	cac	gga	tcc	aaa	act	tcc	ccc	ccc	aga	act	gat
M	R	G	S	H	H	H	H	H	H	G	S	K	T	S	P	P	R	T	D
tat	tga	agc	agg	acg	cac	gga	gcg	tca	gta	ttg	gca	aga	cct	atg	gcg	tta	ccg	gga	att
Y	-	S	R	T	H	G	A	S	V	L	A	R	P	M	A	L	P	G	I

Figure 8- Comparison of the theoretical DNA and amino acid sequence (partial) of *his6-kpsM* (A) and that obtained by sequencing pSEVA351::$P_{trc.x.tetO2}$::B0030::*his6-kpsM* construct (B). The start codon of the genes is highlighted in green; the sequence encoding the 6x histidine tag is highlighted in blue. Cytidine insertion mutation is highlighted in red. The resulted frameshift caused by the insertion is highlighted in yellow.

A similar strategy was followed to clone *his6-kpsM* under the regulation of a different promoter, namely P_{psbA2*}, thus generating plasmid pSEVA351::P_{psbA2*}::B0030::his6-*kpsM*. After transformation in *E. coli,* plasmids were isolated from selected colonies and their restriction profiles were analysed (Figure 8).

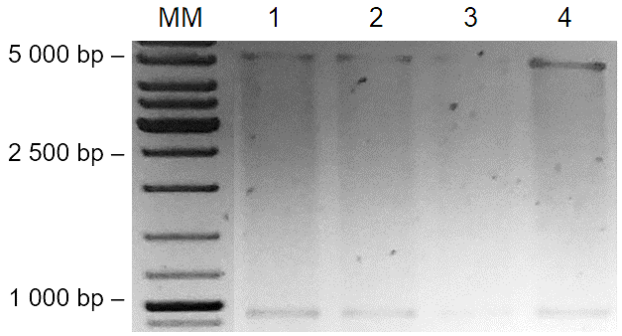

Figure 9 - Restriction analysis of pSEVA351::P_{psbA2*}::B0030::*his6-kpsM* isolated from different *E. coli* colonies. Plasmids were digested with *Eco*RI and *Pst*I. Expected band sizes are 5088 bp, 921 bp and 69 bp MM - GeneRuler™ DNA Ladder Mix.

The restriction profile obtained for the plasmids was similar to what was expected for pSEVA351::P_{psbA2*}::B0030::*his6-kpsM*, the construct was confirmed by sequencing. Transformation of *Synechocystis kpsM* knockout mutant with this plasmid is on-going.

A similar procedure is ongoing to generate the pSEVA351::P_{psbA2*}::B0030::*his6-wzc* and pSEVA351::P_{psbA2*}::B0030::*his6-wzc*$_{Trunc}$ plasmids to clone into *Synechocystis* Δ*wzc* mutant with the native Wzc and the his-tagged version, respectively. This will complement the results obtained in the *in vitro* pull-down assays.

3.2. Generation of plasmids to characterize the role of Wzc *in vivo*

Previous work demonstrated the role of Wzc and Wzb in *Synechocystis* EPS production (Pereira, Santos, et al., 2019). Importantly, it provided the first data on the function and/or structure of two proteins involved in this process in cyanobacteria and the first indication that tyrosine phosphorylation/dephosphorylation may be involved in the regulation of this process. To further characterize Wzc and elucidate its role *in vivo*, we initiated the generation of plasmids to express *Synechocystis* Wzc or its truncated version (lacking the C-terminal domain where its phosphorylation and dephosphorylation occurs) (Pereira, Santos, et al., 2019) in acapsular *E. coli* mutants and evaluate its ability to restore capsule formation. This approach allows the use of a well-characterized polysaccharide production system, namely, the production of *E. coli* group 1 capsules (Whitfield, 2006; Whitfield et al., 2020), while avoiding *Synechocystis* genomic redundancy (Pereira et al., 2015). For that, we use the model strain *E. coli* K30 strain E69 and its mutant CWG655, in which the chromosomal region spanning the (*wza-wzb-wzc*)$_{K30}$ cluster was deleted and the expression on the second copy of *wza, wzb* and *wzc* (22-min locus) was eliminated by polar insertion of an antibiotic cassette (Reid & Whitfield, 2005; Whitfield, 2006). This strain will be complemented with a plasmid containing the arabinose P_{BAD} promoter (Guzman et al., 1995) encoding the native *wza* and *wzb* genes of *E. coli* K30 E69 and either its native *wzc* (as control) or the *Synechocystis wzc* or *wzc*$_{Trunc}$.

E. coli K30 E69 native *wza* (*e69wza*) and *wzb* (*e69wzb*) were successfully amplified from genomic DNA (Figure 9).

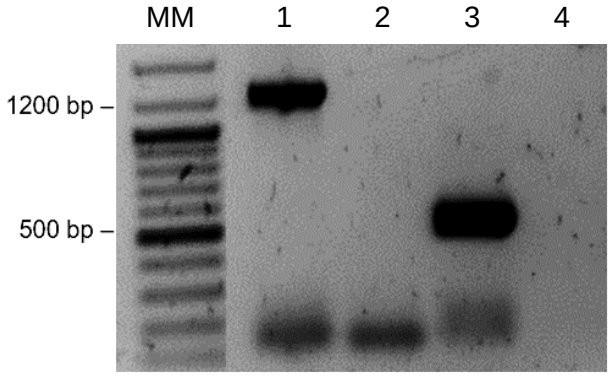

Figure 10 – Agarose gel electrophoresis analysis of PCR products amplified with specific primers for *E. coli* K30 E69 native *wza* (*e69wza*; lane 1) and *wzb* (*e69wzb*; lane 3). Lanes 2 and 4 - Negative controls with no genomic DNA. Expected band sizes are 1 179 bp or 485 bp for *e69wza* or *e69wzb*, respectively. MM - GeneRuler™ DNA Ladder Mix.

The PCR products were purified from agarose gels, digested with restriction enzymes, and cloned into pBAD18-cm. After transformation, plasmids were isolated from selected colonies and their restriction profiles were analysed (Figure 10). For one of the plasmids, a band with expected size was observed and its sequence was confirmed by DNA sequencing.

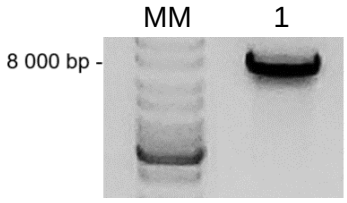

Figure 11 - Restriction analysis of pBAD18-cm::*e69wza*::*e69wzb* isolated from *E. coli*.. Plasmid was digested with *Sph*I. Expected band size is 7 650 bp. MM - GeneRulerTM DNA Ladder Mix.

Subsequently, the *wzc* gene from *E. coli* K30 E69 (*e69wzc*), *wzc* from *Synechocystis (wzc)* and its truncated version (wzc_{Trunc}) were then amplified and purified (Figure 11).

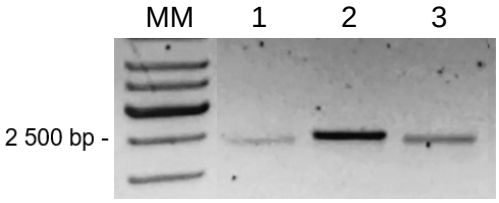

Figure 12 – Agarose gel electrophoresis of the purified PCR products amplified with specific primers for *e69wzc* (lane 1), *wzc* (lane 2) and wzc_{Trunc} (lane 3). Expected band sizes are 2225, 2326 and 2243 bp, respectively. MM - GeneRuler^{TM} DNA Ladder Mix.

