Identification of promoter elements in the Dolichospermum circinale AWQC131C saxitoxin gene cluster and the experimental analysis of their use for heterologous expression.

Paul Michael D'Agostino  
Technische Universitat Munchen

Bakir Al-Sinawi  
University of New South Wales

Rabia Mazmouz  
University of New South Wales

Julia Muenchhoff  
University of New South Wales

Brett Anthony Neilan  
University of Newcastle

Michelle Moffitt (✉ m.moffitt@westernsydney.edu.au)  
University of Western Sydney  https://orcid.org/0000-0001-9872-3175

Research article

Keywords: cyanobacteria, Dolichospermum circinale, Anabaena circinalis, saxitoxin, paralytic shellfish toxins, regulation, promoter, transcription, heterologous expression

Posted Date: February 3rd, 2020

DOI: https://doi.org/10.21203/rs.2.15635/v3

License: ☕️ ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at BMC Microbiology on February 17th, 2020. See the published version at https://doi.org/10.1186/s12866-020-1720-3.
Abstract

Abstract Background: Dolichospermum circinale is a filamentous bloom-forming cyanobacterium responsible for biosynthesis of the paralytic shellfish toxins (PST), including saxitoxin. PSTs are neurotoxins and in their purified form are important analytical standards for monitoring the quality of water and seafood and biomedical research tools for studying neuronal sodium channels. More recently, PSTs have been recognised for their utility as local anaesthetics. Characterisation of the transcriptional elements within the saxitoxin (sxt) biosynthetic gene cluster (BGC) is a first step towards accessing these molecules for biotechnology. Results: In D. circinale AWQC131C the sxt BGC is transcribed from two bidirectional promoter regions encoding five individual promoters. These promoters were identified experimentally using 5’ RACE and their activity assessed via coupling to a lux reporter system in E. coli and Synechocystis sp. PCC 6803. Transcription of the predicted drug/metabolite transporter (DMT) encoded by sxtPER was found to initiate from two promoters, P sxtPER1 and P sxtPER2. In E. coli, strong expression of lux from P sxtP, P sxtD and sxtPER1 was observed while expression from P orf24 and P sxtPER2 was remarkably weaker. In contrast, heterologous expression in Synechocystis sp. PCC6803 showed that expression of lux from P sxtP, P sxtPER, and P orf24 promoters was statistically higher compared to the non-promoter control, while P sxtD showed poor activity under the described conditions. Conclusions: Both of the heterologous hosts investigated in this study exhibited high expression levels from three of the five sxt promoters. These results indicate that the majority of the native sxt promoters appear active in different heterologous hosts, simplifying initial cloning efforts. Therefore, heterologous expression of the sxt BGC in either E. coli or Synechocystis could be a viable first option for producing PSTs for industrial or biomedical purposes.

Background

Saxitoxin (STX) is a neurotoxin produced by cyanobacteria and dinoflagellates and is a member of the broader group of alkaloids known as the paralytic shellfish toxins (PSTs) [1]. When high concentrations of PSTs are consumed by humans, acute poisoning can lead to death due to respiratory paralysis [2-4]. Therefore, PSTs are needed as analytical standards for the monitoring and protection of commercial seafood and freshwater reservoirs, as well as for use in biomedical research. While the PSTs pose a significant public health risk and economic burden on society during algal bloom events, their scientific and pharmaceutical potential is well known [3, 4]. Purified PSTs have been a critical tool for researchers investigating neuronal sodium channels, where the toxins specifically block site 1 of voltage-gated sodium channels [5]. Under controlled administration, PSTs are potent anaesthetics, particularly in combination with other local anaesthetics [6, 7]. Further attempts to utilise STX in clinical trials are hindered by its toxicity, but more recent approaches, such as generating liposomal formulations of STX, resulted in blockage of the sciatic nerves in rats with no myotoxic, cytotoxic or neurotoxic effects [8]. It has been postulated that the same delivery could provide effective localised treatment for severe joint pain [9]. Other PSTs, such as the gonyautoxins (GTXs), also have clinical potential and have been utilised for the treatment of anal fissures and chronic tension type headaches [10-12].
Obtaining significant quantities of purified PSTs for clinical research or water quality analysis is difficult. Chemical synthesis and biocatalytic synthesis of PSTs is complex, difficult to scale up and may not produce all relevant naturally occurring isoforms [13, 14]. The most common form of obtaining purified compounds involves extraction and isolation from dinoflagellate blooms, cyanobacterial cultures or contaminated shellfish coupled with synthetic conversion to additional PSTs[15-17]. Analytical calibration standards are commercially available from the National Research Council Canada (NRC), which are obtained from extractions of PST producing dinoflagellates or contaminated shellfish, and in some instances, semi-synthetic conversions of PSTs are required to obtain or broad array of analogues. Thus, the process is difficult, inefficient and costly [18]. These issues clearly highlight the need for an alternative and reliable method for the production and purification of commercial quantities of PSTs.

