# Improving the production of the heterologous compatible solute betaine in *Synechocystis*-based chassis

David Luís de Faria Rodrigues

Dissertation for the Degree of Master in Bioengineering
Specialization in Molecular Bioengineering

Supervisor: Dr. Catarina C. Pacheco, Ph.D.[a,b,c]
Co-supervisor: Prof. Paula Tamagnini, Ph.D.[a,b,c]

[a] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Rua Alfredo Allen, 208, 4200-135 Porto, Portugal.
[b] IBMC - Instituto de Biologia Molecular e Celular, Universidade do Porto, Rua Alfredo Allen, 208, 4200-135 Porto, Portugal.
[c] FCUP - Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

September 15th, 2023

“But success **shall** crown my endeavours. Wherefore not? Thus far I have gone, tracing a secure way over the pathless seas, the very stars themselves being witnesses and testimonies of my triumph. Why not still proceed over the untamed yet obedient element? What can stop the determined heart and resolved will of man?”

Mary Wollstonecraft Shelley, *Frankenstein*

This work received funding from the European Union’s
Horizon-EIC-2021-PathFinderChallenge through the project PhotoSynH2
(Project 101070948).

*This page was intentionally left blank*

# Resumo

Os solutos compatíveis (CS) são compostos de baixo peso molecular que são acumulados intracelularmente para manter a homeostasia celular em condições de crescimento desfavoráveis, como alta salinidade ou temperatura. Os CS formam um invólucro de hidratação em torno de macromoléculas-chave, como as proteínas, permitindo a sua estabilização e evitando a desnaturação e a perda de função. Estas propriedades protetoras conferem aos CS, como a glicina betaína (GB), um elevado valor comercial com aplicações em formulações para alimentação animal, cosmética e farmacêutica. Atualmente, os métodos de produção de GB dividem-se entre a síntese química, que requer produtos químicos perigosos e com uma elevada pegada de carbono, e a sua extração a partir de *sugar beets* (*Beta vulgaris*), que é mais ecológico e seguro, mas tem um fraco rendimento. Assim, a produção de GB requer uma melhoria da eficiência que está a ser abordada através da engenharia de microrganismos para produzir este soluto. As cianobactérias, sendo micróbios fotoautotróficos, são procuradas para este fim, dada a sua rápida taxa de crescimento e elevada produtividade (em comparação com plantas), e requisitos nutricionais simples em comparação com as bactérias heterotróficas, utilizando luz, dióxido de carbono e água como fontes de energia, carbono e eletrões, respetivamente. Entre as cianobactérias, uma das principais candidatas é *Synechocystis* sp. PCC 6803 (*Syn*6803), devido à sua facilidade de engenharia genética, às ferramentas moleculares disponíveis e aos modelos metabólicos à escala do genoma para efetuar simulações. Em trabalhos anteriores, foi introduzido nesta cianobactéria um módulo sintético para a produção de GB (Ahbet), permitindo a síntese deste CS e melhorando o seu crescimento em condições de salinidade. Estes dados também sugerem que o carbono é redirecionado do reservatório de glicogénio para a produção de CS e substâncias poliméricas extracelulares (EPS). Neste projeto, foi caracterizado o duplo mutante *Syn*6803 Δ*ggpS*Δ*slr0977*, deficiente no CS nativo glucosilglicerol e putativamente deficiente na produção de EPS, com e sem o módulo Ahbet. O crescimento destes mutantes foi semelhante na ausência ou presença de 3% **NaCl**, o que sugere que a introdução do módulo Ahbet não impôs uma sobrecarga metabólica. Além disso, foi observada uma diminuição do glicogénio e um aumento dos polissacarídeos capsulares (CPS) em condições de salinidade. A avaliação da produção de CS nos mutantes Δ*ggpS*Δ*slr0977* e Δ*ggpS*Δ*slr0977* Ahbet mostra que este último é incapaz de produzir GB, possivelmente devido a mutações no módulo Ahbet. No entanto, os resultados obtidos até agora revelaram que a deleção do gene *slr0977* não resulta no comprometimento da produção de EPS nas condições testadas, em contraste com trabalhos anteriores realizados na ausência de sal e sob maior intensidade de luz. Por conseguinte, é necessário compreender as alterações que ocorrem no chassis *Syn*6803 utilizado para a produção de GB, através da aplicação de abordagens transcriptómicas, proteómicas e metabolómicas. Este conhecimento representará um passo em frente para a engenharia racional de *Syn*6803, permitindo a identificação de potenciais alvos para otimizar a produção de GB.

*This page was intentionally left blank*

# Abstract

Compatible solutes (CS) are low-weight molecular compounds that are accumulated intracellularly to maintain cell homeostasis in unfavourable growth conditions like high salinity or temperature. CS form a hydration shell around key macromolecules such as proteins, enabling stabilization and preventing denaturation and loss of function. These protective properties grant CS, such as glycine betaine (GB), high commercial value with applications in feed, cosmetic and pharmaceutical formulations. Currently, GB production methods are split into chemical synthesis, requiring hazardous chemicals and with a high carbon footprint, and extraction from sugar beets (*Beta vulgaris*) that is greener and safer but has poor yields. Thus, GB production requires efficiency improvement that is being addressed by engineering microorganisms to produce this solute. Cyanobacteria, as photoautotrophic microbes, are sought for this purpose given their fast growth rate and high productivity (compared with plants), and simple nutritional requirements in comparison with heterotrophic bacteria, using light, carbon dioxide, and water as sources of energy, carbon, and electrons respectively. Among cyanobacteria, a leading candidate is *Synechocystis* sp. PCC 6803 (*Syn*6803) due to its amenability to genetic engineering, available molecular tools and genome scale metabolic models to perform simulations. In previous work, a synthetic device for the production of GB (Ahbet) was introduced into this cyanobacterium, enabling the synthesis of this CS and improving growth under salinity conditions. These data also suggest that carbon is redirected from the glycogen sink to the production of CS and extracellular polymeric substances (EPS). In this project, the *Syn*6803 double mutant Δ*ggpS*Δ*slr0977*, deficient in the native CS glucosylglycerol and putatively impaired in EPS production, with and without the Ahbet device was characterized. The growth of these mutants was similar in absence or presence of 3% NaCl, suggesting that the introduction of the Ahbet device did not impose a metabolic burden. In addition, a decrease of glycogen and an increase of capsular polysaccharides (CPS) was observed under salinity conditions. The evaluation of CS production in the Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet mutants shows that the latter does not produce GB, likely due to mutations in the Ahbet device. Nevertheless, the results obtained so far revealed that the deletion of the *slr0977* gene does not result in the impairment of EPS production under the conditions tested, in contrast to previous work carried out in the absence of salt and under higher light intensity. Therefore, there is a need to understand the changes that occur in *Syn*6803 chassis used for GB production, by applying transcriptomic, proteomic, and metabolomic approaches. This knowledge will represent a step-forward for the rational engineering of *Syn*6803 by enabling the identification of potential targets to optimize production of GB.

*This page was intentionally left blank*

# Agradecimentos

Este trabalho representa o culminar de uma longa jornada de 5 anos. Ao terminar esta etapa, posso finalmente olhar para trás e agradecer a todos aqueles que contribuíram para que eu conseguisse chegar a este ponto.

Em primeiro lugar, devo a maior gratidão à minha orientadora, Catarina Pacheco, por me ter acompanhado com muita paciência ao longo de já dois anos. Tudo o que sei agora deve-se a todo o teu trabalho, e por isso ficarei eternamente agradecido. Quero também agradecer à Prof. Paula Tamagnini, por me ter acolhido no seu grupo "Bioengineering and Synthetic Microbiology" (BSM) e possibilitar este meu crescimento. Impossível seria esquecer todo o contributo da família BSM, com especial atenção aos meus colegas Manuel Correia, João Pissarra, João Marques, Daniel Santos, Malte Marquardt, Jorge Cardoso, e Diogo Rodrigues, pelo ambiente de camaradagem que potenciaram, e por assegurarem que quando tudo parecia estar perdido havia sempre uma luz ao fundo do túnel.

Gostaria também de agradecer a contribuição do Dr. Pedro Lamosa (CERMAX, ITQB, Universidade Nova de Lisboa), por possibilitar a quantificação dos solutos compatíveis que complementou este trabalho.

Não poderia deixar de agradecer à minha família, que sempre me apoiou durante todo este tempo, e sempre me encorajou a conseguir tudo o que eu desejava. Em especial, tenho que agradecer à minha mãe e à minha irmã, por serem o meu porto seguro e por partilharem tanto as minhas conquistas como as minhas frustrações.

Finalmente, preciso de agradecer a 10(!) pessoas que caminharam comigo lado a lado durante todo este tempo: Daniela Guedes, João Ribeiro, Rita Bessa, Joana Santos, Joana Sousa, Margarida Castro, Inês Baptista, Ana Sofia Teixeira, Maria Rocha, e Leonel Neves. Obrigado por todos os momentos que passámos juntos, por serem os primeiros tanto a festejar os sucessos como a oferecer ajuda quando eu precisava, e por estarem do meu lado tanto agora como para a frente.

This page was intentionally left blank

# List of contents

# List of figures

# List of tables

# Abbreviations

| | |
|---|---|
| **ADP** | Adenosine diphosphate |
| **ADP-Glu** | ADP-glucose |
| **BADH** | Betaine-aldehyde dehydrogenase |
| **BSA** | Bovine serum albumin |
| **CDS** | Coding sequence |
| **CHDH** | Choline dehydrogenase |
| **Chl *a*** | Chlorophyll *a* |
| **CHT** | Total carbohydrate content |
| **Cm** | Chloramphenicol |
| **CPS** | Capsular polysaccharides |
| **CRISPR** | Clustered regularly interspaced short palindromic repeats |
| **CRISPRi** | CRISPR interference |
| **CS** | Compatible solute |
| **DMT** | Dimethylglycine *N*-transferase |
| **DNA** | Deoxyribonucleic acid |
| **dNTP** | Deoxynucleotide triphosphate |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **EPS** | Extracellular polymeric substances |
| **Fru** | Fructose |
| **Fru-6-P** | Fructose-6-phosphate |
| **G-3-P** | Glycerol-3-phosphate |
| **GB** | Glycine betaine |
| **GG** | Glucosylglycerol |
| **GG-6-P** | Glucosylglycerol-6-phosphate |
| **GgpP** | Glucosylglycerol-phosphate phosphatase |
| **Glu** | Glucose |
| **GOE** | Great Oxidation Event |
| **GSMT** | Glycine sarcosine *N*-methyltransferase |
| **Gyr** | Billion years |
| **kDa** | Kilo Daltons |
| **Km** | Kanamycin |
| **log** | Natural logarithm |
| **metX** | *S*-adenosylmethionine synthase |
| **MW** | Molecular weight |
| **MWM** | Molecular weight marker |

| | |
|---|---|
| **NAD(P)** | Nicotinamide adenine dinucleotide (phosphate) |
| **NAD(P)H** | Reduced nicotinamide adenine dinucleotide (phosphate) |
| **NMR** | Nuclear magnetic resonance |
| **NOE** | Neoproterozoic Oxidation Event |
| **O/N** | Overnight |
| **OD** | Optical density |
| **PAL** | Present atmospheric level |
| **PCC** | Pasteur culture collection |
| **PCR** | Polymerase chain reaction |
| **PHB** | Polyhydroxybutyrate |
| **$P_i$** | Inorganic phosphorous |
| $p_{O_2}$ | Atmospheric partial pressure of $O_2$ |
| **RBS** | Ribosome binding site |
| **RIN** | RNA integrity number |
| **RNA** | Ribonucleic acid |
| **rpm** | Revolutions per minute |
| **RPS** | Released polysaccharides |
| **RT** | Room temperature |
| **SAH** | *S*-adenosylhomocysteine |
| **SAM** | *S*-adenosylmethionine |
| **Sps** | Sucrose-phosphate synthase |
| **Suc** | Sucrose |
| **Suc-6-P** | Sucrose-6-phosphate |
| **TAE** | Tris-acetate-EDTA |
| **TE** | Tris-EDTA |
| **UDP** | Uridine diphosphate |
| **UDP-Glu** | UDP-glucose |
| **vol/vol** | Volume/volume |
| **wt/vol** | Weight/volume |

*This page was intentionally left blank*

# Introduction

## 1.1 - Cyanobacteria: small organisms with great potential

Cyanobacteria, also known as blue-green algae, are prokaryotic organisms with an autotrophic metabolism and ability to perform oxygenic photosynthesis. Through this process, they convert solar energy, carbon dioxide, and water into complex organic carbon compounds and reducing power under the form of NAD(P)H, while releasing $O_2$ as a by-product (Garcia-Pichel et al., 2020; Hall and Rao, 1999). Ever since their genesis, 3.5 billion years (Gyr) ago, these organisms have been influencing and shaping the evolution of most other lifeforms on Earth (Knoll, 1980; Olson, 2006; Rosing and Frei, 2004). The prime example of the enormous impact they have had can be found in the Proterozoic era, a period ranging from 2.5 to 0.54 Gyr ago aptly named "the age of blue-green algae" (**Figure 1**). During the Great Oxidation Event (GOE, 2.32-2.5 Gyr ago), the increasing oxygen content in the atmosphere caused by the proliferation of cyanobacteria originated a mass extinction event with a drop in global biomass over 80% (Hodgskiss et al., 2019; Ostrander et al., 2019; Planavsky et al., 2014). Afterwards, in the Neoproterozoic Oxidation Event (NOE, 0.8-0.6 Gyr ago), the stabilization of these oxygen levels allowed the appearance of complex lifeforms such as animals, and the diversification of marine planktonic algae that helped develop the food webs necessary to support them (Brocks et al., 2017; Lyons et al., 2014; Sánchez-Baracaldo, 2015; Sánchez-Baracaldo & Cardona, 2020; Scott et al., 2008). In the present day, cyanobacteria remain in a very prestigious position in the tree of life, as the foundation of many of the planet's food chains, producers of oxygen, and agents of nitrogen fixation (Knoll, 2008).