At the moment, *e69wzc*, *wzc* and wzc_{Trunc} are being cloned into pBAD18-cm::*e69wza*::*e69wzb.*

4. Discussion

The main goal of this work was to generate the molecular tools (plasmids) and develop the methodology to elucidate the role of *Synechocystis* Wzc and KpsM in the synthesis of EPS. To achieve that, different approaches were followed, namely (i) the analysis of the functional partners and of Wzc by *in vitro* pull-down assays, (ii) the construction of plasmids to generate *Synechocystis* mutants expressing tagged versions of Wzc and KpsM for *in vivo* crosslink followed by pull-down assays, and (iii) generation of plasmids for the *in vivo* characterization of Wzc in the well characterized EPS production system of *E. coli* group 1 capsules.

For the identification of Wzc functional partners using purified His6-Wzc as bait, the protein was overexpressed, purified and the methodology to isolate His6-Wzc in IMAC beads was established. However, it is still necessary to optimize the purification/isolation of Wzc, reduce the unspecific binding to the IMAC beads and optimize the recovery of His6-Wzc/functional partner complexes in elution. This could be achieved by changing the salt and detergent concentrations in the buffers, increasing the number of washes and/or the concentration of imidazole in the loading and elution of His6-Wzc in IMAC. In addition, the amount of week chelators mainly present in the periplasm of *E. coli*, and its lysates that affect the binding of the target to the IMAC columns can also be reduced (Magnusdottir et al., 2009). Regarding the generation of plasmids for complementation of the *Synechocystis kpsM* knockout mutant with the gene encoding the tagged version of the protein, the inability to obtain pSEVA351::$P_{trc.x.tetO2}$::B0030*::his6-kpsM* raises several hypothesis. Thus, the cytosine insertion observed may result from a replication slippage event (Schlotterer & Tautz, 1992) or sequencing error. Also, considering that $P_{trc.x.tetO2}$ was obtained by modification of P_{trc1O} (Ferreira et al., 2018) that, in *E. coli* DH5α, is 881 times stronger compared to the reference P_{rnpB} (Huang et al., 2010), it is possible that high levels of KpsM may be detrimental to *E. coli* or cause a high metabolic burden, making it difficult to obtain transformants, although this hypothesis is less likely. It is expected that, the results obtained from the pull-downs using purified His6-Wzc and the *in vivo* crosslink of *Synechocystis* expressing His6-KpsM will complement and expand the yet limited knowledge on the network of protein interactions comprising EPS-related proteins (Sato et al., 2007; Xu et al., 2021).

It was previously suggested that cyanobacterial EPS production may imply a more complex scenario than that observed for other organisms (Pereira et al., 2015). The elucidation of this process in cyanobacteria is even more difficult due to the presence of multiple copies of several EPS-related genes (Pereira et al., 2015). Therefore, the use of a well characterized EPS production system, such as that of *E. coli* K30 E69 (Reid & Whitfield, 2005; Whitfield, 2006) can help provide important insight about the *in vivo* functional role of

the proteins. In the case of *E. coli* K30 E69, Wzc regulates both the polymerization and the translocation of the polymer events occurring at its C-terminal domain (Morona et al., 2009; Whitfield et al., 2020). These events comprise its autophosphorylation followed by its dephosphorylation by the low molecular weight tyrosine phosphatase Wzb, with the cycling between different Wzc phosphorylation states being critical for polymer production (Wugeditsch et al., 2001). Regarding *Synechocystis* Wzc, it was shown that it is important for EPS production and that it has auto-kinase and ATPase activities *in vitro* (Pereira, Santos, et al., 2019). Therefore, the complementation of the acapsular *E. coli* CWG655 mutant (devoid of *wza*, *wzb* and *wzc*) with the native *wza*, *wzb* and, *Synechocystis wzc* or wzc_{Trunc}, can help clarify the role of the C-terminal domain of *Synechocystis* Wzc. The *E. coli* native *wza* and *wzb* were successfully cloned in pBAD18-cm, and at the moment, the native *wzc* (control), *Synechocystis wzc* or wzc_{Trunc} are also being cloned. Once the plasmids are generated, they will be used to transform *E. coli* mutant CWG655 and the ability to restore capsule formation will be evaluated.

5. Conclusions and future perspectives

This work contributed to develop the methodology and generate plasmids to further characterize *Synechocystis* Wzc and KpsM, including the identification of their functional partners and/or its *in vivo* characterization in a heterologous system. However, further studies are required:

Regarding the identification of the Wzc and KpsM protein interactions, it is necessary to:

(i) Complete the identification of the proteins co-eluted with His6-Wzc in the pull-down assays.

(ii) Generate plasmids to complement *Synechocystis* Δ*wzc* with genes encoding His6-Wzc or His6-Wzc_{Trunc}, followed by the identification of their protein interactors *in vivo* by crosslink followed by pull-down assays.

(iii) Use pSEVA351::P_{psbA2*}::B0030::*his6-kpsM* to transform the *Synechocystis kpsM* knockout mutant and identify KpsM protein interactors *in vivo* by crosslink followed by pull-down assays.

To clarify the role of Wzc *in vivo* it is necessary to:

(i) Clone the genes encoding *Synechocystis* Wzc or Wzc_{Trunc} (and the native *E. coli* Wzc, as control) in pBAD18-cm::*e69wza*::*e69wzb*, transform *E. coli* CWG655 with the plasmid obtained and evaluate capsule formation in the resulting mutants.

In addition, it will be important to expand these studies to other proteins, as more components of the cyanobacterial EPS production machinery are unveiled.

6. References:

Allen, R., Rittmann, B. E., & Curtiss, R., 3rd. (2019). Axenic Biofilm Formation and Aggregation by *Synechocystis sp.* Strain PCC 6803 Are Induced by Changes in Nutrient Concentration and Require Cell Surface Structures. *Applied and environmental microbiology*, *85*(7), e02192-02118. https://doi.org/10.1128/AEM.02192-18

Anderson, S. L., & McIntosh, L. (1991). Light-activated heterotrophic growth of the cyanobacterium *Synechocystis sp.* strain PCC 6803: a blue-light-requiring process. *Journal of bacteriology*, *173*(9), 2761-2767. https://doi.org/10.1128/jb.173.9.2761-2767.1991

Bahat-Samet, E., Castro-Sowinski, S., & Okon, Y. (2004). Arabinose content of extracellular polysaccharide plays a role in cell aggregation of *Azospirillum brasilense*. *FEMS Microbiology Letters*, *237*(2), 195-203. https://doi.org/10.1111/j.1574-6968.2004.tb09696.x

Beyl, T., Louw, T. M., & Pott, R. W. M. (2019). Cyanobacterial Growth in Minimally Amended Anaerobic Digestion Effluent and Flue-Gas. *Microorganisms*, *7*(10), 428. https://www.mdpi.com/2076-2607/7/10/428

Carmichael, W. W. (1992). Cyanobacteria secondary metabolites - the cyanotoxins. *Journal of Applied Bacteriology*, *72*(6), 445-459. https://doi.org/https://doi.org/10.1111/j.1365-2672.1992.tb01858.x