Heterologous expression of cyanobacterial biosynthetic gene clusters (BGCs) is not well established, in comparison to heterologous expression of Actinomycete BGCs. *Streptomyces* expression hosts successfully produced more than 90 Actinomycete NPs, most of which are heterologously expressed using native promoters [19]. On the other hand, only 10 cyanobacterial NPs have been successfully produced from heterologous systems using both native and heterologous promoters [20], highlighting the need to better understand the function of native promoters in heterologous hosts. *E. coli* is a suitable host for the heterologous expression of cyanobacterial pathways based on its fast growth rate as previously demonstrated [21, 22]. Initial studies used native promoters to produce the ribosomal peptides patellamide A and C, and the microviridins [23, 24]. Recently, there has been a focus on the heterologous expression of cyanobacterial natural product biosynthetic gene clusters (BGC) including the lyngbyatoxin (ltx) BGC in *E. coli* using the tetracycline-inducible PtetO promoter [21, 25]. The ltx BGC has been the focus of multiple heterologous expression studies due to its relatively small size. It has also been expressed in the cyanobacterium *Anabaena* sp. PCC 7120, and *E. coli* GB05-MtaA [26, 27]. While the native ltx promoters were active in *Anabaena* sp. PCC 7120 and drive the production of lyngbyatoxin A, the native promoters were not active in *E. coli*. The addition of *Anabaena* sp. PCC 7120 sigma factors to the *E. coli* host also failed to induce expression of lyngbyatoxin A, suggesting that the heterologous host was unable to recognise the cyanobacterial ribosome binding sites [28]. Subsequently, titres of lyngbyatoxin A close to the native producer were achieved when using the cyanobacteria *Anabaena* sp. PCC 7120 as a heterologous host [26]. This highlights that discrepancies between cyanobacterial promoter efficiencies in different host organisms remain poorly understood. Therefore, it is beneficial to test the activity of promoters in heterologous hosts and expression vectors using reporter systems prior to cloning complex biosynthetic pathways for biotechnological applications [29].

For the most part, cyanobacterial transcription machinery is similar to that found in *E. coli*, with the main difference being the widespread absence of the -35 hexamer in cyanobacteria, which is believed to be replaced by a transcription factor binding site to initiate transcription [30]. In *E. coli*, σ70 is able to recognise the majority of promoters while in cyanobacteria, a range of different sigma factors have been identified [31, 32].
The saxitoxin (sxt) BGC has been characterised in six cyanobacterial species from the order Nostocales and one from the order Oscillatoriales [33-37]. Each sxt BGC encodes a ‘core’ set of enzymes putatively responsible for STX biosynthesis, supplemented with ‘tailoring’ and ‘auxiliary’ genes that give rise to PST analogues or perform functions after PST biosynthesis. Information regarding the regulation of transcriptional elements of cyanobacterial secondary metabolite biosynthesis remains limited to the microcystin (mcy) and jamaicamide (jam) BGCs [38-43]. However, the regulation of PSTs at the molecular level, including the transcriptional elements of the sxt BGC remains largely unknown [44, 45].

Here, we identify the transcription units of the sxt BGC within the cyanobacterium Dolichospermum circinale AWQC131C which enabled the experimental isolation of five promoter regions. We then assessed the reliability of a lux reporter system to assay the activity of uncharacterised cyanobacterial promoters in the heterologous hosts E. coli and Synechocystis sp. PCC 6803 for the first time. Characterisation of these cyanobacterial promoters and determination of their activity in E. coli and Synechocystis, highlights the unpredictability of cyanobacterial promoters of natural product BGCs in heterologous hosts. This study is the first essential phase of understanding the cloning strategy expression of PST biosynthesis, identifying the need for promoter engineering or exchange in future experiments.

**Results**

**Identification of transcriptional units within the sxt biosynthetic gene cluster**

Reverse-transcriptase PCR revealed that the sxt BGC in D. circinale AWQC131C is transcribed as five transcriptional units from two bidirectional promoter regions (Figure 1; Figure S1). All five transcripts appear to be constitutively expressed under standard laboratory conditions, since sxt mRNA was detected across all time points. Operon 1, sxtDV*EABC (* indicates disrupted ORF of sxtV[34]), spans 7.3 kb, is transcribed in the reverse direction and encodes several proteins predicted to be involved early in PST biosynthesis. Operon 2, sxtPQR, spans 3.5 kb and is transcribed in the forward direction. The catalytic functions of SxtP, SxtQ and SxtR are unknown, but they are likely to be essential for PST biosynthesis as their presence and organisation is conserved amongst all reported sxt clusters. The third transcriptional unit is monocistronic and encodes SxtPER, a putative permease of the drug/metabolite transporter family of proteins and is transcribed from two promoters, as discussed further below. Operon 4, is transcribed in the forward direction and spans 12.8 kb. Operon 4 encodes a protein of unknown function, Orf24 that is conserved in most sxt clusters, followed by genes encoding twelve enzymes involved in PST biosynthesis, resulting in the polycistron orf24sxtSTUNGHMIJKLO.