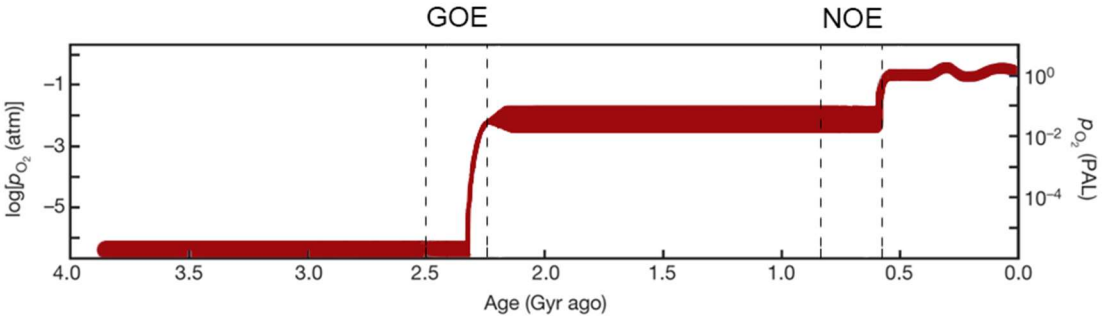



**Figure 1.** Evolution of Earth's atmospheric oxygen content through time. The major events of oxygen production, the Great Oxidation Event (GOE) and Neoproterozoic Oxidation Event (NOE), are highlighted (dashed lines). Right axis, atmospheric partial pressure of $O_2$ ($p_{O_2}$) relative to the present atmospheric level (PAL); left axis, log $p_{O_2}$; Gyr – billon years. Modified from Lyons et al. (2014).

Cyanobacteria are sought after for both research and industrial purposes. For the latter, heterotrophic microorganisms are mostly used, requiring the supply of carbon sources and thus limiting the sustainability of the production processes. In comparison, cyanobacteria require $CO_2$, water, and light to grow, as well as inorganic nutrients and a source of nitrogen in the case of the non-diazotrophic strains (Markou et al., 2014). These simple nutritional requirements make cyanobacteria much more cost-effective (Jodlbauer et al., 2021), especially as emergent technologies reduce the cost of artificial lighting, traditionally a major setback to process up-scaling (Chen et al., 2011; Johnson et al., 2018). Moreover, the faster growth rate and higher production ability of cyanobacteria over plants, the amenability to genetic engineering of several strains, and the possibility to use alternative water sources such as seawater for cultivation, combine to make them extremely appealing chassis for added value compound production (Machado et al., 2012; Nozzi et al., 2013; Rodionova et al., 2017). The products synthesized by cyanobacteria are very diverse, ranging from bioplastics to compounds with health benefits that display antiviral, antibacterial, antifungal, and even anticancer activity (Abed et al., 2009). Among all the organisms that encompass the extremely diverse group that is cyanobacteria, *Synechocystis* sp. PCC 6803 (hereafter *Syn*6803) has gained relevance as a cell-factory, due to the early sequencing and annotation of its genome (Kaneko et al., 1996) that generated a vast amount of data, along with the development of genome-scale metabolic models (Fu, 2009; Montagud et al., 2011). Customized synthetic biology tools have also been developed for this organism, such as synthetic promoters, ribosome binding sites (RBS), and other regulatory elements like riboswitches, expanding the available portfolio and insulating the device from the native regulatory network (Ferreira et al., 2018; Huang and Lindblad, 2013; Nagy et al., 2021; Ohbayashi et al., 2016; for a review see Pacheco et al., 2021). All together, these advancements allow a more rational approach to increase the production of desirable compounds using *Syn*6803 as chassis.

## 1.2 - Compatible solutes: properties, structure, and synthesis

To be able to grow under salinity conditions, cyanobacteria require specific adaptation mechanisms. Exposure to salt generates an osmotic pressure that forces water out of the cell and causes a high influx of ions, which leads to a decrease in turgor pressure and cell volume as well as to the denaturation of macromolecules and complex structures (such as proteins and membranes), and inevitably causing cell death (Ladas and Papageorgiou, 2000) (**Figure 2A**). The "salt-out" strategy (**Figure 2B**) is one of the mechanisms that enables cells to maintain the osmotic balance by actively pumping $Na^+$ and $Cl^-$ out of the cell and by, simultaneously, synthesizing a particular set of substances known as compatible solutes (CS). CS are characterized by very high solubility, absence of net charge, and inability to directly interact with macromolecules (Galinski, 1995; Klähn and Hagemann, 2011). Instead, these solutes interact with water generating a hydration shell around the macromolecules, which has a thermodynamic stabilizing effect that counteracts denaturation (Galinski, 1995; Jadhav et al., 2018). Moreover, these solutes can be accumulated intracellularly at molar concentrations that balance osmotic pressure and prevent the loss of turgor (Galinski, 1995; Klähn and Hagemann, 2011). By protecting key players in cell metabolism, CS are not only associated with resistance to salt stress, but also to extreme temperature shifts, desiccation, oxidative and chaotropic stress (Cray et al., 2015; de Lima Alves et al., 2015; Galinski, 1995). The breadth of their functions is rivalled only by their applications in various fields: in molecular biology, protecting

enzymes and DNA; applied in cosmetics as sunscreen and antioxidants; in pharmaceuticals as excipients and active compounds alike; and also used in food/feed industries as stabilizers (Shivanand and Mugeraya, 2011).

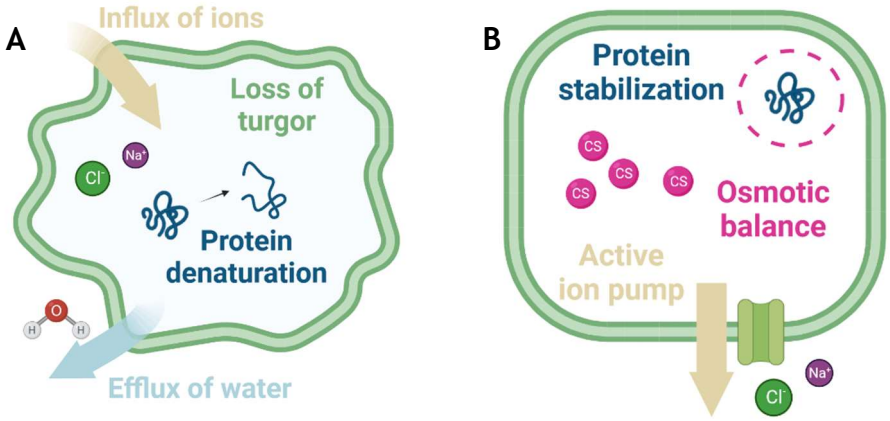



**Figure 2.** Schematic representation of the effects of salt stress in cyanobacteria **(A)**, and of the "salt-out" response as a mechanism to cope with salinity **(B)**. CS – compatible solute. Figures from Klähn and Hagemann (2011) modified using biorender.com.

Only a few molecules have the properties necessary to function as CS, belonging to specific groups such as sugars, polyols, heterosides, amino acids and their derivatives. In addition, not all CS provide the same level of protection towards osmotic stress. The sugars, sucrose and trehalose, are associated with freshwater strains that have a low salt tolerance (NaCl concentration up to 3% or 0.51 M, also considered seawater-like conditions). Glucosylglycerol and glucosylglycerate, both heterosides, are characteristic of marine strains found in environments with up to 10.5% NaCl (1.8 M). The most effective CS are the amino acid derivatives, glycine betaine and glutamate betaine, mostly present in halophile strains that grow in hypersaline environments (with salt saturated concentrations ~17.5%/3 M NaCl) (Hagemann, 2011, 2013; Kirsch et al., 2019; Klähn and Hagemann, 2011) (**Figure 3**).

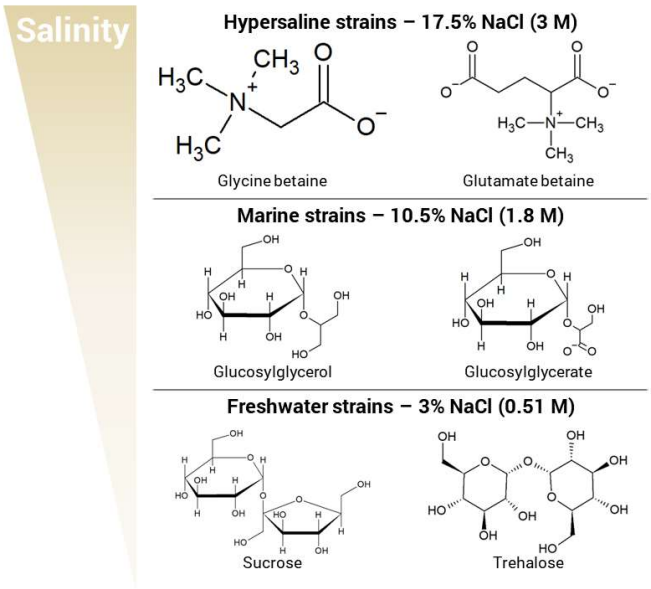



**Figure 3.** Structure of different compatible solutes commonly found in cyanobacteria and correlation with their habitat. Figures from Hagemann (2011) and Klähn and Hagemann (2011) modified using biorender.com.

Although *Syn*6803 was isolated from freshwater, it is a moderately halotolerant organism that produces sucrose (Suc) and glucosylglycerol (GG) as its main CS (Reed and Stewart, 1985). The syntheses of these solutes require a two-step reaction catalysed by two enzymes. For Suc, the enzyme sucrose-phosphate synthase (Sps) acts on UDP-glucose and fructose-6-phosphate to synthesize sucrose-6-phosphate; followed by sucrose-phosphate phosphatase (Spp) which converts it into the final compound (**Figure 4A**). Similarly, the synthesis of glucosylglycerol begins with the conversion of ADP-glucose and glycerol-3-phosphate into glucosylglycerol-6-phosphate by glucosylglycerol-phosphate synthase (GgpS); and culminates in the dephosphorylation of this product by glucosylglycerol-phosphate phosphatase (GgpP) to generate GG (Klähn and Hagemann, 2011) (**Figure 4B**). In both cases, the first enzymatic step is regulated and is responsible for the production of the CS in a salt-dependent manner. As the intracellular ion concentration increases, *sps* and *ggpS* are derepressed and, subsequently, the translated protein is activated (Kirsch et al., 2019). While the mechanism behind this activation is still unclear, it is likely based on the salt-dependent phosphorylation of Sps and GgpS (Hagemann and Erdmann, 1994; Hagemann et al., 1993).

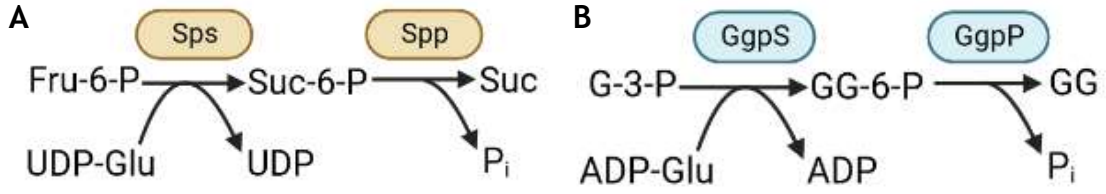



**Figure 4.** Biosynthetic pathways for the production of sucrose (Suc) **(A)** and glucosylglycerol (GG) **(B)**. UDP – uridine diphosphate; UDP-Glu – UDP-glucose; Fru-6-P – fructose-6-phosphate; Suc-6-P – sucrose—6-phosphate; $P_i$ – inorganic phosphorous; Sps - sucrose-phosphate synthase; Spp - sucrose-phosphate phosphatase; ADP – adenosine diphosphate; ADP-Glu – ADP-glucose; G-3-P – glycerol-3-phosphate; GG-6-P – glucosylglycerol-6-phosphate; GgpS - glucosylglycerol-phosphate synthase; GgpP - glucosylglycerol-phosphate phosphatase.