De Philippis, R., Colica, G., & Micheletti, E. (2011). Exopolysaccharide-producing cyanobacteria in heavy metal removal from water: molecular basis and practical applicability of the biosorption process. *Appl Microbiol Biotechnol*, *92*(4), 697-708. https://doi.org/10.1007/s00253-011-3601-z

De Vuyst, L., & Degeest, B. (1999). Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiol Rev*, *23*(2), 153-177. https://doi.org/10.1111/j.1574-6976.1999.tb00395.x

Dittmann, E., Fewer, D. P., & Neilan, B. A. (2013). Cyanobacterial toxins: biosynthetic routes and evolutionary roots. *FEMS Microbiology Reviews*, *37*(1), 23-43. https://doi.org/10.1111/j.1574-6976.2012.12000.x

Ferreira, E. A., Pacheco, C. C., Pinto, F., Pereira, J., Lamosa, P., Oliveira, P., Kirov, B., Jaramillo, A., & Tamagnini, P. (2018). Expanding the toolbox for *Synechocystis sp.* PCC 6803: validation of replicative vectors and characterization of a novel set of promoters. *Synthetic Biology*, *3*(1). https://doi.org/10.1093/synbio/ysy014

Fisher, M. L., Allen, R., Luo, Y., & Curtiss, R., 3rd. (2013). Export of extracellular polysaccharides modulates adherence of the *Cyanobacterium Synechocystis*. *PloS one*, *8*(9), e74514-e74514. https://doi.org/10.1371/journal.pone.0074514

García-Meza, J. V., Barrangue, C., & Admiraal, W. (2005). Biofilm formation by algae as a mechanism for surviving on mine tailings. *Environ Toxicol Chem*, *24*(3), 573-581. https://doi.org/10.1897/04-064r.1

Gromova, I., & Celis, J. E. (2006). Chapter 27 - Protein Detection in Gels by Silver Staining: A Procedure Compatible with Mass Spectrometry. In J. E. Celis (Ed.), *Cell Biology (Third Edition)* (pp. 219-223). Academic Press. https://doi.org/https://doi.org/10.1016/B978-012164730-8/50212-4

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *Journal of bacteriology*, *177*(14), 4121-4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995

Hagemann, M., & Hess, W. R. (2018). Systems and synthetic biology for the biotechnological application of cyanobacteria. *Curr Opin Biotechnol*, *49*, 94-99. https://doi.org/10.1016/j.copbio.2017.07.008

Hammer, U. T. (1986). *Saline lake ecosystems of the world*. Dr. W. Junk Publishers.

Huang, H. H., Camsund, D., Lindblad, P., & Heidorn, T. (2010). Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res*, *38*(8), 2577-2593. https://doi.org/10.1093/nar/gkq164

Ikeuchi, M., & Tabata, S. (2001). *Synechocystis sp.* PCC 6803—A useful tool in the study of the genetics of cyanobacteria. *Photosynth Res.*, *70*, 73-83. https://doi.org/10.1023/A:1013887908680

Jittawuttipoka, T., Planchon, M., Spalla, O., Benzerara, K., Guyot, F., Cassier-Chauvat, C., & Chauvat, F. (2013). Multidisciplinary evidences that *Synechocystis* PCC 6803 exopolysaccharides operate in cell sedimentation and protection against salt and metal stresses. *PloS one*, *8*(2), e55564-e55564. https://doi.org/10.1371/journal.pone.0055564

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., & Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp.* strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res*, *3*(3), 109-136. https://doi.org/10.1093/dnares/3.3.109

Kannenberg, E. L., & Brewin, N. J. (1994). Host-plant invasion by Rhizobium: the role of cell-surface components. *Trends in microbiology*, *2*(8), 277-283. https://doi.org/10.1016/0966-842x(94)90004-3

Kasting, J. F. (2013). What caused the rise of atmospheric O_2? *Chemical Geology*, *362*, 13-25. https://doi.org/https://doi.org/10.1016/j.chemgeo.2013.05.039

Kehr, J.-C., & Dittmann, E. (2015). Biosynthesis and function of extracellular glycans in cyanobacteria. *Life (Basel, Switzerland)*, *5*(1), 164-180. https://doi.org/10.3390/life5010164

Komárek, J. (2013). *Cyanoprokaryota: 3rd Part: Heterocystous Genera. In: Büdel, B., Gärtner, G., Krienitz, L. and Schagerl, M., Eds., Süßwasserflora von Mitteleuropa, Bd. 19 (3)*. Springer Spektrum.

Komárek, J. i. A. K. (1999). *Cyanoprokaryota. 1 Teil, 1 Teil*. G. Fischer.

Kufryk, G., Sachet, M., Schmetterer, G., & Vermaas, W. (2002). Transformation of the cyanobacterium *Synechocystis sp.* PCC 6803 as a tool for genetic mapping: Optimization of efficiency. *FEMS Microbiology Letters*, *206*, 215-219. https://doi.org/10.1016/S0378-1097(01)00540-7

Kumar, J., Singh, D., Tyagi, M. B., & Kumar, A. (2019). Cyanobacteria: Applications in Biotechnology. In *Cyanobacteria* (pp. 327-346). https://doi.org/10.1016/b978-0-12-814667-5.00016-7

Lau, N.-S., Matsui, M., & Abdullah, A. A.-A. (2015). Cyanobacteria: Photoautotrophic Microbial Factories for the Sustainable Synthesis of Industrial Products. *BioMed Research International*, *2015*, 754934. https://doi.org/10.1155/2015/754934

Li, X., Luo, K., Ren, J., Wang, X., & Mu, Q. (2017). Characterisation of extracellular polymeric substances from different cyanobacterial species and their influence on biocalcification processes. *Environmental Chemistry*, *14*. https://doi.org/10.1071/EN17068

Magnusdottir, A., Johansson, I., Dahlgren, L.-G., Nordlund, P., & Berglund, H. (2009). Enabling IMAC purification of low abundance recombinant proteins from E. coli lysates. *Nature Methods*, *6*(7), 477-478. https://doi.org/10.1038/nmeth0709-477

Meeks, J. C., & Castenholz, R. W. (1971). Growth and photosynthesis in an extreme thermophile, *Synechococcus lividus* (Cyanophyta). *Arch Mikrobiol*, *78*(1), 25-41. https://doi.org/10.1007/bf00409086

Miller, J. H. (1972). Experiments in molecular genetics. Cold sprong harbor laboratory, *Cold spring harbor*, New York

Mitschke, J., Georg, J., Scholz, I., Sharma, C. M., Dienst, D., Bantscheff, J., Voß, B., Steglich, C., Wilde, A., Vogel, J., & Hess, W. R. (2011). An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis sp.* PCC 6803. *Proceedings of the National Academy of Sciences*, *108*(5), 2124-2129. https://doi.org/10.1073/pnas.1015154108

Morona, R., Purins, L., Tocilj, A., Matte, A., & Cygler, M. (2009). Sequence-structure relationships in polysaccharide co-polymerase (PCP) proteins. *Trends Biochem Sci*, *34*(2), 78-84. https://doi.org/10.1016/j.tibs.2008.11.001

Myszka, K., & Czaczyk, K. (2009). Characterization of Adhesive Exopolysaccharide (EPS) Produced by *Pseudomonas aeruginosa* Under Starvation Conditions. *Current microbiology*, *58*, 541-546. https://doi.org/10.1007/s00284-009-9365-3

Nouha, K., Kumar, R. S., Balasubramanian, S., & Tyagi, R. D. (2018). Critical review of EPS production, synthesis and composition for sludge flocculation. *Journal of Environmental Sciences*, *66*, 225-245. https://doi.org/https://doi.org/10.1016/j.jes.2017.05.020