The 3’ ends of operons 1-4 were bioinformatically screened for putative Rho-dependent and Rho-independent transcriptional termination sites using the programs TransTerm and TranstermHP, respectively [46, 47]. Rho-independent transcriptional termination sites were identified in the non-coding regions of three out of four sxt mRNA transcripts (Table S1). Rho-dependent or Rho-independent termination sites were not identified in the sequence of mRNA encoding operon 1.
Transcription start sites and promoter regions of the \( \text{sxt} \) operons

The transcriptional start sites (TSSs) of each operon were experimentally identified via 5' rapid amplification of cDNA ends (5' RACE) (Table 1; Figure S2). The upstream region of each TSS was screened for a promoter sequence consistent with conserved binding sequences of group 1, 2 and 3 sigma factors [31]. All promoters identified in this study displayed sequence similarity to the consensus -10 hexamer (Pribnow box) of the prokaryotic RNA polymerase binding site, while there was sporadic presence of the -35 hexamer binding site (Table 1). These results suggest that the \( \text{sxt} \) promoters of \( \text{D. circinale} \) AWQC131C are activated by an RNA polymerase core enzyme in conjunction with a group 1 or group 2 sigma factor [31, 48]. For the -10 promoter sequences identified, a search was conducted for an extended -10 binding site and upstream (UP) element. The 5' untranslated region (UTR) of each operon was also bioinformatically screened for the presence of consensus ribosomal binding site (RBS) sequences, although previously reported bioinformatics surveys of cyanobacterial genomes were not able to identify the consensus RBS sequence in all genes [49, 50]. Based on the 5' RACE and bioinformatics data, the \( \text{D. circinale} \) AWQC131C \( \text{sxt} \) BGC includes of a total of five TSSs under standard culture conditions (Figure 2).

Operon 1 (\( \text{sxtDV*EABC} \)) contains a short 5' UTR of -32 bp upstream of the translation start site and a promoter (\( \text{PsxtD} \)) with high sequence similarity to the \( \text{E. coli} \) \( \sigma_70 \) -10 and -35 hexamers. \( \text{PsxtP} \) initiates the transcription of operon 2, possesses a short 5' UTR spanning 34 bp and contains both -10 and -35 regions. The transcript initiated by \( \text{PsxtP} \) also displayed a likely RBS (AAGA) 6 nucleotides upstream of the \( \text{sxtP} \) translation start site. A conserved -35 sequence was also identified 21 bp upstream of the extended -10 sequence, resulting in an unusually long distance between the two hexamers. \( \text{Orf24} \) has a perfectly conserved -10 consensus sequence, including the extended -10 TGn motif (Table 1). The 5' UTR for \( \text{orf24} \) is 160 bp in length.

Unusually, transcription of the putative transporter, \( \text{sxtPER} \), was initiated from two promoters, \( \text{PsxtPER1} \) and \( \text{PsxtPER2} \). \( \text{PsxtPER1} \) is located 91 bp upstream of the annotated TSS of \( \text{sxtPER} \) (Figure 2) and contains a highly conserved -10 and -35 RNA polymerase binding site. \( \text{PsxtPER2} \) is located 94 bp downstream of the translational start site and contains a highly conserved -10 sequence, including the single nucleotide seen in extended -10 promoters as well as an RBS (AAAGAAG).

The activity of \( \text{sxt} \) promoters in \( \text{E. coli} \)

The five promoters identified in the \( \text{D. circinale} \) \( \text{sxt} \) cluster using 5' RACE, \( \text{PsxtP}, \text{PsxtD}, \text{PsxtPER1}, \text{PsxtPER2} \) and \( \text{Orf24} \), were amplified by PCR and cloned into the \( \text{E. coli} \) expression vector, pET28b (Novagen), directly in front of a luciferase reporter (\( \text{lux} \)) operon (Figure S4). Expression of luciferase from each of these promoters was measured and compared with negative controls; \( \text{pET28-lux} \) harbouring a non-promoter region from within the \( \text{sxtO} \) gene and the \( \text{pET28-lux} \) plasmid with no added promoter. Unpaired t-tests showed that all promoters exhibited significant levels of expression (Table S2) when compared to the \( \text{pET28-lux} \) negative control. Under the described culture conditions, the heterologous
PsxtD, PsxtP, and PsxtPER1 promoters mediated the highest levels of luciferase expression in \textit{E. coli} (Table S3). There was a statistically significant difference ($p<0.0001$) between the highest performing promoter PsxtD and all the other promoters, as well as the controls (sxtO and pET28-lux) (Table S4).

The promoter responsible for the transcription of \textit{orf24} and the second promoter of \textit{sxtPER}, PsxtPER2 were weaker than the other promoters, but still significantly stronger than the controls (Figure 3A). The incorporation of both promoters into the \textit{lux} expression constructs resulted in a 12-27-fold increase in luciferase expression over \textit{sxtO-lux} (Figure 3B), and 810-1770-fold increase in luciferase expression over the pET28-\textit{lux} control. These results indicate that the promoters are active, albeit weaker than the other three promoters.