## 1.3 - Glycine betaine as a compatible solute of industrial interest

Among the CS mentioned above, glycine betaine (GB) is particularly attractive for commercial purposes given its applications in: feed supplementation (Aquilina et al., 2013); cryoprotection (Cleland et al., 2004); the optimization of PCRs (Weissensteiner and Lanchbury, 1996); cosmetics (Estrin et al., 1997); and in pharmaceutical formulations. Currently, GB is used to treat homocystinuria (European Medicines Agency, 2016) with additional potential uses in liver disease (Day and Kempson, 2016) as well as clotting and haemorrhagic conditions (Messadek, 2002). Traditionally, GB has been obtained by extraction from sugar beets (*Beta vulgaris*), but this process results in low yields, becoming therefore expensive (Heikkila et al., 1981). Alternatively, GB can be produced by chemical synthesis in a cheaper fashion but it requires the use of strong acids and bases, and consumes a lot of energy for heating and cooling (Kegang, 1997), thus generating a large carbon footprint (DuPont, 2015). As such, greener and more effective processes for GB synthesis are required, with its biological production by microorganisms coming to the forefront as a viable option given its sustainability (Sheldon and Woodley, 2018). Several microbes have been identified as producers of GB, including phototrophic and methanogenic bacteria (Galinski, 1995), but also chemoheterotrophs. Among these producers, both halophiles and non-halophiles can be found (Imhoff and Rodriguez-Valera, 1984; Ollivier et al., 1994; Oren, 1990), including *Escherichia coli*, an ubiquitously used organism for industrial synthesis processes (Landfald and Strøm, 1986).

The biosynthesis of GB can occur via two distinct pathways. The most commonly found metabolic route is the one shared by the chemoheterotrophs, involving a two-step oxidation of choline by choline dehydrogenase (CHDH) and betaine-aldehyde dehydrogenase (BADH), with glycine betaine-aldehyde as the intermediate (Landfald and Strøm, 1986) (**Figure 5A**). However, this requires the presence of choline in the growth medium, as these organisms rely on its uptake for the production of GB (Hagemann, 2011). The other pathway is restricted to phototrophic organisms such as plants and cyanobacteria (Hagemann, 2011), and allows the *de novo* synthesis of GB through three successive methylation steps of glycine using *S*-adenosylmethionine (SAM) as a donor of methyl groups. The methylations are catalysed by two distinct enzymes: glycine sarcosine *N*-methyltransferase (GSMT), which is responsible for the first two methylations; and dimethylglycine *N*-methyltransferase (DMT), which converts dimethylglycine into GB (Nyyssölä et al., 2000; Waditee et al., 2003) (**Figure 5B**). This tri-methylation pathway would be preferred in an industrial setting, as it does not require the supply of choline, lowering costs and improving economic feasibility.

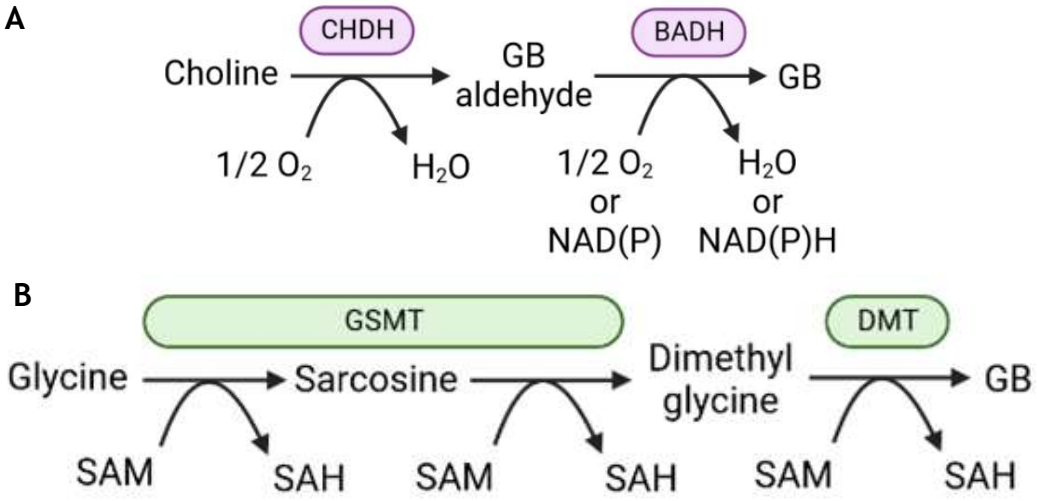



**Figure 5.** Biosynthetic pathways of glycine betaine (GB), via the oxidation of choline **(A)** or the tri-methylation of glycine **(B)**. CHDH - choline dehydrogenase; BADH - betaine-aldehyde dehydrogenase; NAD(P) – nicotinamide adenine dinucleotide (phosphate); NAD(P)H – reduced nicotinamide adenine dinucleotide (phosphate); GSMT - glycine sarcosine *N*-methyltransferase; DMT - dimethylglycine *N*-methyltransferase; SAM - *S*-adenosylmethionine; SAH – *S*-adenosylhomocysteine.

## 1.4 - Optimization of *Syn*6803 as a chassis for glycine betaine production

The application of *Syn*6803 as an industrial-scale producer of GB requires the implementation of the biosynthetic pathway, as it is not one of this cyanobacterium's native CS. For the proof-of-concept, Ferreira et al. (2022) selected the tri-methylation pathway described for *Aphanotece halophytica* and, based on the available information, designed a synthetic device envisaging GB production in *Syn*6803-based chassis, designated Ahbet. This device contained not only the coding sequences (CDS) - *gsmt* and *dmt* - for the methyltransferases involved in the synthesis of GB, but also included the CDS for the *S*-adenosylmethionine synthase - *metX*. The latter was included to prevent the shortage of the SAM cofactor, which could potentially hinder GB production. The Ahbet device was introduced into CS-deficient strains of *Syn*6803, generated by single or double deletion of the *ggpS* and *sps* genes (encoding the enzymes responsible for the first step of CS synthesis). Subsequently, CS quantification showed that the implementation of the Ahbet device enables *Syn*6803 to produce GB, achieving a maximum production of 1.89 µmol GB/mg protein in the glucosylglycerol deficient chassis (Δ*ggpS* Ahbet). Furthermore, it also allowed the survival of this strain under 5% NaCl, an otherwise lethal salt concentration for the Δ*ggpS* background. In addition, carbohydrate quantification demonstrated that, in salt conditions, the pool of glycogen decreases while the amount of CS and extracellular polymeric substances (EPS) increase, namely, capsular polysaccharides (CPS). In cyanobacteria such as *Syn*6803, EPS are complex polysaccharides with several roles, including preventing dehydration and making them key players in the salt stress response (Jittawuttipoka et al., 2013; Pereira et al., 2009). Put together, these results allowed to devise a strategy for optimizing the production of GB by removing competing carbon sinks (such as glycogen and EPS), and diverting resources to the synthesis of the heterologous CS (Ferreira et al., 2022). As mentioned above, one of the competing carbon sinks are EPS. On other bacteria, the synthesis of EPS relies on three

distinct pathways: the Wzy-, the ABC transporter-, or the Synthase-dependent (Whitfield et al., 2020). In *Syn*6803, genes related to all three are present, but not the complete set that defines one pathway (Pereira et al., 2015). Studies focusing on key proteins related to these pathways have shown that a deficiency in the KpsM (Slr0977) protein, a transmembrane domain of the ABC transporter involved in EPS export, results in a significant decrease in EPS production, with a 20% and 50% drop on CPS and released polysaccharides (RPS), respectively (Santos et al., 2021a). In *Syn*6803, the *kpsM* gene has three putative homologues: *slr0977*, *slr2107*, and *sll0574*. The simultaneous inhibition of the transcription of these homologues by CRISPRi suggested that, in *Syn*6803, Slr0977 is the main KpsM homologue involved in RPS export (in standard laboratory conditions) (Santos et al., 2021a; Santos et al., 2021b). Considering that the Δ*slr0977* mutant has impaired EPS production (Santos et al., 2021a), previously suggested to be a competing pathway for the production of CS in *Syn*6803 under salinity conditions (Ferreira et al., 2022); the use of this mutant could result in the optimization of GB production in this cyanobacterium. Furthermore, the combination of the deletion of the *slr0977* and *ggpS* genes is also envisaged as a chassis that could further increase the production of GB.

## 1.5 - Aims

This project aimed at the optimization of the production of the heterologous compatible solute (CS) glycine-betaine (GB) using a cyanobacterial (*Syn*6803)-based chassis. Building upon previous work, the synthesis of the extracellular polymeric substances (EPS) - previously identified as a potential competing pathway (Ferreira et al., 2022) - was targeted to redirect metabolic fluxes towards the production of GB. For this purpose, *Syn*6803 mutants expected to have impaired EPS production were investigated as chassis for GB production, namely Δ*slr0977* and the double mutant Δ*ggpS*Δ*slr0977*. The characterization of these mutants harbouring the device for GB production included:

i) assessment of growth fitness under salinity conditions,

ii) quantification of carbohydrates (capsular and released polysaccharides, glycogen and polyhydroxybutyrate);

iii) evaluation of CS production (glycine betaine and the native CS, glucosylglycerol and sucrose);

iv) the initial work for a transcriptome analysis that will allow to gain insights on the key players in *Syn*6803 GB-production and unveiling targets for future optimizations.

# Materials and methods

## 2.1 - Reagents and enzymes

The media components and other reagents were obtained from Fisher Scientific (USA), Merck (Germany), or Sigma Aldrich (USA). All DNA-modifying enzymes and polymerases were purchased from ThermoFisher Scientific (USA) and Promega (USA).

## 2.2 - Organisms and culture conditions

The *Syn*6803 wild-type and mutants used in this work were maintained in Erlenmeyer flasks batch cultures with BG11 medium (Stanier et al., 1971), supplemented with antibiotics when appropriate (**Table 1**). The cultures were incubated at 30 °C with orbital shaking (150 rpm) under a 12 h light/12 h dark regimen. In all experiments, the light intensity was 25 $\mu E/m^2/s$, and the Cosine-corrected irradiance was measured using a Dual Solar/Electric Quantum Meter (Spectrum Technologies, Inc., United States).

**Table 1.** List of strains used in this work

| Strain | Description | Antibiotic | Source/Reference |
|---|---|---|---|
| *Synechocystis* sp. PCC 6803 (*Syn*6803) | Unicellular, non-motile *Synechocystis* sp. PCC 6803 substrain GT-Kazusa | - | Pasteur Culture Collection, France (Kanesaki et al., 2012; Trautmann et al., 2012) |
| *Syn*6803 Δ*slr0977* | *Syn*6803 mutant with a kanamycin resistance cassette replacing the *slr0977* locus | Km 400 µg/mL | Santos et al. (2021a) |
| *Syn*6803 Δ*slr0977* Ahbet | *Syn*6803 Δ*slr0977* mutant harbouring the Ahbet device | Km 400 µg/mL<br>Cm 10 µg/mL | Host group unpublished work |
| *Syn*6803 Δ*ggpS* Ahbet | *Syn*6803 mutant with the *ggpS* locus deleted, harbouring the Ahbet device | Cm 10 µg/mL | Ferreira et al. (2022) |

| | | | |
|---|---|---|---|
| *Syn*6803 Δ*ggpS*Δ*slr0977* | *Syn*6803 Δ*ggpS* mutant with a kanamycin resistance cassette replacing the *slr0977* locus | Km 400 µg/mL | Host group unpublished work |
| *Syn*6803 Δ*ggpS*Δ*slr0977* Ahbet | *Syn*6803 Δ*ggpS*Δ*slr0977* mutant harbouring the Ahbet device | Km 400 µg/mL Cm 10 µg/mL | Host group unpublished work |

## 2.3 - Growth experiments

Pre-cultures of *Syn*6803 mutants were inoculated in BG11 medium (supplemented with 200-250 µg/mL Km and/or 5 µg/mL Cm, when appropriate), and maintained in the conditions described above. The cultures were grown to an $OD_{730}$ ≈ 2 and, subsequently diluted in BG11 medium (without antibiotic) to a final $OD_{730}$ ≈ 0.5. This culture was distributed into 100 mL Erlenmeyer flasks (50 mL per flask) without NaCl or containing 3, 5, or 7% (wt/vol) NaCl (0.51, 0.86, and 1.2 M NaCl, respectively), previously sterilized by autoclaving. These cultures were maintained in the same conditions as the pre-cultures and growth was monitored for 16 days by measuring $OD_{730}$. All experiments included four biological replicates with technical duplicates.

## 2.4 - Chlorophyll *a* measurement

The chlorophyll *a* (chl *a*) content was determined as described by Meeks and Castenholz (1971). Briefly, 1 mL of culture was centrifuged for 2 min at 14,000 *xg*, and the pellet resuspended in an equal volume of 90% (vol/vol) methanol. Samples were left overnight (O/N) in the dark at 4 ºC, after which, were vortexed and centrifuged for 2 min at 14,000 *xg*. The supernatants' OD at 663 nm was then measured and chl *a* concentration calculated using a molar extinction coefficient of 12.7 mL/µg/cm. Measurements were performed for four biological replicates with technical duplicates.