Okajima, M. K., Sornkamnerd, S., & Kaneko, T. (2018). Development of Functional Bionanocomposites Using Cyanobacterial Polysaccharides. *The Chemical Record*, *18*(7-8), 1167-1177. https://doi.org/https://doi.org/10.1002/tcr.201700074

Parikh, A., & Madamwar, D. (2006). Partial characterization of extracellular polysaccharides from cyanobacteria. *Bioresour Technol*, *97*(15), 1822-1827. https://doi.org/10.1016/j.biortech.2005.09.008

Pereira, S., Mota, R., Vieira, C. P., Vieira, J., & Tamagnini, P. (2015). Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. *5*, 14835. https://doi.org/10.1038/srep14835

Pereira, S., Santos, M., Leite, J., Flores, C., Eisfeld, C., Buttel, Z., Mota, R., Rossi, F., De Philippis, R., Gales, L., Morais-Cabral, J., & Tamagnini, P. (2019). The role of the tyrosine kinase Wzc (Sll0923) and the phosphatase Wzb (Slr0328) in the production of extracellular polymeric substances (EPS) by *Synechocystis* PCC 6803. *MicrobiologyOpen*. https://doi.org/10.1002/mbo3.753

Pereira, S., Sousa, A., Santos, M., Araújo, M., Serôdio, F., Granja, P., & Tamagnini, P. (2019). Strategies to Obtain Designer Polymers Based on Cyanobacterial Extracellular Polymeric Substances (EPS). *Int J Mol Sci*, *20*(22). https://doi.org/10.3390/ijms20225693

Pereira, S., Zille, A., Micheletti, E., Moradas-Ferreira, P., De Philippis, R., & Tamagnini, P. (2009). Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. *FEMS Microbiol Rev*, *33*(5), 917-941. https://doi.org/10.1111/j.1574-6976.2009.00183.x

Petit, C., Grill, J., Maazouzi, N., & Marczak, R. (2004). Regulation of polysaccharide formation by *Streptococcus thermophilus* in batch and fed-batch cultures. *Applied Microbiology and Biotechnology*, *36*, 216-221.

Rai, A., Pearson, L., & Kumar, A. (2013). Stress Biology of Cyanobacteria: Molecular Mechanisms to Cellular Responses. In.

Reid, A. N., & Whitfield, C. (2005). functional analysis of conserved gene products involved in assembly of *Escherichia coli* capsules and exopolysaccharides: evidence for molecular recognition between Wza and Wzc for colanic acid biosynthesis. *Journal of bacteriology*, *187*(15), 5470-5481. https://doi.org/10.1128/jb.187.15.5470-5481.2005

Rippka, R. (1988). Isolation and purification of cyanobacteria. *Methods Enzymol*, *167*, 3-27. https://doi.org/10.1016/0076-6879(88)67004-2

Rossi, F., & De Philippis, R. (2015). Role of cyanobacterial exopolysaccharides in phototrophic biofilms and in complex microbial mats. *Life (Basel, Switzerland)*, *5*(2), 1218-1238. https://doi.org/10.3390/life5021218

Salama, Y., Chennaoui, M., Sylla, A., Mountadar, M., Rihani, M., & Assobhei, O. (2016). Characterization, structure, and function of extracellular polymeric substances (EPS) of microbial biofilm in biological wastewater treatment systems: a review. *Desalination and Water Treatment*, *57*(35), 16220-16237. https://doi.org/10.1080/19443994.2015.1077739

Sambrook, J. R. D. W. (2001). *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press.

Santos, M., Pereira, S. B., Flores, C., Príncipe, C., Couto, N., Karunakaran, E., Cravo, S. M., Oliveira, P., & Tamagnini, P. (2021). Absence of KpsM (Slr0977) Impairs the Secretion of Extracellular Polymeric Substances (EPS) and Impacts Carbon Fluxes in *Synechocystis sp.* PCC 6803. *mSphere*, *6*(1). https://doi.org/10.1128/mSphere.00003-21

Sarma, T. A. (2013). *Handbook of Cyanobacteria* ((1st ed.) ed.). CRC Press. https://doi.org/https://doi.org/10.1201/b14316

Sato, S., Shimoda, Y., Muraki, A., Kohara, M., Nakamura, Y., & Tabata, S. (2007). A large-scale protein protein interaction analysis in *Synechocystis sp.* PCC 6803. *DNA Res*, *14*(5), 207-216. https://doi.org/10.1093/dnares/dsm021

Savakis, P., & Hellingwerf, K. J. (2015). Engineering cyanobacteria for direct biofuel production from CO_2. *Curr Opin Biotechnol*, *33*, 8-14. https://doi.org/10.1016/j.copbio.2014.09.007

Schlotterer & Tautz, (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, vol 20, 211-215 https://doi.org/10.1093/nar/20.2.211

Schmid, J., Sieber, V., & Rehm, B. (2015). Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies [Review]. *Frontiers in microbiology*, *6*(496). https://doi.org/10.3389/fmicb.2015.00496

Shah, S. A. A., Akhter, N., Auckloo, B. N., Khan, I., Lu, Y., Wang, K., Wu, B., & Guo, Y.-W. (2017). Structural Diversity, Biological Properties and Applications of Natural Products from Cyanobacteria. A Review. *Marine drugs*, *15*(11), 354. https://www.mdpi.com/1660-3397/15/11/354

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent, T., Goodwin, L., Nolan, M., Davenport, K. W., Han, C. S., Rubin, E. M., Eisen, J. A., Woyke, T., Gugger, M., & Kerfeld, C. A. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences*, *110*(3), 1053-1058. https://doi.org/10.1073/pnas.1217107110

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., Páez-Espino, A. D., Durante-Rodríguez, G., Kim, J., Nikel, P. I., Platero, R., & de Lorenzo, V. (2013). The Standard European Vector Architecture (SEVA): a

coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res*, *41*(Database issue), D666-675. https://doi.org/10.1093/nar/gks1119

Singh, P., Khan, A., & Srivastava, A. (2020). Heterocyst and akinete differentiation in cyanobacteria: a view toward cyanobacterial symbiosis. In (pp. 235-248). https://doi.org/10.1016/B978-0-12-819311-2.00016-4

Stanier, R. Y., Kunisawa, R., Mandel, M., & Cohen-Bazire, G. (1971). Purification and properties of unicellular blue-green algae (order *Chroococcales*). *Bacteriological reviews*, *35*(2), 171-205. https://doi.org/10.1128/br.35.2.171-205.1971

Suresh Kumar, A., Mody, K., & Jha, B. (2007). Bacterial exopolysaccharides – a perception. *Journal of Basic Microbiology*, *47*(2), 103-117. https://doi.org/https://doi.org/10.1002/jobm.200610203

Sutherland, I. W. (2001). Extracellular Polysaccharides. In *Biotechnology Set* (pp. 613-657). https://doi.org/https://doi.org/10.1002/9783527620999.ch16f

Vaningelgem, F., Zamfir, M., Adriany, T., & De Vuyst, L. (2004). Fermentation conditions affecting the bacterial growth and exopolysaccharide production by *Streptococcus thermophilus* ST 111 in milk-based medium. *Journal of Applied Microbiology*, *97*(6), 1257-1273. https://doi.org/https://doi.org/10.1111/j.1365-2672.2004.02418.x

Whitfield, C. (2006). Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli. Annu Rev Biochem*, *75*, 39-68. https://doi.org/10.1146/annurev.biochem.75.103004.142545

Whitfield, C., & Larue, K. (2008). Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. *Nat Struct Mol Biol*, *15*(2), 121-123. https://doi.org/10.1038/nsmb0208-121

Whitfield, C., Wear, S. S., & Sande, C. (2020). Assembly of Bacterial Capsular Polysaccharides and Exopolysaccharides. *Annual Review of Microbiology*, *74*(1), 521-543. https://doi.org/10.1146/annurev-micro-011420-075607

Whitton, B., & Potts, M. (2000). Introduction to the cyanobacteria, in the Ecology of Cyanobacteria: Their Diversity in Time and Space (Eds B. *A. Whitton, M. Potts*, 1-11.