**The activity of sxt promoters in \textit{Synechocystis} sp. PCC 6803**

Four \textit{sxt} promoters were active in \textit{Synechocystis} sp. PCC 6803 (Figure 4). Unpaired t-tests showed that expression of luciferase from the PsxtP, PsxtPER1, and Porf24 were significantly different to expression in the control strain, while expression from PsxtD was not statistically different to the control strain ($P<0.05$; Table S5).

The sxtD promoter regulates the transcription of operon 1 of the \textit{D. circinale} \textit{sxt} cluster, which carries the core biosynthetic genes, including the polyketide synthase-like enzyme, sxtA. Strains harbouring PsxtD had very low luciferase expression levels that were only 1.3-fold higher than expression levels in the promoter-less control strain (Table S6), and were statistically lower than the other 3 promoters (Table S7). The lack of statistically significant expression from PsxtD indicates the promoter as the only candidate for exchange for heterologous expression of PSTs in \textit{Synechocystis} sp. PCC 6803. PsxtP and Porf24 mediated consistent levels of luciferase expression per OD$_{730}$ throughout the experiment (Figure 4B) PsxtPER1 mediated expression levels that were initially up to three-fold higher than PsxtP, however, the rate of expression decreased over the course of growth.

**Discussion**

Information regarding the transcriptional elements of cyanobacterial secondary metabolite biosynthesis is an essential first step in both understanding how these metabolites may be regulated within the native producers as well as harnessing these genes for future heterologous expression experiments. Here, we targeted the transcriptional units of the \textit{D. circinale} AWQC131C \textit{sxt} BGC and measured promoter activity within two possible future heterologous hosts, \textit{E. coli} and \textit{Synechocystis}.

The entire \textit{D. circinale} AWQC131C \textit{sxt} cluster was found to be transcribed on a total of five transcriptional units from two bidirectional promoter sites. The isolated promoters all contained a -10 Pribnow box, as expected, but sporadically contained other transcriptional elements including the -35 hexamer, UP element and RBS, which is commonly observed in cyanobacteria[30, 49, 50]. A further feature was the identification of a second promoter region initiating transcription of the proposed permease SxtPER. One of these included an intragenic promoter region, which results in the expression of a 258 aa truncated
isoform of SxtPER. While uncommon, the use of a second TSS to produce two protein isoforms has previously been reported [51]. For example, the bacterocin colicin V is secreted by a membrane transporter, where both the full length CvaA and shorter CvaA* proteins, translated from the same cvaA gene, are both required for optimal excretion [52]. Interestingly, BLASTp analysis revealed that a complete RhaT super domain of the drug/metabolite transporter family [53] is present within both isoforms of the protein. The larger encoded protein contains additional sequence at the N-terminus. While our previous proteomics experiments in D. circinale were unable to detect both isoforms of SxtPER protein under standard conditions, further experiments are required to determine if the two isoforms of the proteins are required for excretion of PST in D. circinale.

The expression of PsxtP is an interesting example of the promoter elements required for heterologous expression of cyanobacterial promoters in E. coli. PsxtP does not seem to have a discernible -35 binding region yet does have a RBS and promoted high expression levels in E. coli. Previous studies have shown that while the distance between the -10 and -35 sequences can effect transcription in cyanobacteria, the -35 hexamer is not always required [54-56]. Thus, the competing preferences between the TSS sequence and position, taken together with other elements of the promoter such as -10, and -35 sequences, transcription factors, the sequence length between the -10 and -35 regions, and the RBS, highlights the complexity of transcriptional regulation and shows the importance of experimental validation of promoter activation data to further improve bioinformatics databases.

It was found that expression levels of PsxtPER1 initiated at a high level but decreased over the course of growth. This indicates the majority of toxins could be exported from the cell early in culture and retained in the cell as the culture progresses. This would allow future research to optimise PST extraction at different culture stages, either from the cell free component or the cell mass. Alternatively, since PsxtPER1 is active at the early growth stages within the heterologous host, it could be a target for repression to limit toxin export and therefore retaining toxin within the cell. This will increase the efficiency of toxin isolation from the cell biomass.

It is known that gene expression levels will have significant impact on the amount of PST molecule produced. Here, the promoters that regulate PST biosynthesis mediated lowered expression of luciferase in Synechocystis sp. PCC 6803 compared to E. coli. The significant decrease in luciferase expression by the cyanobacterial promoters in Synechocystis sp. PCC 6803 has previously been observed in studies of the zinc inducible promoter, Psmt, from Synechococcus sp. PCC7002. Psmt mediated higher levels of protein synthesis and therefore higher levels of ethylene production in E. coli, compared to Synechocystis sp. PCC 6803, which only produced residual levels [57]. Conversely, high expression levels impacted lynbyatoxin biosynthesis in E. coli. Heterologous expression of lynbyatoxin (ltxA-D) in E. coli was only successful when the strong T7 phage promoter was replaced with the weaker PtetO promoter [27]. PtetO has since been exploited for the heterologous expression of multiple cyanobacterial BGCs in E. coli [21, 22, 25]. Subsequent heterologous expression of lynbyatoxin in Anabaena using non-native promoters was more successful than the E. coli system while expression from the native promoter did not occur in either case [23]. Together, these results suggest that native promoters are recognised differently in
heterologous hosts and that while successful transcription of cyanobacterial BGCs in heterologous hosts is important, other factors play a role in the efficiency of the host production of secondary metabolites. This study has identified each of the five native sxt promoters and established through the use of the lux reporter, which of those were not recognised in both *E. coli* and *Synechocystis*.