## 2.5 - Total carbohydrate content, released and capsular polysaccharides measurements

Total carbohydrate content (CHT), released polysaccharides (RPS) and capsular polysaccharides (CPS) were determined as previously described (Mota et al., 2013). Ten millilitres of culture samples were dialyzed (12–14 kDa molecular weight cut-off; CelluSepT4, Orange Scientific, Belgium) against, at least, 10 volumes of distilled water or 3% (wt/vol) NaCl solution (identical to the growth medium) for a minimum period of 24 h. One millilitre of the collected sample was used to spectrophotometrically quantify CHT by the phenol-sulphuric acid method (Dubois et al., 1956). Briefly, 1 mL of 5% (vol/vol) phenol and 5 mL of concentrated sulphuric acid were added to the sample, mixed by inverting, and allowed to

react for 10 min before being incubated in a water bath at RT for 15 min. The solution's OD at 488 nm was measured and carbohydrate concentration determined from a glucose standard curve equation (**Supplementary Figure 1)**. To determine the RPS, 5 mL of the dialyzed sample were centrifuged at 3,220 *xg* for 10 min at RT, and 1 mL of the cell-free supernatant was used for quantification using the same method described for the CHT. For CPS quantification, five millilitres of dialyzed cultures were centrifuged at 3,220 *xg* for 10 min at RT, after which the cell pellet was resuspended in distilled water and boiled in a mineral oil bath at 96 ºC for 15 min. The samples were centrifuged as described previously, and 1 mL of the cell-free supernatant was used for CPS quantification by the phenol-sulphuric acid method. CHT, RPS, and CPS were normalized by chl *a* content. All experiments included four biological replicates with technical triplicates.

## 2.6 - Glycogen extraction and quantification

Glycogen extraction was performed as described in Ernst et al. (1984). The cell samples were collected 1 h after the transition between the dark and the light phase, 10 mL of the cultures were centrifuged for 10 min at 3,220 *xg*, and the cell pellets stored at -80 ºC until further use. The cells were thawed, resuspended in 100 µL of distilled water and transferred to a 2 mL screw-cap tube. Four hundred microlitres of 30% (wt/vol) KOH were added and the mixture was incubated at 100 °C for 90 min, then quickly cooled on ice. Six hundred microlitres of ice-cold absolute ethanol were added and the mixture was incubated on ice for 2 h, and subsequently centrifuged for 5 min at 16,000 *xg*, 4 °C. The supernatant was discarded, and the isolated glycogen was washed 6 times with 500 µL of ice-cold absolute ethanol, separating the pellet by centrifugation as previously described. After the washing steps, the pellet was dried at 60 °C for 5-10 mins, then resuspended in 1 mL of distilled water and transferred to a 15 mL Falcon. The sample was diluted by adding 5 mL of distilled water, and used for glycogen quantification by the phenol-sulphuric acid method (Dubois et al., 1956), as described above. The results from the quantification were normalized by chl *a* content. Experiments included four biological replicates with technical triplicates.

## 2.7 - PHB extraction and quantification

PHB extraction was performed as described previously (Thiel et al., 2017) with adaptations. Ten millilitres of cell culture were centrifuged at 3,220 *xg* for 10 min, with the cell pellets being stored at -80 °C until further use. For PHB quantification, the alkaline lysis of the polymer to D-3-hydroxybutyric acid was performed by resuspending the pellets in 400 µL of 0.5 M NaOH. The suspension was then heated at 85 °C with shaking at 300 rpm for 1 h, after which the pH was adjusted to 8.0 by adding ~100 µL of 1 M HCl. Then, 150 µL of lysate were used for PHB quantification performed with the D-3-Hydroxybutyric Acid Assay Kit (Megazyme, USA). The assay was set up in 96-well plates including technical triplicates for each sample, and the absorbance at 492 nm was measured using a microplate reader (CLARIOstar®, Germany). The PHB concentration was determined using a D-3-hydroxybutyric acid standard curve equation (**Supplementary Figure 2**), and normalized by chl *a* content. Experiments included at least two biological replicates with technical triplicates.

## 2.8 - Compatible solutes quantification

Cultures of *Syn*6803 mutants Δ*slr0977* and Δ*ggpS*Δ*slr0977* with and without the Ahbet device were grown in BG11 or BG11 supplemented with 3% (wt/vol) NaCl, at an initial $OD_{730}$ ≈ 0.5, as described in section 2.3 - . The quantification of the native and heterologous CS (sucrose, glutamate, glucosylglycerol, and glycine betaine), was performed using 500 mL of culture (grown in ten 100 mL Erlenmeyer flasks,50 mL of culture per flask. Three or four days after inoculation ($OD_{730}$ ~1.0), the cells were harvested by centrifugation at 3,220 *xg* for 10 min at RT. Cells were washed using 100 mL of ice-cold distilled water or 3% (wt/vol) NaCl solution (identical to the growth medium). Centrifugation was repeated and the cell pellets were resuspended in 50 mL of the respective washing solutions. From this suspension, a 0.5 mL aliquot was collected and centrifuged for 2 min at 14,000 *xg*; later on, the cell pellet was used for protein quantification. The remaining cell suspension was centrifuged at 3,220 *xg*, 4 °C, and the pellet stored at −20 °C. Ethanol-chloroform extraction of the CS was performed as described in Ferreira et al. (2022). Briefly, cell pellets were suspended in 25 mL of 80% (vol/vol) ethanol and subsequently transferred to a 100 mL round flask containing a magnetic stir bar. The flask was connected to a coil condenser (circulating cold water), and heated at 100 °C with stirring for 10 min. The suspension was transferred to a 50 mL tube and centrifuged at 3,220 *xg* for 10 min, RT. The supernatant was stored and the pellet resuspended in 20 mL of 80% (vol/vol) ethanol for a new extraction process, and the supernatants from both extraction processes were pooled. The remainder protocol was performed as described in Santos et al. (2006). The ethanol from the supernatants was removed by rotary evaporation using a Buchi R300 rotary evaporator (Buchi, Switzerland) using the following settings: water bath at 40 °C, condenser at 4 °C and 120 mbar of pressure. The remaining mixture was transferred to a glass centrifuge tube and half a volume of chloroform was added and mixed by vortexing. The emulsion was then centrifuged at 3,220 *xg* for 15 min at RT, and the upper, aqueous phase was transferred to a 50 mL Falcon. To the remaining chloroform phase, half a volume of $dH_2O$ was added, and the extraction process repeated. The resulting supernatants were then pooled and frozen at -80 °C, and subsequently lyophilized. Detection, identification, and quantification of CS was performed by proton NMR (this step was performed by Dr. Pedro Lamosa at CERMAX, ITQB, Universidade Nova de Lisboa). For this effect, freeze-dried extracts were dissolved in 1 mL of $D_2O$ and a known amount of sodium formate was added to serve as an internal concentration standard. Spectra were acquired at 25 °C on a Bruker Avance III 800 spectrometer (Bruker, Rheinstetten, Germany) working at a proton operating frequency of 800.33 MHz, equipped with a 5 mm, three channel, inverse detection cryoprobe TCI-z H&F/C/N with pulse-field gradients. A 3 s soft pulse was applied before the excitation pulse, to pre-saturate the water signal. Spectra were acquired under fully relaxed conditions (flip angle 60°; repetition delay of 60 s) so that the area of the NMR signals was proportional to the amount of the different protons in the sample. Integration of the signals was performed using the tools available in the TopSpin software (Bruker, Rheinstetten, Germany) version 3.6.2.

The concentration of CS was expressed as µmol per mg of protein. For protein quantification, the cells were resuspended in TE buffer (10 mM Tris, 1mM EDTA pH 8.0), and disrupted by vortexing for 2 min. The suspension was centrifuged at 12,000 *xg* for 10 min, and protein quantification was performed using the Bio-Rad Protein Assay (Bio-Rad, USA) following the manufacturer's instructions. Protein concentration was calculated from a bovine serum

albumin (BSA) standard curve equation (**Supplementary Figure 3).** Experiments included two biological replicates.

## 2.9 - PCR

The presence or absence of the Ahbet device in the Δ*slr0977* and Δ*slr0977* Ahbet strains was assessed by PCR using primers targeting the *gsmt* ORF (**Table 2**) of the Ahbet device. Each reaction mixture contained: 0.5 U of GoTaq® G2 Flexi DNA Polymerase (Promega, USA), 1x Green GoTaq Flexi buffer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.25 µM of each *gsmt* primer (**Table 2**), and 4 µL of cell extract (used for protein quantification, see previous section). The PCR profile was: initial denaturation of 5 min at 95 °C; followed by 25 cycles of 20 s at 95 °C (denaturation), 20 s at 56 °C (annealing), and 20 s at 72 °C (extension); and a final extension at 72 °C for 5 min. For the amplification of the Ahbet device sequence, the reaction mixtures contained: 0.4 U of Phusion™ High-Fidelity DNA Polymerase (ThermoFisher Scientific ,USA), 1x Phusion™ HF buffer, 200 µM of each dNTP, 0.25 µM of each PS primer (**Table 2**), and 8 µL of cell extract (used for protein quantification, see previous section). The PCR profile was: initial denaturation of 30 s at 98 °C; followed by 30 cycles of 10 s at 98 °C (denaturation), 20 s at 67 °C (annealing), and 1 min 45 s at 72 °C (extension); and a final extension at 72 °C for 10 min. The PCR reactions were run on 1% (wt/vol) agarose gel electrophoresis performed by standard protocols using TAE buffer(Sambrook and Russell, 2006).

## 2.10 - RNA extraction

Cell sampling was performed by centrifuging 100 mL of *Syn*6803 culture at $OD_{730}$ ≈ 1 (or equivalent volume), for 10 min at 3,220 x*g*. Cell pellets were treated with RNAprotect Bacteria Reagent (Qiagen, Germany) according to instructions and stored at −80 °C. Total RNA was extracted using the TRIzol® Reagent (Ambion, USA) in combination with the PureLink™ RNA Mini Kit (Invitrogen, USA), including the On-column PureLink® DNase Treatment. Briefly, the cells were disrupted in 1 mL TRIzol® Reagent containing 0.2 g of 0.2 mm-diameter glass beads (acid washed, Sigma, USA) using FastPrep®-24 (MP Biomedicals, USA), with 2 cycles of 1 min at 4.0 m/s and 1 min incubation on ice in between. For the remaining extraction protocol, all steps were performed according to the PureLink™ RNA Mini Kit instructions except DNase treatment, in which a 90 min incubation period at 25 °C was used. After elution, the RNA samples (30 µL) were stored at -80 °C until further use. RNA concentration and purity (the ratios $A_{260}/A_{280}$ and $A_{260}/A_{230}$) were measured using NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA). The RNA quality and integrity was checked using the RNA 6000 Pico Kit (Agilent technologies, USA). The absence of genomic DNA contamination was checked by PCR, in reaction mixtures containing: 0.5 U of GoTaq® G2 Flexi DNA Polymerase (Promega, USA), 1x Green GoTaq Flexi buffer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.25 µM of each *rnpB* primer (**Table 2**), and 200 ng of total RNA. The PCR profile was: initial denaturation of 5 min at 95 °C; followed by 25 cycles of 20 s at 95 °C (denaturation), 20 s at 56 °C (annealing), and 20 s at 72 °C (extension); and a final extension at 72 °C for 5 min. The PCR reactions were run on 1% (wt/vol) agarose gel electrophoresis performed by standard protocols using TAE buffer(Sambrook and Russell, 2006).

**Table 2.** List of oligonucleotides used in this work

| Primer name | Primer sequence (5'→3') | Tm (°C) | Amplicon size (bp) | Purpose | Reference |
|---|---|---|---|---|---|
| rnpBF1 | CGTTAGGAT AGTGCCACAG | 56 | 136 | Assessment of genomic DNA contamination | Pinto et al. (2012) |
| rnpBR1 | CGCTCTTAC CGCACCTTTG | | | | |
| SD_GSMT_F | TGCTAAGCGG GTACTAGATGC | 56 | 350 | Confirmation of presence of the Ahbet device | Ferreira et al. (2022) |
| SD_GSMT_R | CCTTCGTCCA AGATCAAATCG | | | | |
| PS1 | AGGGCGGC GGATTTGTCC | 60 | 3500 | Amplification of the Ahbet device sequence | Silva-Rocha et al. (2013) |
| PS2 | GCGGCAACC GAGCGTTC | | | | |

## 2.11 - Statistical analysis

The statistical analysis was performed by means of one-way analyses of variance (ANOVAs) followed by Tukey's multiple comparisons test, or by the *t*-test, using GraphPad Prism v9.5.0 (GraphPad Software Inc., USA).

# Results

In previous work done by the host group, the heterologous production of glycine betaine (GB) was achieved in several *Syn*6803-based chassis deficient in the production of native compatible solutes (CS), with a Δ*ggpS* Ahbet mutant displaying the highest GB production and improved growth under 3% (wt/vol) NaCl (Ferreira et al., 2022). Alongside this, Ferreira et al. (2022) identified two key carbon sinks that compete for resources with the synthesis of CS: glycogen, a carbon storage polymer; and the extracellular polymeric substances (EPS), which have several protective roles including resistance to salt stress. Therefore, targeting these pathways could redirect metabolic fluxes toward the production of GB. Considering this, a *Syn*6803 mutant reported to have impaired EPS production - Δ*slr0977* – (Santos et al., 2021a), was used as chassis. The introduction of the Ahbet device into Δ*slr0977* did not lead to an improvement of growth in the absence or presence of salt. The analysis of the carbohydrate profiles of Δ*slr0977* and Δ*slr0977* Ahbet also revealed that, in contrast to what was expected, the impairment in EPS production was abolished under the conditions tested. Nevertheless, the production of GB was 3 µmol/mg protein (unpublished work by the host group), that represents a 59% increase compared with the obtained using the Δ*ggpS* Ahbet. Here, in an attempt to further improve the production of GB, the use of a *Syn*6803 double mutant that combines the gene deletions that led to highest GB production -Δ*ggpS*Δ*slr0977* - as chassis for GB production was investigated.