Willis, L. M., & Whitfield, C. (2013). Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydrate Research*, *378*, 35-44. https://doi.org/https://doi.org/10.1016/j.carres.2013.05.007

Wugeditsch, T., Paiment, A., Hocking, J., Drummelsmith, J., Forrester, C., & Whitfield, C. (2001). Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *J Biol Chem*, *276*(4), 2361-2371. https://doi.org/10.1074/jbc.M009092200

Xu, C., Wang, B., Yang, L., Zhongming Hu, L., Yi, L., Wang, Y., Chen, S., Emili, A., & Wan, C. (2021). Global Landscape of Native Protein Complexes in *Synechocystis sp.* PCC 6803. *Genomics Proteomics Bioinformatics*. https://doi.org/10.1016/j.gpb.2020.06.020

Zakhia, F., Jungblut, A.-D., Taton, A., Vincent, W. F., & Wilmotte, A. (2008). Cyanobacteria in Cold Ecosystems. In R. Margesin, F. Schinner, J.-C. Marx, & C. Gerday (Eds.), *Psychrophiles: from Biodiversity to Biotechnology* (pp. 121-135). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74335-4_8

7. Annexes:

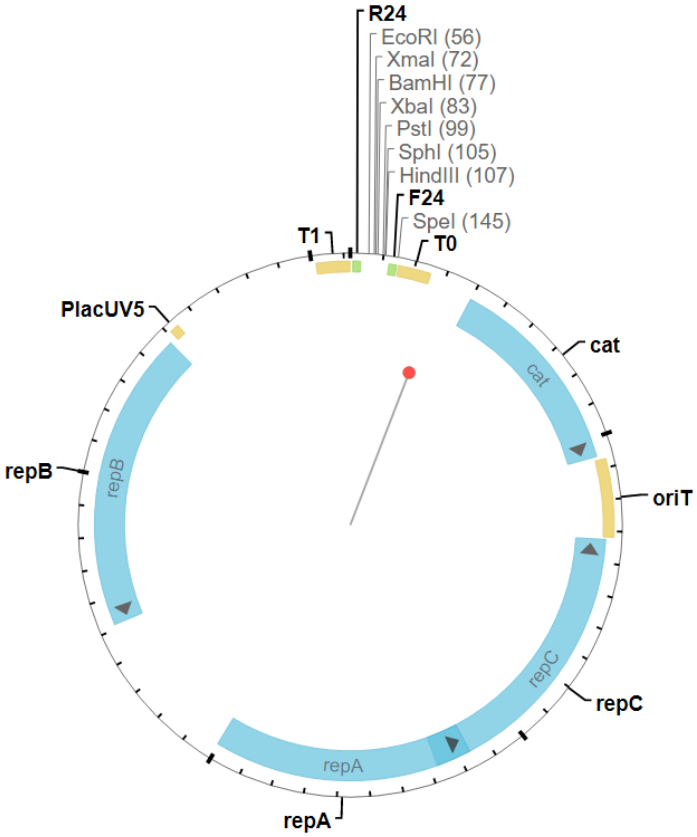

Figure A – pSEVA351 plasmid map. Source: http://beta.labgeni.us/registries/SEVA/pSEVA351/

>pGEM-Teasy::*his6-kpsM*

gggcgaattgggcccgacgtcgcatgctcccggccgccatggcggccgcgggaattcgattgcctgcagatgaga
ggatcgcatcaccatcaccatcacggatccaaaacttcccccccagaactgattattgaagcaggacgcacggag
cgtcagtattggcaagacctatggcgttaccgggaattgttttacaccctggcttggcgggacattgcggtacgg
tacaaacaaacggcgatcggtatagcttgggccttaatccggccatttttgaccatggtggtgtttacggtggta
tttggtaagttggctaatttaccttcggaggggtgccctatcccattctggtgtttgcgggaatgttgccctgg
cagtttttttccacttcccttagttccgccagcgatagtctaattgccaatgccaatctaatttctaaggtgtat
tttcctcgcttagtggtgcctaccagtgccgtggtgactagctttgttgatttttaatttctgggatgattatg
ttggggctgatggcttggtataatttcttgcccagttggcatgtgattacattgcctttcttcattttgattgcc
tttatggcttccatgggagcagggttatggctttgttccctcaatgtcaaataccgagattttcgctacattgtg
ccattcattgtccaatttggtttgtacatttccccggtgggttttagtagtaatgtggtgccggaaaaatggcga
ttgctctattccattaacccgatggtgagtgtaattgatggttttcgttgggcgattttgggggggagaatcaact
atatttctgccgggtttttttgttgtctttactgttggtgatcattatttttataacaggaattctctattttcgt
aagatggaacgcaccttcgctgatgtgatttaaccatggcaacaaacaaatggttttatgagctattgaaggatt
ttaaatacagttgcctgaattgggtttgaccactagtgcaatcactagtgaattcgcggccgcctgcaggtcgac
catatgggagagctcccaacgcgttggatgcatagcttgagtattctatagtgtcacctaaatagcttggcgtaa
tcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgag

Figure B – partial sequence (theoretical) of the pGEM-Teasy::*his6-kpsM*. The sequence encoding the 6x histidine tag is highlighted in yellow; *his6-kpsM* sequence highlighted in grey; the start and stop codon are highlighted in green and red, respectively.

>pSEVA351::P*trc.x.TetO2*::B0030::*his6-kpsM*

acaaacaacagataaaacgaaaggcccagtctttcgactgagcctttcgttttatttgatgcctttaattaaagc
ggataacaatttcacacaggaggccgcctaggccgcggccgcgcgaattcgagctcggtacccggggatcctcta
gaggagctgttgacaattaatcatccggctcgtataatgtgtggatccccgtcagtgacggagattcacacatac
tagagattaaagaggagaaatactagaatgagaggatcgcatcaccatcaccatcacggatccaaaacttccccc
ccagaactgattattgaagcaggacgcacggagcgtcagtattggcaagacctatggcgttaccgggaattgttt
tacaccctggcttggcgggacattgcggtacggtacaaacaaacggcgatcggtatagcttgggccttaatccgg
ccatttttgaccatggtggtgtttacggtggtatttggtaagttggctaatttaccttcggaggggtgccctat
cccattctggtgtttgcgggaatgttgccctggcagttttttccacttcccttagttccgccagcgatagtcta
attgccaatgccaatctaatttctaaggtgtattttcctcgcttagtggtgcctaccagtgccgtggtgactagc
tttgttgatttttaatttctgggatgattatgttggggctgatggcttggtataatttcttgcccagttggcat
gtgattacattgccttttcttcattttgattgcctttatggcttccatgggagcagggttatggctttgttccctc
aatgtcaaataccgagattttcgctacattgtgccattcattgtccaatttggtttgtacatttccccggtgggt
tttagtagtaatgtggtgccggaaaaatggcgattgctctattccattaacccgatggtgagtgtaattgatggt
tttcgttgggcgattttgggggggagaatcaactatatttctgccgggtttttgttgtctttactgttggtgatc
attattttataacaggaattctctattttcgtaagatggaacgcaccttcgctgatgtgatttaaccatggcaa
caaacaaatggttttatgagctattgaaggattttaaatacagttgcctgaattgggtttgaccactagtagcgg
ccgctgcaggcatgcaagcttgcggccgcgtcgtgactgggaaaaccctggcgactagtcttggactcctgttga
tagatccagtaatga