**Conclusion**

PSTs have a range of biomedical applications and thus heterologous expression of the sxt BGC should be explored as a potential tool for the characterisation, manipulation and sustainable production of these compounds. Heterologous expression of cyanobacterial natural product BGCs has had mixed success in the past and further characterisation of cyanobacterial promoters is required for successful expression of complex biosynthetic pathways such as the PST biosynthetic pathway. This study identified five putative sxt promoters in *D. circinale* AWQC131C and tested their activity in *E. coli* and *Synechocystis* sp. PCC 6803. In *E. coli*, P*sxtD*, P*sxtP* and P*sxtPER1* promoted high levels of luciferase expression while P*orf24* was significantly weaker. Further, if *Synechocystis* sp. PCC6803 is to be used as a host, the inactive P*sxtD* should be replaced by host-compatible promoters. Successfully manipulating the sxt BGC within a heterologous host at the transcription level is the first step to efficiently access the PSTs for a range of biotechnological applications.

**Methods**

**Strains and culture conditions**

*D. circinale* AWQC131C was maintained in Jaworski's medium (JM) [58] at 24°C ±1°C and illuminated with 11 μmol m⁻² s⁻¹ of photons on a 12:12 h light/dark cycle. *Synechocystis* sp. PCC 6803 was maintained in BG11 medium supplemented with 100 µg mL⁻¹ spectinomycin when required, at 30°C under constant illumination. Unless otherwise specified, *E. coli* strains (Table 2) were maintained in Luria broth or on agar plates supplemented with 100 µg mL⁻¹ ampicillin or 50 µg mL⁻¹ kanamycin and grown at 37°C.

**Total RNA extraction, cDNA synthesis and transcriptional analysis**

To extract high quality total RNA, cell pellets were snap-frozen in liquid nitrogen and ground into a fine powder with a mortar and pestle prior to extraction with the RNeasy Plant Mini kit (QIAGEN). Residual genomic (g) DNA was removed from total RNA samples using TURBO DNA-free™ DNase as described by the manufacturer (Ambion). Removal of contaminating gDNA was confirmed via PCR with the 27F/809R PCR primer set targeting the cyanobacterial 16S rRNA gene [60]. RNA quality was also checked via formaldehyde gel electrophoresis, while gDNA was checked by agarose gel electrophoresis.

The Superscript® III First Strand synthesis system (Invitrogen) was used to reverse transcribe 1 µg of total RNA primed with an antisense gene-specific primer (GSP). Transcriptional units were determined by PCR amplification in a 20 µL reaction mixture containing 2.5 mM MgCl₂, 1 × PCR buffer (Fisher Biotec, Geneworks), 10 pmol dNTPs (Austral Scientific), 10 pmol of GSP, 1 U of *Taq* polymerase (Fisher Biotec,
Geneworks) and sterile Milli-Q water. Thermal cycling was performed in a Bio-Rad 96-well iCycler (Bio-Rad) and began with an initial denaturation cycle of 95°C for 4 min, followed by 35 cycles of DNA denaturation at 95°C for 20 s and primer annealing at 55°C for 20 s. DNA strand extension was altered to 1 min for every 1 kb of amplified product. A final extension at 72°C for 7 min and a final holding temperature of 4°C completed the thermal cycling. Each reaction contained cDNA as the template and two primers (Table S8), which were designed to target an adjacent gene. Amplification was observed if the two adjacent genes were located on the same mRNA transcript. The positive control for each PCR contained gDNA. Two negative control reactions were performed by adding template from a cDNA synthesis reaction, the first omitting reverse transcriptase and the second reaction omitting a nucleic acid template.

Isolation of *D. circinale* AWQC131C *sxt* biosynthetic gene cluster transcription start sites (TSS) and promoters using 5′ rapid amplification of cDNA ends (5′RACE)

To isolate the promoter of each transcriptional unit, TSSs were localised with the FirstChoice® RLM-RACE kit for 5′ RACE (Ambion) with 10 µg total RNA as starting material. The 5′ RACE adapter was ligated directly onto RNA followed by reverse transcription cDNA synthesis. First round PCR reactions were performed using a 5′ outer adapter primer in conjunction with four reverse GSPs at approximately 50-100 bp intervals (Figure S3, Table S9). Reactions containing amplified products from the first round PCR became the template for second round nested PCR containing a 5′ adapter inner primer in conjunction with the same four reverse primers. Amplicons of interest were analysed on a 2% (w/v) agarose gel and purified using a QIAquick spin gel extraction kit (QIAGEN). Purified PCR products were then cloned into the pGEM-T Easy vector (Promega) and sequenced using an ABI 3730 capillary sequencer at the Ramaciotti Centre for Genomics, UNSW.