## 3.1 - Effect of NaCl on the growth of Δ*ggpS*Δ*slr0977* strains

The Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet strains were cultivated in standard BG11 medium or in BG11 supplemented with 3, 5, and 7% (wt/vol) NaCl (corresponding to 0.51, 0.86, and 1.2 M, respectively), and the growth was monitored over a period of 16 days by measuring the $OD_{730}$ and chl *a* content (**Figure 6**). In the absence of salt, both Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet showed growth similar to the wild-type (**Figure 6A**). Challenging cells with 3% NaCl did not have a significant impact in the behaviour of both strains (**Figure 6B & 6A**), with the chl *a* content showing an identical trend to the $OD_{730}$. Further increasing the NaCl concentration to 5% caused a growth arrest that was accompanied by a severe decline in chl *a* content, as by day 16 this pigment showed a decrease of ~44% for the Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet mutants (**Figure 6C**). At 7% NaCl, the higher salinity tested, none of the strains were able to grow, leading to the total loss of chl *a* at day 14 (**Figure 6D**).

In summary, the implementation of the Ahbet device into Δ*ggpS*Δ*slr0977* did not have any discernible effect on growth compared with the Δ*ggpS*Δ*slr0977* background, under the conditions tested.

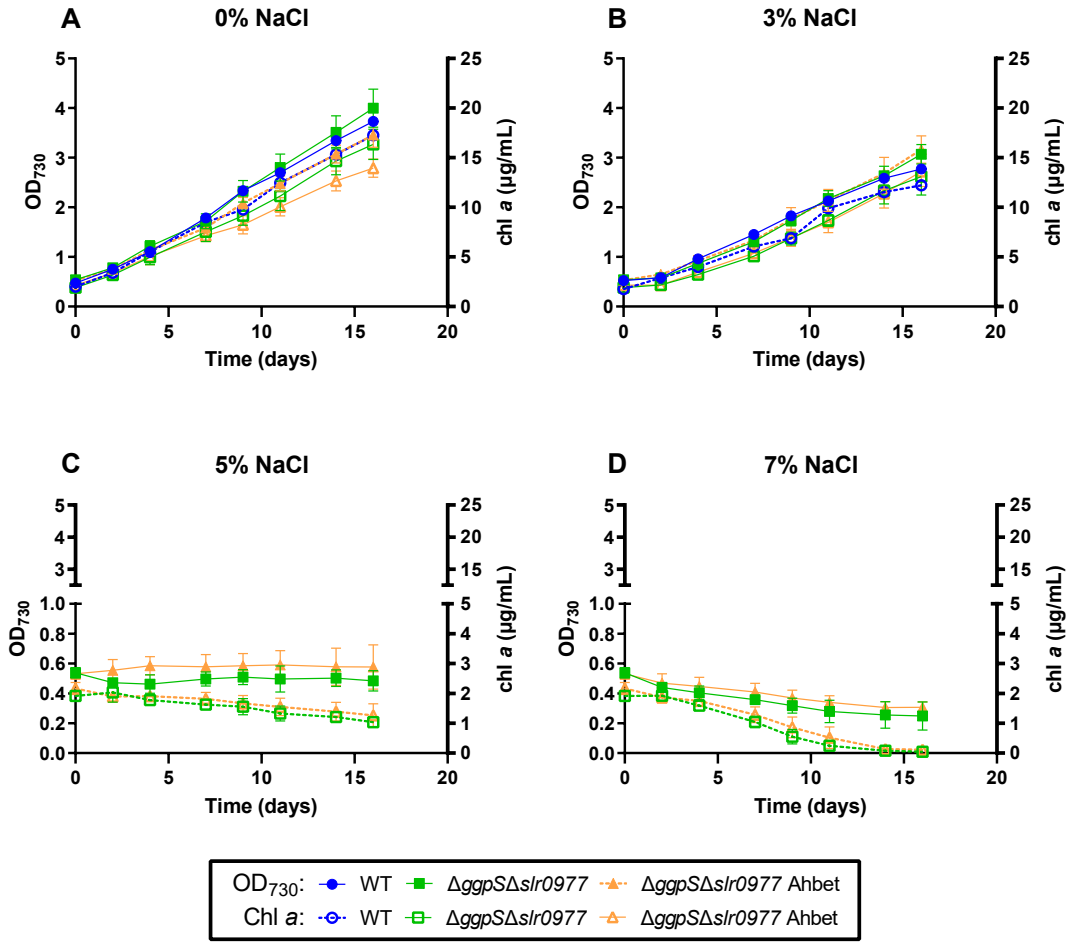



**Figure 6.** Growth curves of *Syn*6803 wild-type (WT) and Δ*ggpS*Δ*slr0977* mutant with and without the Ahbet device. Cultures were grown in BG11 **(A)** or BG11 supplemented with 3% **(B)**, 5% **(C)**, or 7% NaCl **(D)** at 30ºC with shaking at 150 rpm under a 12 h (25 µE/m$^2$/s) light/12 h dark cycle. Growth was monitored by measuring optical density at 730 nm ($OD_{730}$), and chlorophyll *a* (chl *a*) (full and dotted lines, respectively). Error bars correspond to standard deviations from four biological replicates with technical duplicates. Statistical analysis was performed using the *t*-test. Data for the wild-type (WT) correspond to one biological replicate that was included merely for comparison purposes.

## 3.2 - Effect of NaCl on the carbohydrate content of Δ*ggpS*Δ*slr0977* strains

The total carbohydrate (CHT) content of Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet strains grown in standard BG11 medium or in BG11 supplemented with 3% (wt/vol) NaCl was measured (**Figure 7A**). The CHT showed no variation with salinity in the Δ*ggpS*Δ*slr0977* mutant, whereas for the Δ*ggpS*Δ*slr0977* Ahbet an 8% increase ($p \leq 0.0373$) was observed when cells were grown under 3% NaCl (**Figure 7A**). To clarify the carbon distribution under salinity conditions, the pools of extracellular polymeric substances, CPS (capsular polymeric substances) and RPS (released polymeric substances), as well as the reserve polymers

glycogen and polyhydroxybutyrate (PHB) were analysed. The CPS content of both strains increased by two-fold under 3% NaCl (**Figure 7B**), while no change was observed in RPS (**Figure 7C**). In terms of carbon storage polymers, the amount of glycogen on both strains decreased by two-fold in the presence of salt (**Figure 7D**), whereas PHB was not detected under the conditions tested. The latter was detected in the WT strain both under the conditions used in this study and the conditions used by Santos et al. (2021a) (data not shown).

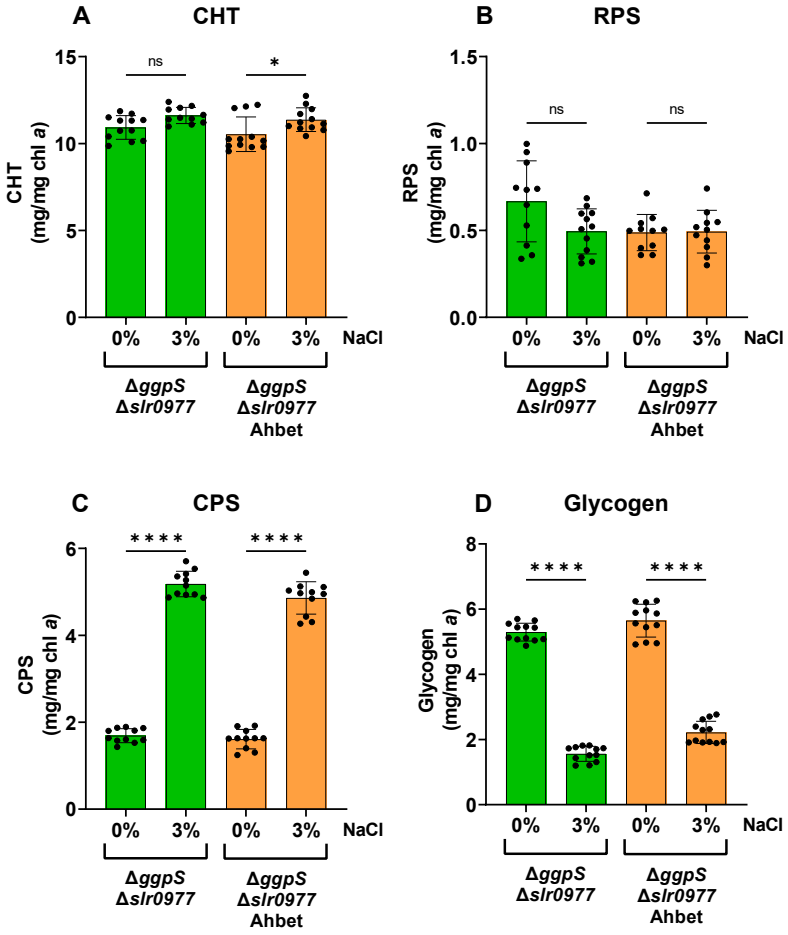



**Figure 7.** Effect of NaCl on CHT – total carbohydrates **(A)**, CPS – capsular polysaccharides **(B)**, RPS – released polysaccharides **(C)**, and glycogen **(D)** produced by *Syn*6803 Δ*ggpS*Δ*slr0977* mutant with and without the Ahbet device. Cultures were grown in BG11 or BG11 supplemented with 3% NaCl at 30ºC with shaking at 150 rpm under a 12 h (25 µE/$m^2$/s) light/12 h dark cycle. Results were expressed as milligrams per milligram of chlorophyll *a* (chl *a*), and obtained from four biological replicates with technical triplicates. Statistical analysis was performed using one-way ANOVA. Statistically significant differences are identified: **** ($p \leq 0.0001$), * ($p \leq 0.05$), and ns (not significant).

## 3.3 - Quantification of compatible solutes in Δ*ggpS*Δ*slr0977* strains

In addition to carbohydrates, the CS content (namely glutamate, sucrose, glucosylglycerol and glycine betaine) was quantified in Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet strains grown in BG11 (0% NaCl) or BG11 supplemented with 3% NaCl (Figure 8). In both strains, the absence of glucosylglycerol production was confirmed whereas glutamate and sucrose accumulated in a salinity dependent manner, with the latter only being detected in presence of NaCl as expected. The sucrose contents found in the Δ*ggpS*Δ*slr0977* (8.72 µmol/mg protein, preliminary results) and Δ*ggpS*Δ*slr0977* Ahbet (12.79 µmol/mg protein) were higher than what was reported for Δ*ggpS* Ahbet (4.57 µmol/mg protein) (Ferreira et al., 2022). Surprisingly, the Δ*ggpS*Δ*slr0977* Ahbet mutant was unable to produce GB.

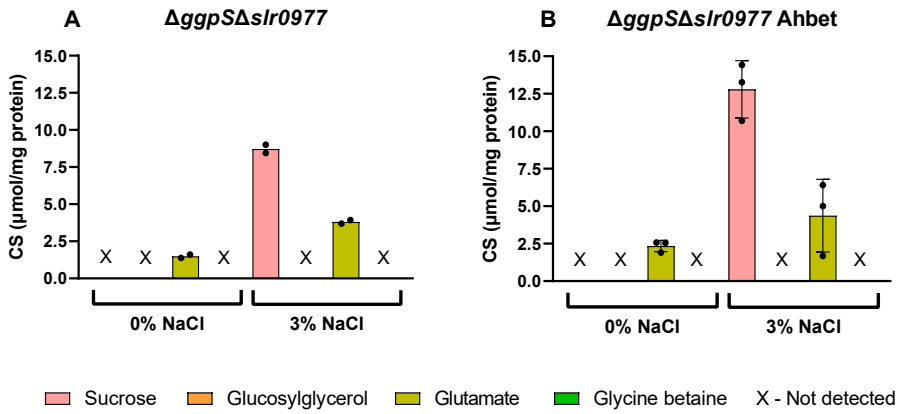



**Figure 8.** Effect of NaCl on the synthesis of the compatible solutes sucrose, glucosylglycerol, glutamate and glycine betaine by Δ*ggpS*Δ*slr0977* Ahbet (**A**) and Δ*ggpS*Δ*slr0977* Ahbet (**B**). Cultures were grown in BG11 or BG11 supplemented with 3% NaCl at 30ºC with shaking at 150 rpm under a 12 h (25 µE/$m^2$/s) light/12 h dark cycle. Quantification of the compatible solutes was performed by proton NMR (in collaboration with Dr. Pedro Lamosa, CERMAX, ITQB, Universidade Nova de Lisboa). Results were expressed as micromoles of CS per milligram of protein, and obtained from at least two biological replicates. Statistical analysis was performed where applicable using the *t*-test.