Figure C – partial sequence (theoretical) of pSEVA351::P*trc.x.TetO2*::B0030::*his6-kpsM.* The sequence encoding the 6x histidine tag is highlighted in yellow; P*trc.x.TetO2* in pink; B0030 in brown; *his6-kpsM* in grey; restriction sites for *Xba*I, *Spe*I and *Pst*I are in purple, cyan and blue, respectively; the start and stop codons are in green and red, respectively.

>pSEVA351::P*psbA2::B0030::*his6-kpsM***

caaacaacagataaaacgaaaggcccagtctttcgactgagcctttcgttttatttgatgcctttaattaaagcg
gataacaatttcacacaggaggccgcctaggccgcggccgcgcgaattcgcggccgcttctagagagctttacaa
aactctcattaatcctttagactaagtttagtcagttccaatctgtactagagattaaagaggagaaatactaga
atgagaggatcgcatcaccatcaccatcacggatccaaaacttcccccccagaactgattattgaagcaggacgc
acggagcgtcagtattggcaagacctatggcgttaccgggaattgttttacaccctggcttggcgggacattgcg
gtacggtacaaacaaacggcgatcggtatagcttgggccttaatccggccatttttgaccatggtggtgtttacg
gtggtatttggtaagttggctaatttaccttcggaggggtgccctatcccattctggtgtttgcgggaatgttg
ccctggcagttttttccacttcccttagttccgccagcgatagtctaattgccaatgccaatctaatttctaag
gtgtatttcctcgcttagtggtgcctaccagtgccgtggtgactagctttgttgatttttaatttctgggatg
attatgttggggctgatggcttggtataatttcttgcccagttggcatgtgattacattgcctttcttcattttg
attgcctttatggcttccatgggagcagggttatggctttgttccctcaatgtcaaataccgagattttcgctac
attgtgccattcattgtccaatttggtttgtacatttccccggtgggttttagtagtaatgtggtgccggaaaaa
tggcgattgctctattccattaacccgatggtgagtgtaattgatggttttcgttgggcgattttgggggggaga
tcaactatatttctgccgggtttttgttgtctttactgttggtgatcattatttttataacaggaattctctat
tttcgtaagatggaacgcaccttcgctgatgtgatttaatactagtagcggccgctgcaggcatgcaagcttgcg
gccgcgtcgtgactgggaaaaccctggcgactagtcttggactcctgttgatagatccagtaatga

Figure D – partial sequence (theoretical) of the pSEVA351::P*psbA2**::B0030::*his6-kpsM.* The sequence encoding the 6x histidine tag is highlighted in yellow; P*psbA2** in pink; B0030 in brown; *his6-kpsM* in grey; restriction sites for *Eco*RI, *Xba*I, *Spe*I and *Pst*I are in marron, purple, cyan and blue, respectively; the start and stop codons are in green and red, respectively.

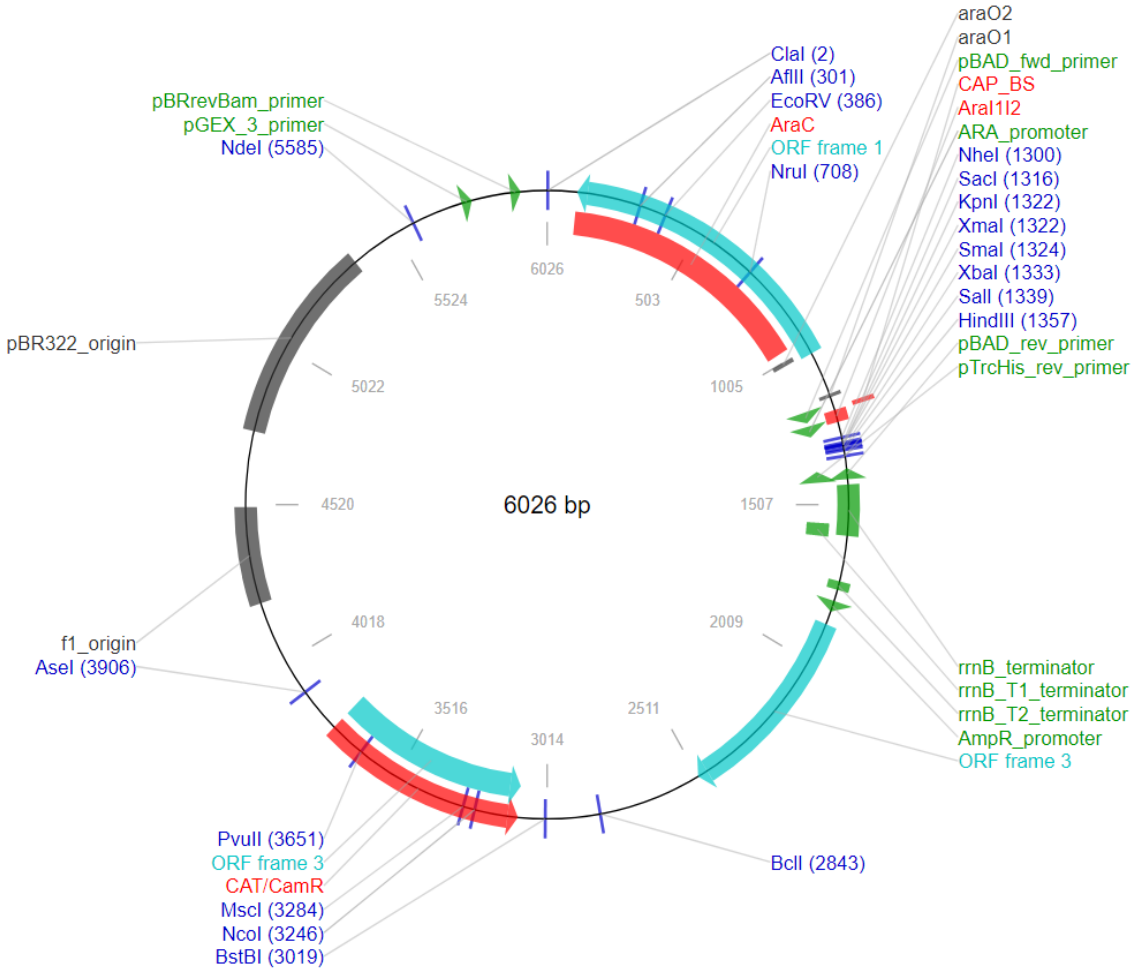

Figure E – pBAD18-cm plasmid map.Source:https://www.addgene.org/vector-database/1840/?gclid=EAIaIQobChMIleKQ9e-z9AIVBwOLCh3jbggoEAAYASAAEgLPQ_D_BwE