Cloning and transformation

The TOPO TA cloning® kit (Invitrogen) and the pGEM®-T Easy Vector kit (Promega) were used for the cloning and transformation of *E. coli* (Table 2). Cloning with the TOPO TA cloning® kit involved setting up a ligation reaction containing 4 µL of PCR product, 1 µL of Invitrogen salt solution (1.2 M NaCl, 0.06 M MgCl₂) and 10 ng pCR®2.1-TOPO® plasmid DNA (Invitrogen). The ligation reaction was incubated for 20 min at room temperature and was then ready for transformation. The pGEM®-T Easy vector ligation reaction contained 1 × rapid ligation buffer (Promega), 50 ng of pGEM®-T Easy vector DNA (Promega), 3 Weiss U of T4 DNA ligase (Promega) and 3 µL of PCR product. The ligation reaction was left to incubate overnight at 4°C and was then ready for transformation. Positive transformants were selected by blue and white colony screening and the presence of a cloned insert was confirmed by colony PCR using either the primer sets M13F and M13R (pCR®2.1-TOPO) or T7F and M13R (pGEM-T Easy). Plasmids containing an insert were then sequenced.

Engineering the *sxt* promoter-luciferase reporter constructs for expression in *E. coli* DH5α
Five promoters (PsxtD, PsxtP, PsxtPER1, Orf24, PsxtPER2) and a non-promoter region within the sxtO open reading frame were cloned into the pET28b expression vector along with the luciferase reporter (luxCDABE) operon from *Photorhabdus luminescens* (Figure S4A). The luciferase operon (luxCDABE) was amplified from the pLuxNSII plasmid [59] via PCR (denaturation at 98°C for 3 min, followed by 30 cycles of denaturation at 98°C for 15 s, annealing step at 60°C for 20s, extension at 72°C for 30s/kb, and a final extension at 72°C for 10 min), and the pET28b backbone was also PCR amplified to remove the T7 promoter region. All primers were designed using the NEBuilder assembly tool (Table S10). Double-stranded PCR fragments were amplified using the KAPA HiFi Hotstart DNA polymerase (KAPA biosystems). The pET28b backbone, sxt promoter (Psxt), and lux operon were assembled using the Gibson assembly master mix (NEB) [61], and incubated at 50°C, for 1 h. The reaction was transformed into chemically competent *E. coli* DH5α and positive colonies selected as above.

**Engineering the sxt promoter-luciferase reporter constructs for expression in *Synechocystis* sp. PCC 6803**

The Psxt-lux integration vector was engineered through classical restriction/ligation cloning using the restriction enzymes *NotI* and *KpnI* (NEB). The Psxt-lux fragments were amplified from the pET28b-P-lux vectors using the lacI-P-lux_NotI_F, and lacI-P-lux_KpnI_R primers (Table S10). Linear DNA fragments were digested, purified and ligated into the pSYN_6.3 vector (Figure S4B) using T4 DNA ligase at 22°C for 1 h, followed by transformation into *E. coli* DH5α and colony screening. Plasmid constructs were confirmed by terminal-end sequencing.

The integration of the Psxt-lux fragments into the *Synechocystis* sp. PCC 6803 genome (Figure S4C) was achieved via the natural competence of the host [62]. *Synechocystis* sp. PCC 6803 was grown at 30°C, with shaking at 100 rpm, under constant light until exponential phase, and used to inoculate a 50 mL BG11 medium to an initial OD$_{730}$ of 0.05. After ~4 days of photoautotrophic growth, the cells reached an OD$_{730}$ of 0.5, and were harvested by centrifugation at 2,750 g for 5 min. Cells were resuspended in 2 mL of fresh BG11 medium, divided into 0.5 mL aliquots (OD$_{730}$ of 2.5), combined with 10 µg of DNA and incubated at 30°C for 6 h. A sterile Immobilon Transfer membrane (Merk Millipore) was placed on each BG11 agar plate, overlaid with 200 µL of the transformation mixture and incubated for 12 h under constant illumination at 30°C. Membranes were transferred to BG11 agar plates containing 25 µg mL$^{-1}$ spectinomycin. The plates were incubated for a further two days at 30°C under constant illumination, before the membrane was transferred to plates containing 50 µg mL$^{-1}$ spectinomycin and incubated for a further 7-10 days until colonies become visible. Recombinant *Synechocystis* sp. PCC 6803 colonies were picked and streaked onto BG11 agar plates supplemented with 100 µg mL$^{-1}$ spectinomycin and subcultured a further three times to achieve integration of the cloned reporter fragment and full chromosomal segregation. Transformants were confirmed using the PhaCaF and PhaCbR PCR primers (Table S10).