Subsequently, it was investigated whether the absence of GB production in the Δ*ggpS*Δ*slr0977* Ahbet was due to absence of the device. For this purpose, a PCR targeting the *gsmt* gene present in the Ahbet device was performed using as template the cell extracts obtained from an aliquot of the same biomass used for CS quantification (Figure 9). The presence of an amplicon on the PCR reactions performed, using the Δ*ggpS*Δ*slr0977* Ahbet samples, confirms the presence of the synthetic device and ruled out the possibility of the absence of GB production due to loss of the Ahbet device.

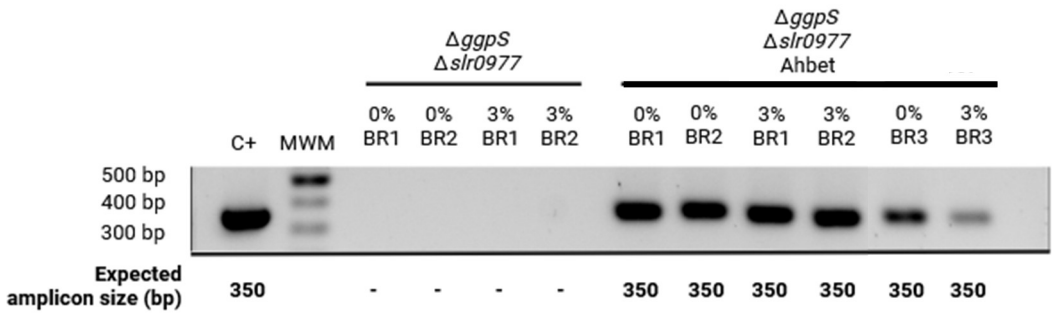



**Figure 9.** Agarose gel showing the polymerase chain reactions (PCR) using the cell crude extracts as template and the *gsmt* primers targeting the Ahbet cluster. The presence of an amplicon confirms the presence of the Ahbet device in the Δ*ggpS*Δ*slr0977* Ahbet samples. Reactions were performed using the primers SD_GSMT_F and SD_GSMT_R (**Table 2**). C+ - positive control using the pSEVA 351 plasmid harbouring the Ahbet device. MWM – molecular weight marker GeneRuler DNA Ladder Mix (Thermo Scientific™).

Given this result, the Δ*ggpS*Δ*slr0977* Ahbet samples were used for the amplification of the full sequence of the Ahbet device (3500 bp) by PCR. Currently, the Ahbet amplicon is being sequenced to verify the existence of mutations in the Ahbet device that could possibly prevent its functionality and therefore justify the absence of GB production in the Δ*ggpS*Δ*slr0977* Ahbet strain.

## 3.4 - Extraction of RNA for the transcription analysis by RNAseq

Previously, the characterization of *Syn*6803 GB-producing strains Δ*ggpS* Ahbet and Δ*slr0977* Ahbet revealed key differences between them. The Δ*ggpS* Ahbet showed improved growth in comparison with the WT and Δ*ggpS* under 0% and 3% NaCl (Ferreira et al., 2022). In contrast, the Δ*slr0977* Ahbet did not display any growth improvement, but the production of GB was significantly increased when compared with Δ*ggpS* Ahbet (unpublished work), that was previously described as the best GB producer (Ferreira et al., 2022). As such, these strains were selected for a transcriptome analysis by RNAseq, to gain insights into the biological mechanisms underlying the "improved growth" and the "improved production" phenotypes. For this purpose, samples for RNA extraction were collected from *Syn*6803 WT, Δ*slr0977*, Δ*slr0977* Ahbet, and Δ*ggpS* Ahbet cultures grown in standard BG11 medium or in BG11 supplemented with 3 and 5% (wt/vol) NaCl. After sample processing, the RNA purity was assessed by the determination of the $A_{260}/A_{280}$ and $A_{260}/A_{230}$ ratios, which showed values within the expected ranges (2.07-2.14 and 2.3-2.7, respectively; **Supplementary Table 1**). These results indicate that the levels of protein or organic compounds that co-purified with the RNA are not significant. Subsequently, the absence of RNA contamination with genomic DNA was confirmed by PCR, with no amplification of DNA in any of the samples (**Figure 10**). The final RNA quality control performed was the evaluation of the sample integrity (**Supplementary Table 1**). With the exception of samples WT 0% NaCl and Δ*slr0977* 0% and 5% NaCl, all other RNAs showed an RNA integrity number (RIN) above 6.0, which is required for the transcriptome analysis. For the samples with RINs below 6, new samples have already been collected and RNA will be extracted.

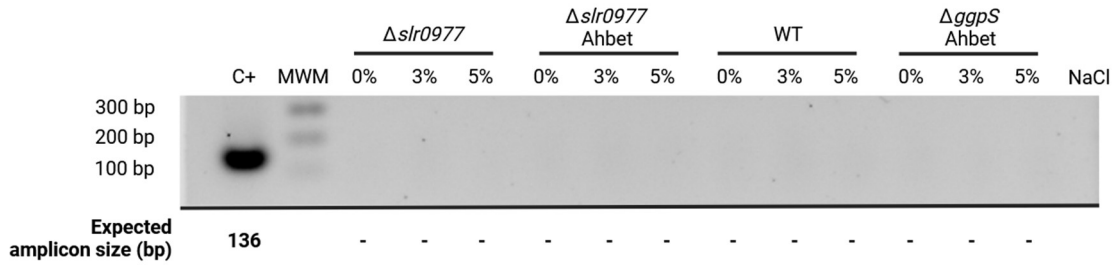



**Figure 10.** Agarose gel showing the polymerase chain reactions (PCR) using the RNA samples as template and the *rnpB* primers. The absence of amplification confirms that there is no contamination of the RNA with genomic DNA. Reactions were performed using the primers rnpBF1 and rnpBR1 (Pinto et al., 2012). C+ - positive control using wild-type (WT) genomic DNA; MWM – molecular weight marker GeneRuler DNA Ladder Mix (Thermo Scientific™).

# Discussion

Glycine betaine (GB) is an extremely valuable compound from a commercial standpoint, however, its production methods leave much to be desired. The use of cyanobacteria, specifically *Syn*6803, as a photoautotrophic chassis for the heterologous production of GB is appealing given its sustainability. *Syn*6803 mutants Δ*ggpS* (deficient in the production of the native solute glucosylglycerol) and Δ*slr0977* (putatively impaired in the production of extracellular polymeric substances - EPS) yielded the highest GB production in this cyanobacterium, with the latter surpassing the former by 59% (Ferreira et al., 2022; unpublished work). In an attempt to improve production further, the double mutant Δ*ggpS*Δ*slr0977* was generated (unpublished work), and characterized here.

The introduction of the Ahbet device into the Δ*ggpS*Δ*slr0977* background did not compromise its growth under the salinities tested, suggesting that there is no metabolic burden associated with the production of GB. This finding is in line with the reported by Ferreira et al. (2022), as simulations performed using the *i*Syn811 metabolic model showed that the synthesis of GB would have a lesser impact on cell growth than the native compatible solutes (CS). At a 3% salt concentration, Δ*ggpS*Δ*slr0977* Ahbet exhibits similar growth as the Δ*slr0977* Ahbet strain, suggesting that, under this salinity, the phenotype of the Δ*slr0977* mutant prevails in the double mutant. However, when increasing the salt concentration to 5% NaCl, the Δ*ggpS*Δ*slr0977* Ahbet exhibits a growth arrest similar to the observed for Δ*ggpS* Ahbet (Ferreira et al., 2022). This implies that GG is required for growth at 5% NaCl; or alternatively, the amount of GB synthesized is not sufficient to support growth at this salinity.

Previous work has demonstrated that there is a trade-off between the synthesis of various CS and other carbon sinks, either in the form of reserve polymers such as glycogen and PHB, or the production of EPS (Baran et al., 2017; Du et al., 2013; Ferreira et al., 2022; Kirsch et al., 2017). To study this interplay in the Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet strains, the carbohydrate pools were analysed. In terms of total carbohydrates, a small increment (8%, $p \leq 0.05$) was detected in the Δ*ggpS*Δ*slr0977* Ahbet strain grown under 3% NaCl; however, it is difficult to determine its biological relevance. In contrast, a significant increase in capsular polysaccharides (CPS) was detected, further demonstrating that these polysaccharides play a protective role under salinity conditions (Ferreira et al., 2022; Jittawuttipoka et al., 2013; Pereira et al., 2009). This seems to assume particular importance in mutants deficient in the production of glucosylglycerol (GG) grown under 3% NaCl, such as Δ*ggpS* and Δ*ggpS*Δ*slr0977* with or without the Ahbet, for which the CPS increase by more than two-fold (Ferreira et al., 2022; unpublished work). The released polysaccharides (RPS) remain unchanged when cells are exposed to salt, thus implying a smaller contribution to the salt stress response (compared

with CPS). These results obtained for the EPS (CPS and RPS) stress out that the deletion of *slr0977*, which was reported to be a key player in EPS export in the absence of salt (Santos et al., 2021a), does not result in the impairment of EPS under salinity conditions, pointing towards the existence of alternative producing pathways or redundancy of existing ones. In fact, the analysis of *Syn*6803's transcriptome grown under 4% NaCl showed the upregulation of the *kpsM* homologue *slr2107*, and also of *slr0896* and *slr0488* (Klähn et al., 2021). The latter two are *wzx* homologues, encoding the protein responsible for EPS export in the Wzy-dependent pathway (Pereira et al., 2015; Whitfield et al., 2020). In addition to EPS, the pools of the reserve polymers were also analysed, with glycogen showing a decrease in presence of salt, which is in agreement with previous reports (Ferreira et al., 2022; unpublished work). This observation gives strength to the assumption that, under salinity conditions, glycogen is degraded to redirect resources for cell protection through the synthesis of EPS and CS (Ferreira et al., 2022) (**Figure 11**).

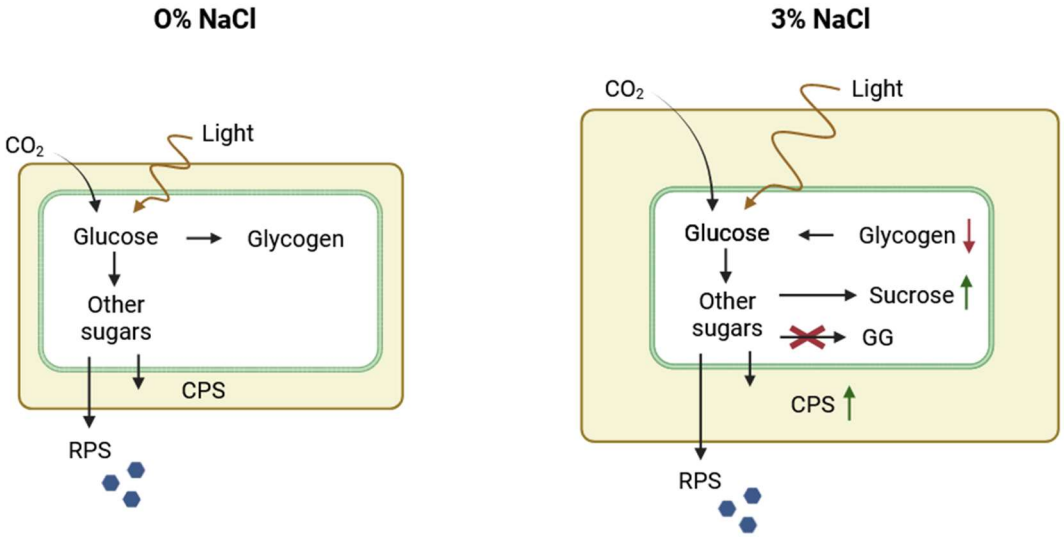



**Figure 11.** Schematic representation of the putative redirection of resources under salinity conditions in glucosylglycerol (GG) deficient *Syn*6803 strains, such as Δ*ggpS* and Δ*ggpS*Δ*slr0977* with and without the Ahbet device. CPS – capsular polysaccharides; RPS – released polysaccharides. Created using biorender.com.

Previously, the levels of polyhydroxybutyrate (PHB), another carbon storage polymer, were shown to be increased in a *Syn*6803 Δ*slr0977* mutant grown in the absence of salt (Santos et al., 2021a). Therefore, the quantification of PHB in Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet was pursued to verify whether resources are also deviated from the synthesis of this polymer to the production of CS. In the conditions tested here, PHB could not be detected in the absence or presence of salt, showing that it is not accumulated in these mutants. This suggests that the contribution of PHB to the resources that are directed to cell protection and homeostasis under salinity conditions is null or very limited.