>pBAD18-cm::*e69wza*::*e69wzb*

ctactgtttctccatacccgttttttgggctagcgaattcgagctcggtacccggggatcctctagacattaaa
gaggagaaatactagatgaagaaaaaacttgttagattttcggcattagcattggcaattgggtttttatcaggt
tgtacaatcatccctggtcagggattgaacagtctgcgtaaaaacgtggtagaactcccggatagcgactacgat
ttggataagctggtgaatgtttatcctatgacgccaggtctaatagatcaactccgtccggaaccggttatagca
cgctcgaatcctcagttggataacctattaaaaagttatgagtatcgcattggtgtgggtgatgtgctcatggtt
acggtatgggatcatccggaattaacgacacctgctggtcaatatcgtagtgcaagtgatacgggtaactgggta
aattctgacggaactatttttatccttacataggtaaggttcaggtagctggaaaaacagtaagtcaagtacga
caagatattacaagccgattaactacatatattgaaagccctcaagttgatgtcagcatagctgcattccggtca
caaaaggtttatgtaactggtgaagttgcaaactctggaaaacaggccattacaaatattcccctaactgtgatg
gatgctatcaatgcggcaggggggcttgcggctgatgctgactggagaaacgttgttcttactcataacggtaaa
gatacaaagatttcattatatgcactaatgcagaaaggagatctaacccagaatcatttattataccatggcgat
atattattcattccaagtaatgatgatctcaaagtatttgtcatgggagaggtcggtaagcaaagtacattgaaa
atggaccgcagcggtatgacccttgcagaggcgcttggtaatgctgagggtatttctcaagaaatgagcgatgct
acagggatatttgttgtacgtcagttaaaaggtgataggacaggaaaaatagcagatatatatcaattgaatgct
caggatgcttcagcaatggtcttaggtacagaatttcaactacaaccttatgacattgtatatgtaacgactgca
cctcttgttcgttggaatcgtgttatatcacagctcgtacctacaatatcaggtgttcatgatatgacggaaaca
gtacgatacattaagagatggccaaactaagtcgacattaaagaggagaaatactagatggccaaactaatgttt

gattcaattttagtgatctgcacaggaaatatctgccgttctccaattggtgagcgtttattaagacggctatta
ccaagcaaaaagattaattccgctggggttggggcattggttgatcatgcagcagatgaatccgcaattcgcgtc
gctgaaaaaaatggtctttgtctcaaaggccaccgtgggacaaaatttacctctgcattagctcgacagtatgat
cttttactcgtgatggaatattctcatctagaacaaattagccggatagcacctgaagcacgtggtaaaactatg
ctgtttggacactggcttgatagtaaagaaatcccagacccataccgtatgagtgatgaagcattcgactcggta
tatcaattgcttgaacaagcaagcaagcgttgggctgagaaattaggtgaataagcatgcaagcttggctgtttt
ggcggatgagagaagatttt

Figure F - partial sequence (theoretical) of the pBAD18-cm::*e69wza*::*e69wzb*. The sequence of the RBS is highlighted in yellow; *e69wza* in dark grey; *e69wzb* in light grey; restriction sites for *Xba*I, *Sal*I and *Sph*I are in purple, brown and cyan, respectively; the start and stop codons are in green and red, respectively. The primer sequences pBAD18_5', FC2E69wzbF, FC2E69wzbR and pBAD18_3' are underlined in cyan, orange, yellow and pink, respectively.

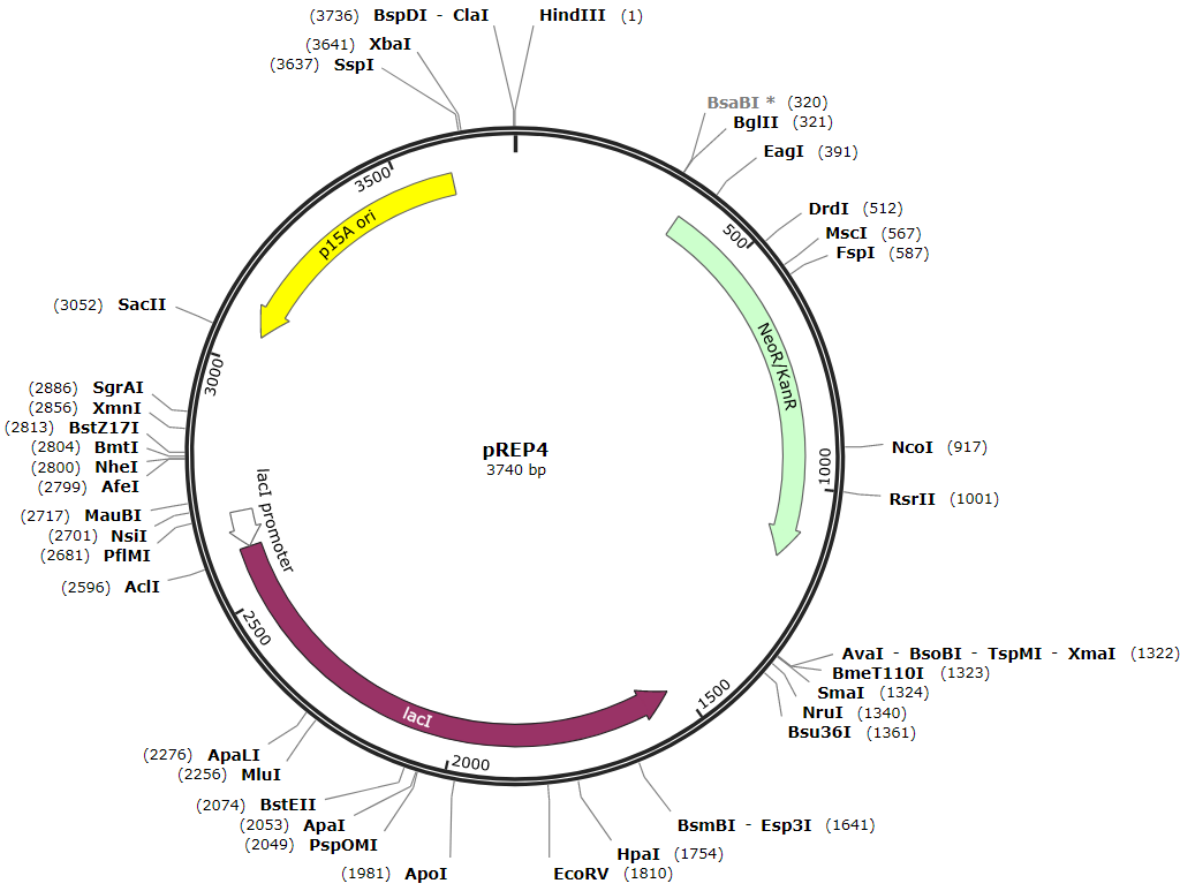

Figure G – pREP4 plasmid map. Source: https://www.snapgene.com/resources/plasmid-files/?set=qiagen_vectors&plasmid=pREP4