**Activity of sxt promoters in *E. coli* DH5α**
Promoter-luciferase reporter constructs were transformed into *E. coli* DH5α and grown on M9 minimal medium supplemented with 50 μg mL⁻¹ kanamycin at 37°C for 24 h. Bioluminescence (RLU) and optical density measurements were measured at one-hour intervals until the OD₆₀₀ reached 0.8. A final measurement was taken at 24 h. The strength of each promoter was measured as the highest bioluminescence, normalised to OD₆₀₀. One-way ANOVA (Graphpad Prism 7) was used to calculate any statistical differences between the promoters. Unpaired t-tests were also used to determine the statistical differences between the strains and the control.

**Activity of sxt promoters in *Synechocystis* sp. PCC 6803**

Promoter-luciferase reporter constructs were transformed into *Synechocystis* sp. PCC 6803 strains, which were inoculated into BG11 medium supplemented with 100 μg mL⁻¹ of spectinomycin and grown at 30°C with shaking under constant illumination. Optical density and relative light units (RLU) were measured every 24 h for 400 h. Promoter strength was measured by determining the highest RLU per OD₇₃₀. One-way ANOVA (Graphpad Prism 7) was used to calculate any statistical differences between the promoters. Unpaired t-tests were also used to determine the statistical differences between the strains and the control.

**Abbreviations**

5' RACE – rapid amplification of cDNA ends

BGC – biosynthetic gene cluster

GSP – gene specific primer

PST – paralytic shellfish toxins

RBS – ribosomal binding site

TSS – transcriptional start site

UP – upstream

UTR – untranslated region

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**
Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information files. Sequences of plasmids used in this study are available at the end of the supplementary information file.

Competing interests

The authors declare that they have no competing interests

Funding

This research was funded by the Australian Research Council, grant number DP0880264. PMD and BA were supported by the Australian Postgraduate award. MM was supported by the Western Sydney University Research Development Strategic Initiative Scheme.

Authors' contributions

PMD, RM, JM, MCM and BAN designed the overall project. PMD performed experimental transcriptional analysis, identification of transcription start site by 5’ RACE and bioinformatic analysis. BA and RM designed, cloned and performed promoter activity assays. PMD and BA wrote the manuscript while all authors edited and approved the manuscript.

Acknowledgements

The authors would like to thank Dr. Adam Jones and Dr. Daniela Ferreira for helpful experimental advice regarding cyanobacterial transcript analysis and identification of the transcription start site. The authors would also like to thank Anna Glöckle and Leanne Pearson for reading and editing the manuscript. The authors would like to thank Professor Carl Johnson for providing the pluxNS II plasmid carrying the luciferase operon, and Professor Dean Price for providing the *Synechocystis* 6803 strain.

References

4. Adams HJ, Blair MR, Jr., Takman BH: The local anesthetic activity of saxitoxin alone and with vasoconstrictor and local anesthetic agents. *Archives Internationales de Pharmacodynamie et de


39. Ginn HP, Pearson LA, Neilan BA: *NtcA from Microcystis aeruginosa PCC 7806 is autoregulatory and binds to the microcystin promoter.* *Applied and Environmental Microbiology* 2010, **76**:4362-4368.


42. Kaebernick M, Dittmann E, Börner T, Neilan BA: *Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide.* *Applied and Environmental Microbiology* 2002, **68**:449-455.


45. D’Agostino PM, Song X, Neilan BA, Moffitt MC: *Proteogenomics of a saxitoxin-producing and nontoxic strain of Anabaena circinalis (cyanobacteria) in response to extracellular NaCl and phosphate depletion.* *Environmental Microbiology* 2016, **18**:461-476.


### Tables

**Table 1: Characteristics of promoter regions in the sxt biosynthetic gene cluster of *D. circinale* AWQC131C.**

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Promoter sequence 5′ à 3′</th>
<th>Position$^\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PsxtD</td>
<td>tgtcTTGTGG....(14bp).....GAgTATACTtgactagtA</td>
<td>-32</td>
</tr>
<tr>
<td>PsxtP</td>
<td>gtatCTATCA....(12bp).....GTgTATACTgtcaagtA</td>
<td>-34</td>
</tr>
<tr>
<td>PsxtPER1</td>
<td>ttccTTGCAA....(15bp).....AGiTACAATtacatgA</td>
<td>-91</td>
</tr>
<tr>
<td>PsxtPER2</td>
<td>tgagATGACA....(21bp).....CGaTATAATtgtgggtG</td>
<td>+94</td>
</tr>
<tr>
<td>Porf24</td>
<td>aaaaTTTCCT....(15bp).....TGcTATAATgaaatcT</td>
<td>-160</td>
</tr>
<tr>
<td><em>E. coli</em> $\sigma^{70}$ consensus</td>
<td>....TTGACA....(14bp).....TGnTATAAT......N</td>
<td></td>
</tr>
</tbody>
</table>