# Final remarks and future work

Glycine betaine (GB) is an industrially relevant compound, with many possible applications. The use of biological strategies for synthesizing GB at a large scale will allow its sustainable production, which current production methods fail to address (DuPont, 2015). In this work, the characterization of the double mutant Δ*ggpS*Δ*slr0977* and its validation as a *Syn*6803-based chassis for GB production were pursued. Analysis of the Δ*ggpS*Δ*slr0977* and Δ*ggpS*Δ*slr0977* Ahbet mutants' behaviour under different salinity conditions reveal that the Ahbet device is not detrimental to growth. In addition, a salt concentration of 3% has no significant impact on growth, which enables the possibility to use seawater instead of potable water for *Syn*6803 cultivation. The carbohydrate content of both mutants was also quantified, showing a trade-off between CPS and glycogen in salinity conditions. The sharp increase of CPS reveals that these strains rely on these polysaccharides to resist salt stress, and also that the *slr*0977 homologue is not necessary for their production in salinity conditions. The quantification of the compatible solutes in Δ*ggpS*Δ*slr0977* strains shows that Δ*ggpS*Δ*slr0977* Ahbet does not produce GB, and the reasons for this need to be clarified. Nevertheless, the results obtained so far show that essential knowledge on the metabolism of *Syn*6803 under salinity conditions is still lacking. To enable rational engineering for GB production, it is essential not only to understand the transcriptomic and proteomic landscape of the cell but also to clarify the resources that are redirected when cells are exposed to salt, for example, using RNAseq, proteomic, and metabolomic approaches, respectively. This would reveal the molecular basis for the differences seen in the phenotypes of *Syn*6803 GB producing strains, namely the Δ*ggpS* Ahbet and the Δ*slr0977* Ahbet mutants, as well as help to refine existing metabolic models for *Syn*6803 (Montagud et al., 2011). With this tool and the knowledge accumulated, generating a GB-producing chassis that could rival *E. coli* in terms of productivity is well within the realm of possibility.

# References


Abed, R. M. M., Dobretsov, S., & Sudesh, K. (2009). Applications of cyanobacteria in biotechnology. *Journal of Applied Microbiology*, *106*(1), 1–12. https://doi.org/10.1111/J.1365-2672.2008.03918.X

Aquilina, G., Bach, A., Bampidis, V., De, M., Bastos, L., Guido Costa, L., Flachowsky, G., Gralak, A., Hogstrand, C., Leng, L., López-Puente, S., Martelli, G., Mayo, B., Ramos, F., Renshaw, D., Rychen, G., Saarela, M., Sejrsen, K., Van Beelen, P., … Westendorf, J. (2013). Scientific opinion on the safety and efficacy of betaine (betaine anhydrous and betaine hydrochloride) as a feed additive for all animal species based on a dossier submitted by VITAC EEIG. *EFSA Journal*, *11*(5), 3210. https://doi.org/10.2903/j.efsa.2013.3210

Baran, R., Lau, R., Bowen, B. P., Diamond, S., Jose, N., Garcia-Pichel, F., & Northen, T. R. (2017). Extensive turnover of compatible solutes in cyanobacteria revealed by deuterium oxide ($D_2O$) stable isotope probing. *ACS Chemical Biology*, *12*(3), 674–681. https://doi.org/10.1021/acschembio.6b00890

Brocks, J. J., Jarrett, A. J. M., Sirantoine, E., Hallmann, C., Hoshino, Y., & Liyanage, T. (2017). The rise of algae in Cryogenian oceans and the emergence of animals. *Nature*, *548*(7669), 578–581. https://doi.org/10.1038/nature23457

Chen, C. Y., Yeh, K. L., Aisyah, R., Lee, D. J., & Chang, J. S. (2011). Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review. *Bioresource Technology*, *102*(1), 71–81. https://doi.org/10.1016/j.biortech.2010.06.159

Cleland, D., Krader, P., McCree, C., Tang, J., & Emerson, D. (2004). Glycine betaine as a cryoprotectant for prokaryotes. *Journal of Microbiological Methods*, *58*(1), 31–38. https://doi.org/10.1016/j.mimet.2004.02.015

Cray, J. A., Stevenson, A., Ball, P., Bankar, S. B., Eleutherio, E. C. A., Ezeji, T. C., Singhal, R. S., Thevelein, J. M., Timson, D. J., & Hallsworth, J. E. (2015). Chaotropicity: a key factor in product tolerance of biofuel-producing microorganisms. *Current Opinion in Biotechnology*, *33*, 228–259. https://doi.org/10.1016/J.COPBIO.2015.02.010

Day, C. R., & Kempson, S. A. (2016). Betaine chemistry, roles, and potential use in liver disease. *Biochimica et Biophysica Acta - General Subjects*, *1860*(6), 1098–1106. https://doi.org/10.1016/j.bbagen.2016.02.001

de Lima Alves, F., Stevenson, A., Baxter, E., Gillion, J. L. M., Hejazi, F., Hayes, S., Morrison, I. E. G., Prior, B. A., McGenity, T. J., Rangel, D. E. N., Magan, N., Timmis, K. N., & Hallsworth, J. E. (2015). Concomitant osmotic and chaotropicity-induced stresses in *Aspergillus wentii*: compatible solutes determine the biotic window. *Current Genetics*, *61*(3), 457–477. https://doi.org/10.1007/s00294-015-0496-8

Du, W., Liang, F., Duan, Y., Tan, X., & Lu, X. (2013). Exploring the photosynthetic production

capacity of sucrose by cyanobacteria. *Metabolic Engineering*, *19*, 17–25. https://doi.org/10.1016/j.ymben.2013.05.001

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Analytical Chemistry*, *28*(3), 350–356. https://doi.org/10.1021/ac60111a017

DuPont. (2015). *Betafin® natural betaine sustainable and substantiated*. http://animalnutrition.dupont.com/fileadmin/user_upload/live/animal_nutrition/documents/open/Betafin__Natural_Betaine_Lifecycle_Analysis.pdf

Ernst, A., Kirschenlohr, H., Diez, J., & Böger, P. (1984). Glycogen content and nitrogenase activity in *Anabaena variabilis*. *Archives of Microbiology*, *140*(2–3), 120–125. https://doi.org/10.1007/BF00454913

Estrin, N. F., Ferguson, J. P., Coale, W. C., & Rhoads, L. C. (Eds.). (1997). *International Cosmetic Ingredient Dictionary and Handbook* (2nd ed.). Cosmetic, Toiletry & Fragrance Association, U.S.

European Medicines Agency. (2016). *Cystadane (betaine anhydrous) | EMA/704352/2016*.

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

Ferreira, E. A., Pacheco, C. C., Rodrigues, J. S., Pinto, F., Lamosa, P., Fuente, D., Urchueguía, J., & Tamagnini, P. (2022). Heterologous production of glycine betaine using *Synechocystis* sp. PCC 6803-based chassis lacking native compatible solutes. *Frontiers in Bioengineering and Biotechnology*, *9*(January), 1–19. https://doi.org/10.3389/fbioe.2021.821075

Fu, P. (2009). Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion. *Journal of Chemical Technology & Biotechnology*, *84*(4), 473–483. https://doi.org/10.1002/JCTB.2065

Galinski, E. A. (1995). Osmoadaptation in bacteria. *Advances in Microbial Physiology*, *37*(C), 273–328. https://doi.org/10.1016/S0065-2911(08)60148-4

Garcia-Pichel, F., Zehr, J. P., Bhattacharya, D., & Pakrasi, H. B. (2020). What's in a name? The case of cyanobacteria. *Journal of Phycology*, *56*(1), 1–5. https://doi.org/10.1111/jpy.12934

Hagemann, M. (2011). Molecular biology of cyanobacterial salt acclimation. *FEMS Microbiology Reviews*, *35*(1), 87–123. https://doi.org/10.1111/j.1574-6976.2010.00234.x

Hagemann, M. (2013). Genomics of salt acclimation: synthesis of compatible solutes among cyanobacteria. In F. Chauvat & C. Cassier-Chauvat (Eds.), *Advances in Botanical Research* (Vol. 65, pp. 27–55). Academic Press. https://doi.org/10.1016/B978-0-12-394313-2.00002-0

Hagemann, M., & Erdmann, N. (1994). Activation and pathway of glucosylglycerol synthesis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology*, *140*(6), 1427–1431. https://doi.org/10.1099/00221287-140-6-1427

Hagemann, M., Golldack, D., Biggins, J., & Erdmann, N. (1993). Salt-dependent protein phosphorylation in the cyanobacterium *Synechocystis* PCC 6803. *FEMS Microbiology Letters*, *113*(2), 205–209. https://doi.org/10.1111/j.1574-6968.1993.tb06515.x

Hall, D. O., & Rao, K. (1999). *Photosynthesis*. Cambridge University Press.

Heikkila, H. O., Melaja, J. A., Millner, D. E. D., & Virtanen, J. J. (1981). *Betaine recovery process | US4359430A*.

Hodgskiss, M. S. W., Crockford, P. W., Peng, Y., Wing, B. A., & Horner, T. J. (2019). A productivity collapse to end Earth's great oxidation. *Proceedings of the National Academy*

*of Sciences of the United States of America*, *116*(35), 17207–17212. https://doi.org/10.1073/pnas.1900325116

Huang, H. H., & Lindblad, P. (2013). Wide-dynamic-range promoters engineered for cyanobacteria. *Journal of Biological Engineering*, *7*(1), 1–11. https://doi.org/10.1186/1754-1611-7-10

Imhoff, J. F., & Rodriguez-Valera, F. (1984). Betaine is the main compatible solute of halophilic eubacteria. *Journal of Bacteriology*, *160*(1), 478–479. https://doi.org/10.1128/jb.160.1.478-479.1984

Jadhav, K., Kushwah, B., & Jadhav, I. (2018). Insight into compatible solutes from halophiles: exploring significant applications in biotechnology. In J. Singh, D. Sharma, G. Kumar, & N. R. Sharma (Eds.), *Microbial Bioprospecting for Sustainable Development* (pp. 291–307). Springer Singapore. https://doi.org/10.1007/978-981-13-0053-0_16

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), 55564. https://doi.org/10.1371/journal.pone.0055564

Jodlbauer, J., Rohr, T., Spadiut, O., Mihovilovic, M. D., & Rudroff, F. (2021). Biocatalysis in green and blue: cyanobacteria. *Trends in Biotechnology*, *39*(9), 875–889. https://doi.org/10.1016/j.tibtech.2020.12.009

Johnson, T. J., Katuwal, S., Anderson, G. A., Gu, L., Zhou, R., & Gibbons, W. R. (2018). Photobioreactor cultivation strategies for microalgae and cyanobacteria. *Biotechnology Progress*, *34*(4), 811–827. https://doi.org/10.1002/btpr.2628

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., … Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes*, *3*(3), 109–136. https://doi.org/10.1093/DNARES/3.3.109

Kanesaki, Y., Shiwa, Y., Tajima, N., Suzuki, M., Watanabe, S., Sato, N., Ikeuchi, M., & Yoshikawa, H. (2012). Identification of substrain-specific mutations by massively parallel whole-genome resequencing of *Synechocystis* sp. PCC 6803. *DNA Research*, *19*(1), 67–79. https://doi.org/10.1093/dnares/dsr042

Kegang, Y. (1997). *Synthesis and separation of trimethyl glycine (betaine) | CN1196351A*.

Kirsch, F., Klähn, S., & Hagemann, M. (2019). Salt-regulated accumulation of the compatible solutes sucrose and glucosylglycerol in cyanobacteria and its biotechnological potential. *Frontiers in Microbiology*, *10*, 2139. https://doi.org/10.3389/fmicb.2019.02139

Kirsch, F., Pade, N., Klähn, S., Hess, W. R., & Hagemann, M. (2017). The glucosylglycerol-degrading enzyme GghA is involved in acclimation to fluctuating salinities by the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Microbiology (United Kingdom)*, *163*(9), 1319–1328. https://doi.org/10.1099/mic.0.000518

Klähn, S., & Hagemann, M. (2011). Compatible solute biosynthesis in cyanobacteria. *Environmental Microbiology*, *13*(3), 551–562. https://doi.org/10.1111/j.1462-2920.2010.02366.x

Klähn, S., Mikkat, S., Riediger, M., Georg, J., Hess, W. R., & Hagemann, M. (2021). Integrative analysis of the salt stress response in cyanobacteria. *Biology Direct*, *16*(1), 1–23. https://doi.org/10.1186/s13062-021-00316-4

Knoll, A. (1980). Archean photoautotrophy: some alternatives and limits. In C. Ponnamperuma & L. Margulis (Eds.), *Limits of Life* (pp. 133–147). Springer Netherlands.

Knoll, A. (2008). Cyanobacteria and earth history. In F. Flores & A. Herrero (Eds.), *The Cyanobacteria: Molecular Biology, Genomics, and Evolution*. (pp. 1–19). Caister Academic Press, Norfolk, UK.