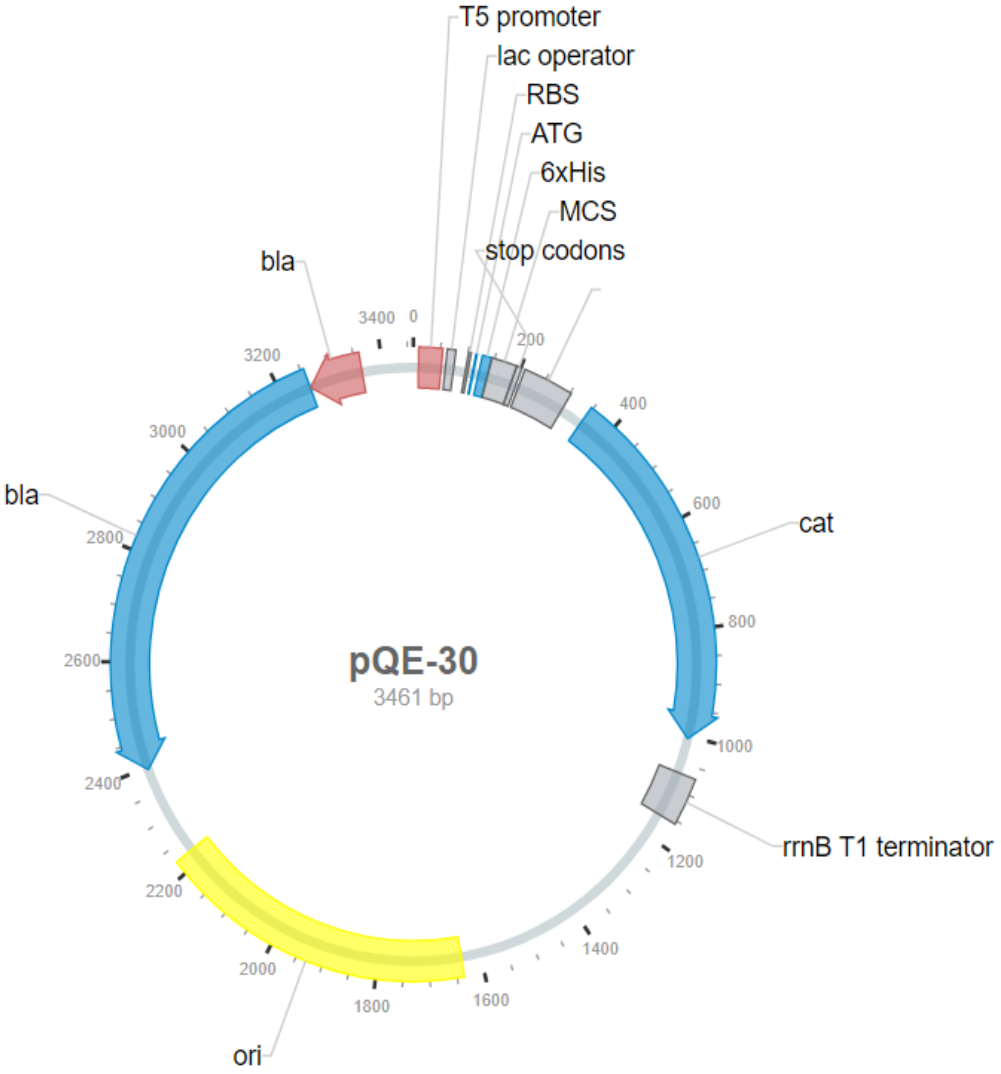

Figure H – pQE-30 plasmid map. Source: https://www.novoprolabs.com/vector/V10794

>pQE_30::*his6-wzc*

cccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacga
ggccctttcgtcttcacctcgagaaatcataaaaaatttatttgctttgtgagcggataacaattatataatagatt
caattgtgagcggataacaatttcacacagaattcattaaagaggagaaattaactatgagaggatcgcatcacc
atcaccatcacggatccacctacagttacccagaattgccccaaggtagccccccccctccgctagttggagcat
acggtgctgaagaagtggactctttttcccttggtaacctgacaggtgtattgaggaggagatggtggattgttg
tcctggtggccctagcggtgggaggcttgaacctcaagcggcaacttgataagccgccggtttttcaatctagtt
tccaattactggtttcccccccccgagaatgaaacggttaatcccctggcccaactccaaggttcgggactgctgg
gcagccttccccggggcaaagaatatttgaacacccagtcggaaattctacaaagtaatattctattggagccgg
tctggagccaaatttatgacgccctccccccagaggagcaaatttcctatagtagttttctggctagcctgtcgg
tgatggtggtgccggacacaggcattgttgagatcagctaccggggcaatgaccaagaaagggtcaaggcggttt
tggaggctctagctgaacaatatttagactacacccgtagagagcaagccagacaggagagtgaaaaactccgct
ttgttaatgaccagttgcccaaattccaagaccgggtggcagagttgcaagcgcaaatgttggcattgcagcagc
aatataagtttttgaccctagccagacagcagaaaatctcagctcacgcctcaacgaaattaatgtgctccggc
aaaatttgtcgatcaaaatccaacaactaaccgccaaacgagatgccctggaacaaaaaacgggcatagataaca
atagtgcctttggcttgagctccctgggacaatcccctgtctatgcaggcttagttagtcggctcaacgaaatca
acctcaaattggcggaagctagcaccatctacactggcaatagcccccagatggagcagttgcaagaagaaaagg
ataacgttttgcagttgatcagagccgaagccgaaaaagaaaaggggacgctcagcaatggagaattggatcagg
tcctcattgccacagacgtgggggggcccaaggcagaagctttgcaggaacttgccagtgtgggcattgagctag
atattctcaacattgaactggaaagtatcattgggtcggaagcggaactgagaaatcaattggaaaatttcacca
ccattgctggcgagtttgccagtctccaacaggaattggagcttcaaagaagtggtttaacccaattagtgacca
caaggcaggcactggaagtggaggtcgcgaggagctttgtaccctggcgcttagtttccgccattaccctacccg
atacaccgattaataccctgcccagggatattttactttctgtcatgctgggacttttggccgggggaggcgcgg
ccctattggttgataagctagatcctgcctatcacagtgtcgaagatttgagagctagtcagcaccacaccatcc
tagggtacattcccttggaaaaggatttgcgagccagtattaagtatggctcccctctgcccaatgaggcaatgc
aagagtcctacgctaggctctattccaacctgtttttcctcaagcgcaaacgccagtgccattcctttgtggtta
cctcggctgagtccggtgatggaaaatctaccacggctttttttcctggcccaagctgccgctaagttggggcaaa
aagtgctgctggtggatggcgatcgctattttcctcaaaaggagtcctggttgaagttagctgaaattaccggct
tcggaggcgaaaacaccacccccgatggtaatggcgtcctcggcagtctggtggccaactctaatggtcacaatg
gggatttacccgaaccttgggcaagaatctgttctacttcaaagtccaagatgacaccatgaccccggagcagt
tggtttctgcttcccaaaattttgtcgtcaaaatgaatcagtggaaagagactttcgacctaattttgatcgata
cgccccccattttgggtctaacggattcccgtctgattgcagatcaaaccgatggtctggtggtggtggtgcgcc
tcaataagacccgcaaggatagcatcaaggaagcttttcgggaattggcgttggctgatctcaacgttatcggaa
ttgtggccaatgctatcacttctacgtccggcggttatggctactactatggtcgttactacaataaccgttact
acgatcgtcaaaaagttgcccaggcagacaattagctgcagccaagcttaattagctgagcttggactcctgttg
atagatccagtaatgacctcagaac

Figure J - partial sequence (theoretical) of the pQE30::*his6-wzc* construct. The sequence encoding the 6x histidine tag is highlighted in yellow; *his6-wzc* sequence highlighted in grey; the start and stop codon are highlighted in green and red, respectively.