The -10 and -35 hexamers are capitalised and conserved nucleotides are in bold print. N indicates transcriptional start site (TSS). $^\Delta$ Position of promoter relative to TSS.
<table>
<thead>
<tr>
<th>Organism</th>
<th>Purpose</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5α</td>
<td>Expression host of lux reporter assay</td>
<td>Promega</td>
</tr>
<tr>
<td>TOP 10</td>
<td>Transformation and propagation of plasmids</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>D. circinale AWQC131C</td>
<td>Organism harbouring native sxt cluster</td>
<td>[34]</td>
</tr>
<tr>
<td>Synechocystis sp. PCC6803</td>
<td>Expression host of lux reporter assay</td>
<td></td>
</tr>
<tr>
<td>Synechocystis SXL7</td>
<td>(\Delta phbC::PsxtD-luxCDABE) - Spec(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>Synechocystis SXL8</td>
<td>(\Delta phbC::PsxtP-luxCDABE) - Spec(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>Synechocystis SXL9</td>
<td>(\Delta phbC::PsxtPER1-luxCDABE) - Spec(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>Synechocystis SXL10</td>
<td>(\Delta phbC::Porf24-luxCDABE) - Spec(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>Synechocystis SXL11</td>
<td>(\Delta phbC::luxCDABE) - Spec(^R)</td>
<td>This study</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Purpose</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCR™2.1 TOPO</td>
<td>Cloning and sequencing of 5′ RACE TSS products</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pGEM T-Easy</td>
<td>Cloning and sequencing of 5′ RACE TSS products</td>
<td>Promega</td>
</tr>
<tr>
<td>pET28b</td>
<td>Cloning and expression of sxt promoters</td>
<td>Novagen</td>
</tr>
<tr>
<td>pSYN_6.3</td>
<td>Cloning and transformation of sxt promoters in Synechocystis sp. PCC 6803</td>
<td>Unpublished</td>
</tr>
<tr>
<td>pLUX NS II</td>
<td>luxCDABE</td>
<td>[59]</td>
</tr>
<tr>
<td>pSXL1</td>
<td>pET28b::PsxtD-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL2</td>
<td>pET28b::PsxtP-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL3</td>
<td>pET28b::PsxtPER1-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL4</td>
<td>pET28b::Porf24-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL5</td>
<td>pET28b::PsxtPER2-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL6</td>
<td>pET28-sxtO-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pET28-lux</td>
<td>pET28-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL7</td>
<td>pSYN_6.3::PsxtD-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL8</td>
<td>pSYN_6.3::PsxtP-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL9</td>
<td>pSYN_6.3::PsxtPER1-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL10</td>
<td>pSYN_6.3::Porf24-luxCDABE</td>
<td>This study</td>
</tr>
<tr>
<td>pSXL11</td>
<td>pSYN_6.3::luxCDABE</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Figures**
Figure 1

Transcriptional analysis of the D. circinale AWQC131C sxt cluster. Genes required for PST biosynthesis are transcribed by a minimum of four mRNA transcripts from two bi-directional promoter regions encoding five promoters; PsxtD (operon 1), PsxtP (operon 2), PsxtPER1 (TU 3), PsxtPER2 (TU 3) and Porf24 (operon 4). Direction of transcription is indicated by black arrows.
Figure 2

Sequence analysis of the five promoters present within the D. circinale AWQC131C sxt biosynthetic gene cluster. Individual promoters include, PsxtD (green), PsxtP (red), PsxtPER2 (light blue), PsxtPER1 (dark blue) and Porf24 (yellow). Transcriptional elements including promoters, TSS, translation start site, and RBS are shown for the five sxt promoters.
Figure 3

Heterologous expression from of luciferase from sxt promoters in E. coli DH5α. (A) Luciferase expression was normalised to the optical density, OD600. The activity of five promoters was tested: PsxtD, PsxtP, PsxtPER1, Porf24, and PsxtPER2. An intergenic sxtO sequence was used as the non-promoter control. The highest RLU/OD for the E. coli DH5α strains expressing luxCDABE. (*) Denotes statistically significant differences (p<0.0001) between the E. coli DH5α strains and the pET28-lux control in unpaired t-tests. (B) Expression relative to the sxtO-lux control. The mean normalised luciferase expression for each promoter was divided by the mean normalised expression of sxtO. The three strongest promoters (PsxtD, PsxtP, PsxtPER) exhibit between 1000-9500 fold higher levels of luciferase expression over the sxtO control. The Porf24 and PsxtPER2 promoters had only a 12 and 27-fold increase in expression, respectively.
Figure 4

Heterologous expression of luciferase from sxt promoters in Synechocystis sp. PCC 6803. (A) Luciferase expression from the four main sxt promoters in Synechocystis sp. PCC 6803. The highest RLU/OD for the Synechocystis sp.PCC 6803 strains expressing luxCDABE. (*) Denotes statistically significant differences (p<0.05) between the Synechocystis sp. PCC 6803 strains and the Synechocystis sp. PCC 6803-lux control in unpaired t-tests. (B) Normalised luciferase expression over time in Synechocystis PCC6803. RLU/OD730 over 400 h of growth. PsxtD normalised expression was low. PsxtP and Porf24 had consistent levels of luciferase expression. PsxtPER showed high initial expression that declined over growth to expression reach levels similar to PsxtP and Porf24.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 200129SupportinginfofinalPD.pdf