Ladas, N. P., & Papageorgiou, G. C. (2000). Cell turgor: a critical factor for the proliferation of cyanobacteria at unfavorable salinity. *Photosynthesis Research*, *65*(2), 155–164. https://doi.org/10.1023/A:1006423221150

Landfald, B., & Strom, A. R. (1986). Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. *Journal of Bacteriology*, *165*(3), 849–855. https://doi.org/10.1128/jb.165.3.849-855.1986

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, *506*(7488), 307–315. https://doi.org/10.1038/nature13068

Machado, I. M. P., & Atsumi, S. (2012). Cyanobacterial biofuel production. *Journal of Biotechnology*, *162*(1), 50–56. https://doi.org/10.1016/J.JBIOTEC.2012.03.005

Markou, G., Vandamme, D., & Muylaert, K. (2014). Microalgal and cyanobacterial cultivation: the supply of nutrients. *Water Research*, *65*, 186–202. https://doi.org/10.1016/j.watres.2014.07.025

Meeks, J. C., & Castenholz, R. W. (1971). Growth and photosynthesis in an extreme thermophile, *Synechococcus lividus* (*Cyanophyta*). *Archiv Für Mikrobiologie*, *78*(1), 25–41. https://doi.org/10.1007/BF00409086

Messadek, J. (2002). *Glycine betaine and its use as anti-hemorrhagic agent | EP1408949A2*.

Montagud, A., Zelezniak, A., Navarro, E., de Córdoba, P. F., Urchueguía, J. F., & Patil, K. R. (2011). Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803. *Biotechnology Journal*, *6*(3), 330–342. https://doi.org/10.1002/BIOT.201000109

Mota, R., Guimarães, R., Büttel, Z., Rossi, F., Colica, G., Silva, C. J., Santos, C., Gales, L., Zille, A., De Philippis, R., Pereira, S. B., & Tamagnini, P. (2013). Production and characterization of extracellular carbohydrate polymer from *Cyanothece* sp. CCY 0110. *Carbohydrate Polymers*, *92*(2), 1408–1415. https://doi.org/10.1016/j.carbpol.2012.10.070

Nagy, C., Thiel, K., Mulaku, E., Mustila, H., Tamagnini, P., Aro, E. M., Pacheco, C. C., & Kallio, P. (2021). Comparison of alternative integration sites in the chromosome and the native plasmids of the cyanobacterium *Synechocystis* sp. PCC 6803 in respect to expression efficiency and copy number. *Microbial Cell Factories*, *20*(1), 1–18. https://doi.org/10.1186/s12934-021-01622-2

Nozzi, N. E., Oliver, J. W. K., & Atsumi, S. (2013). Cyanobacteria as a platform for biofuel production. *Frontiers in Bioengineering and Biotechnology*, *1*, 7. https://doi.org/10.3389/fbioe.2013.00007

Nyyssölä, A., Kerovuo, J., Kaukinen, P., Von Weymarn, N., & Reinikainen, T. (2000). Extreme halophiles synthesize betaine from glycine by methylation. *Journal of Biological Chemistry*, *275*(29), 22196–22201. https://doi.org/10.1074/jbc.M910111199

Ohbayashi, R., Akai, H., Yoshikawa, H., Hess, W. R., & Watanabe, S. (2016). A tightly inducible riboswitch system in *Synechocystis* sp. PCC 6803. *Journal of General and Applied Microbiology*, *62*(3), 154–159. https://doi.org/10.2323/jgam.2016.02.002

Ollivier, B., Caumette, P., Garcia, J. L., & Mah, R. A. (1994). Anaerobic bacteria from hypersaline environments. *Microbiological Reviews*, *58*(1), 27–38. https://doi.org/10.1128/mmbr.58.1.27-38.1994

Olson, J. M. (2006). Photosynthesis in the Archean era. *Photosynthesis Research*, *88*(2), 109–117. https://doi.org/10.1007/s11120-006-9040-5

Oren, A. (1990). Formation and breakdown of glycine betaine and trimethylamine in

hypersaline environments. *Antonie van Leeuwenhoek*, *58*(4), 291–298. https://doi.org/10.1007/BF00399342

Ostrander, C. M., Nielsen, S. G., Owens, J. D., Kendall, B., Gordon, G. W., Romaniello, S. J., & Anbar, A. D. (2019). Fully oxygenated water columns over continental shelves before the Great Oxidation Event. *Nature Geoscience*, *12*(3), 186–191. https://doi.org/10.1038/s41561-019-0309-7

Pacheco, C. C., Ferreira, E. A., Oliveira, P., & Tamagnini, P. (2021). Synthetic biology of cyanobacteria. In R. Kourist & S. Schmidt (Eds.), *The Autotrophic Biorefinery: Raw Materials from Biotechnology* (pp. 131–172). De Gruyter. https://doi.org/10.1515/9783110550603-006

Pereira, S. B., 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. *Scientific Reports*, *5*(1), 1–16. https://doi.org/10.1038/srep14835

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 Microbiology Reviews*, *33*(5), 917–941. https://doi.org/10.1111/j.1574-6976.2009.00183.x

Pinto, F., Pacheco, C. C., Ferreira, D., Moradas-Ferreira, P., & Tamagnini, P. (2012). Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. *PLoS ONE*, *7*(4), e34983. https://doi.org/10.1371/journal.pone.0034983

Planavsky, N. J., Asael, D., Hofmann, A., Reinhard, C. T., Lalonde, S. V., Knudsen, A., Wang, X., Ossa Ossa, F., Pecoits, E., Smith, A. J. B., Beukes, N. J., Bekker, A., Johnson, T. M., Konhauser, K. O., Lyons, T. W., & Rouxel, O. J. (2014). Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nature Geoscience*, *7*(4), 283–286. https://doi.org/10.1038/ngeo2122

Reed, R. H., & Stewart, W. D. P. (1985). Osmotic adjustment and organic solute accumulation in unicellular cyanobacteria from freshwater and marine habitats. *Marine Biology*, *88*(1), 1–9. https://doi.org/10.1007/BF00393037

Rodionova, M. V., Poudyal, R. S., Tiwari, I., Voloshin, R. A., Zharmukhamedov, S. K., Nam, H. G., Zayadan, B. K., Bruce, B. D., Hou, H. J. M., & Allakhverdiev, S. I. (2017). Biofuel production: challenges and opportunities. *International Journal of Hydrogen Energy*, *42*(12), 8450–8461. https://doi.org/10.1016/J.IJHYDENE.2016.11.125

Rosing, M. T., & Frei, R. (2004). U-rich Archaean sea-floor sediments from Greenland - indications of > 3700 Ma oxygenic photosynthesis. *Earth and Planetary Science Letters*, *217*(3–4), 237–244. https://doi.org/10.1016/S0012-821X(03)00609-5

Sambrook, J., & Russell, D. W. (2006). Agarose gel electrophoresis. *CSH Protocols*, *2006*(1), 477–479. https://doi.org/10.1101/PDB.PROT4020

Sánchez-Baracaldo, P. (2015). Origin of marine planktonic cyanobacteria. *Scientific Reports*, *5*(1), 1–10. https://doi.org/10.1038/srep17418

Sánchez-Baracaldo, P., & Cardona, T. (2020). On the origin of oxygenic photosynthesis and cyanobacteria. *The New Phytologist*, *225*(4), 1440–1446. https://doi.org/10.1111/NPH.16249

Santos, H., Lamosa, P., & Borges, N. (2006). Characterization and quantification of compatible solutes in (hyper)thermophilic microorganisms. *Methods in Microbiology*, *35*, 173–199. https://doi.org/10.1016/S0580-9517(08)70011-4

Santos, M., Pereira, S. B., Flores, C., Príncipe, C., Couto, N., Karunakaran, E., Cravo, S. M., Oliveira, P., & Tamagnini, P. (2021a). 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

Santos, M., Pacheco, C. C., Yao, L., Hudson, E. P., & Tamagnini, P. (2021b). CRISPRi as a tool to repress multiple copies of extracellular polymeric substances (EPS)-related genes in the cyanobacterium *Synechocystis* sp. PCC 6803. *Life*, *11*(11). https://doi.org/10.3390/life11111198

Scott, C., Lyons, T. W., Bekker, A., Shen, Y., Poulton, S. W., Chu, X., & Anbar, A. D. (2008). Tracing the stepwise oxygenation of the Proterozoic ocean. *Nature*, *452*(7186), 456–459. https://doi.org/10.1038/nature06811

Sheldon, R. A., & Woodley, J. M. (2018). Role of biocatalysis in sustainable chemistry. *Chemical Reviews*, *118*(2), 801–838. https://doi.org/10.1021/acs.chemrev.7b00203

Shivanand, P., & Mugeraya, G. (2011). Halophilic bacteria and their compatible solutes - osmoregulation and potential applications. *Current Science*, *100*(10), 1516–1521.

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 Research*, *41*(Database issue), D666. https://doi.org/10.1093/NAR/GKS1119

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/mmbr.35.2.171-205.1971

Thiel, K., Vuorio, E., Aro, E. M., & Kallio, P. T. (2017). The effect of enhanced acetate influx on *Synechocystis* sp. PCC 6803 metabolism. *Microbial Cell Factories*, *16*(1), 1–12. https://doi.org/10.1186/s12934-017-0640-x

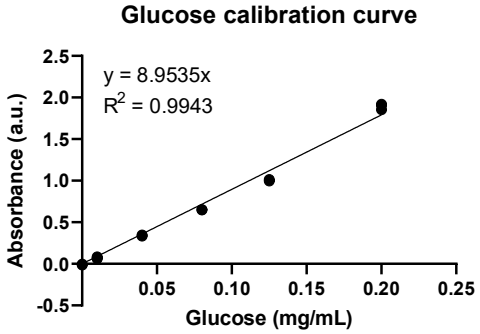
Trautmann, D., Voß, B., Wilde, A., Al-Babili, S., & Hess, W. R. (2012). Microevolution in cyanobacteria: re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. *DNA Research*, *19*(6), 435–448. https://doi.org/10.1093/dnares/dss024

Waditee, R., Tanaka, Y., Aoki, K., Hibino, T., Jikuya, H., Takano, J., Takabe, T., & Takabe, T. (2003). Isolation and functional characterization of *N*-methyltransferases that catalyze betaine synthesis from glycine in a halotolerant photosynthetic organism Aphanothece halophytica. *Journal of Biological Chemistry*, *278*(7), 4932–4942. https://doi.org/10.1074/jbc.M210970200

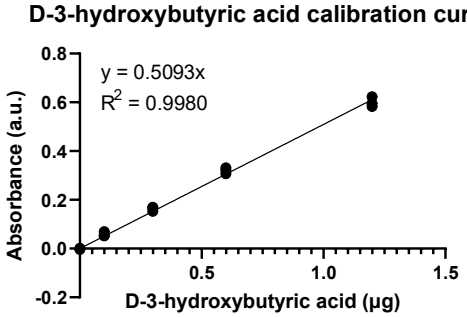
Weissensteiner, T., & Lanchbury, J. S. (1996). Strategy for controlling preferential amplification and avoiding false negatives in PCR typing. *BioTechniques*, *21*(6), 1102–1108. https://doi.org/10.2144/96216rr03

Whitfield, C., Wear, S. S., & Sande, C. (2020). Assembly of bacterial capsular polysaccharides and exopolysaccharides. *Annual Review of Microbiology*, *74*, 521–543. https://doi.org/10.1146/annurev-micro-011420-075607

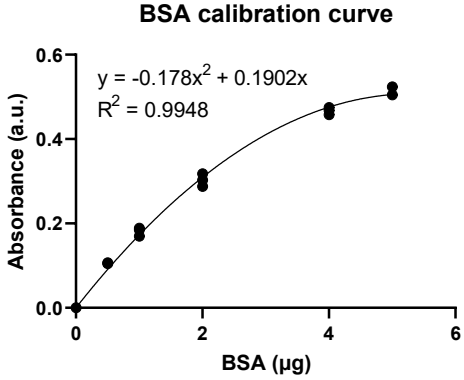
# Annexes




**Supplementary Figure 1.** Calibration curve for glucose concentration versus absorbance.




**Supplementary Figure 2.** Calibration curve for D-3-hydroxybutyric acid mass versus absorbance.




**Supplementary Figure 3.** Calibration curve for bovine serum albumin (BSA) mass versus absorbance.

**Supplementary Table 1.** Quantification of RNA concentration and parameters used to assess RNA purity (260/280 and 260/230), and integrity (RIN – RNA integrity number).

| Sample | [RNA] (ng/µL) | 260/280 | 260/230 | RIN |
|---|---|---|---|---|
| Δ*slr0977* 0% NaCl | 1574.07 | 2.12 | 2.55 | 6.7 |
| Δ*slr0977* 3% NaCl | 1172.92 | 2.12 | 2.56 | 6.2 |
| Δ*slr0977* Ahbet 0% NaCl | 1858.93 | 2.09 | 2.51 | 5.8 |
| Δ*slr0977* Ahbet 3% NaCl | 1038.84 | 2.11 | 2.34 | 7.4 |
| Δ*slr0977* Ahbet 5% NaCl | 1039.58 | 2.1 | 2.46 | 3.7 |
| WT 0% NaCl | 1751.79 | 2.09 | 2.59 | 4 |
| WT 3% NaCl | 1654.65 | 2.11 | 2.54 | 7.2 |
| WT 5% NaCl | 1456.2 | 2.14 | 2.56 | 6.4 |
| Δ*ggpS* Ahbet 0% NaCl | 1859.52 | 2.1 | 2.58 | 7.5 |
| Δ*ggpS* Ahbet 3% NaCl | 2144.96 | 2.07 | 2.58 | 7 |
| Δ*ggpS* Ahbet 5% NaCl | 1652.72 | 2.11 | 2.56 | 6.7 |