Analysis of promoter regions in the saxitoxin gene cluster identifies potential bottlenecks in heterologous biosynthesis

CURRENT STATUS: ACCEPTED

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
University of Western Sydney

✉ m.moffitt@westernsydney.edu.au Corresponding Author
ORCiD: https://orcid.org/0000-0001-9872-3175

DOI: 10.21203/rs.2.15635/v1

SUBJECT AREAS
  Applied & Industrial Microbiology

KEYWORDS
  Promoter, analysis, saxitoxin gene cluster, biosynthesis
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 sxtPER and P sxtPER2. In E. coli, strong expression of lux from P sxtP, P sxtD and P sxtPER was observed while expression from P orf24 and P sxtPER2 was remarkably weaker. In contrast, heterologous expression in Synechocystis sp. PCC 6803 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 future experiments to clone the complete sxt BGC into either heterologous host should result in the transcription of the complete pathway following engineering of the least active promoters. Therefore, heterologous expression of the sxt BGC in either E. coli or Synechocystis could be a viable option for producing PSTs for industrial or biomedical purposes.

Background

Saxitoxin (STX) is a neurotoxin produced by cyanobacteria and microalgae, a member of the broader group of alkaloids known as the paralytic shellfish toxins (PSTs) (Wiese et al., 2010). When high concentrations of PSTs are consumed by humans, acute poisoning can lead to death due to
respiratory paralysis (Etheridge, 2010). 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 (Adams et al., 1976, Lagos, 2014). 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 (Cestèle & Catterall, 2000). Under controlled administration, PSTs are potent anaesthetics, particularly in combination with other local anaesthetics (Barnet et al., 2004, Rodriguez-Navarro et al., 2007). 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 (Epstein-Barash et al., 2009). It has been postulated that the same delivery will provide effective treatment for local and severe joint pain (Chorny & Levy, 2009). Other PSTs, such as the gonyautoxins, also have clinical potential and have been utilised for the treatment of anal fissures and chronic tension type headache (Garrido et al., 2004, Garrido et al., 2005, Lattes et al., 2009).

Obtaining significant quantities of purified saxitoxins for clinical research or water quality analysis is difficult. Additionally, the chemical synthesis of PSTs is complex and unfeasible due to low yields, hence the most common form of obtaining purified compounds involves extraction and isolation from dinoflagellate blooms, cyanobacterial cultures or contaminated shellfish (Laycock et al., 1994). Although analytical calibration standards are commercially available from the National Research Council Canada (NRC), shortages have hampered implementation of routine monitoring of PSTs in bivalves, forcing a reliance on mouse bioassays. These issues clearly highlight the necessity for a feasible and reliable method of commercial PST production and purification.

E. coli is a suitable host for heterologous expression of cyanobacterial pathways based on its fast growth rate and similar GC content. Initial studies used native promoters to produce the ribosomal peptides patellamide A and C, and the microviridins (Long et al., 2005, Ziemert et al., 2008). Recently, there has been a focus on the heterologous expression of cyanobacterial natural product
pathways in *E. coli* using the tetracycline-inducible Ptet₀ promoter (Liu *et al.*, 2017, D’Agostino & Gulder, 2018, Greunke *et al.*, 2018). The lyngbyatoxin biosynthetic gene cluster (*ltx*) has also been expressed in the cyanobacterium *Anabaena* sp. PCC7120 and *E. coli* GB05-MtaA (Ongley *et al.*, 2013, Videau *et al.*, 2016). While the native *ltx* promoters were successfully activated in *Anabaena* sp. PCC7120 to produce lynbyatoxin A, the native promoters were not activated in *E. coli*. Addition of *Anabaena* sp. PCC7120 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 (Wells *et al.*, 2018). While growth of cyanobacteria is much slower than other common microorganisms, they remain a valuable source of significant natural products, and therefore, greater understanding of cyanobacterial promoters is essential. The disconnect between promoter efficiency in different organisms remains poorly understood and therefore, it is essential that promoters are compared in the host organism and heterologous vector (Huang *et al.*, 2010) for biotechnological applications.

For the most part, cyanobacterial transcription machinery conforms to the system in *E. coli*, with the main difference being the widespread absence of the –35 hexamer in cyanobacteria, replaced by a transcription factor binding site to initiate transcription (Curtis & Martin, 1994). In *E. coli*, σ⁷⁰ is able to recognise the majority of promoters while in cyanobacteria, a range of sigma factors have been identified. (Imamura *et al.*, 2003, Imamura & Asayama, 2009).

The saxitoxin biosynthetic (*sxt*) gene cluster has recently been characterised in six cyanobacterial species from the order Nostocales and one from the order Oscillatoriales (Kellmann *et al.*, 2008, Mihali *et al.*, 2009, Stucken *et al.*, 2010, Mihali *et al.*, 2011, Cullen *et al.*, 2018). Each *sxt* cluster 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 metabolites remains limited to the microcystin (*mcy*) and jamaicamde (*jam*) biosynthetic gene clusters (Kaebernick *et al.*, 2000, Kaebernick *et al.*, 2002, Martin-Luna *et al.*, 2006, Jones *et al.*, 2009,
Ginn et al., 2010, Neilan et al., 2012). However, the regulation of PSTs at the molecular level, including the transcriptional elements of sxt remain largely unknown (D’Agostino et al., 2016, Ongley et al., 2016).

Here, we characterise the transcription of sxt within the cyanobacterium Dolichospermum circinale AWQC131C. Transcriptional units were identified which enabled the experimental isolation of two promoter regions. Activity of the promoters was assayed in the heterologous hosts E. coli and Synechocystis sp. PCC6803 using a lux reporter. Characterisation of these cyanobacterial promoters and determination of their activity in E. coli and Synechocystis provides novel expression strategies for PST biosynthesis with biotechnological, pharmaceutical and analytical applications.

Results And Discussion

Identification of transcription units within the sxt gene cluster

Reverse-transcriptase PCR revealed that the encoded enzymes of PST biosynthesis are transcribed on four operons in D. circinale AWQC131C with initiation from two bidirectional promoters (Figure 1). All four transcripts were constitutively expressed in the late exponential growth phase, since sxt mRNA was always detected (data not shown). Transcribed in the reverse direction, operon 1 (sxtDV*EABC; * indicates disrupted ORF of sxtV) spans 7.3 kb and encodes several proteins predicted to be involved early in PST biosynthesis. Operon 2 (sxtPQR) is 3.5 kb and was transcribed in the forward direction. The catalytic functions of SxtPQR are unknown but are likely to be essential for PST biosynthesis as their presence and organisation is conserved amongst all reported sxt clusters. Operon 3 is monocistronic encoding the predicted permease of the drug/metabolite transporter (DMT) superfamily transporter SxtPER. The largest operon, operon 4, is transcribed in the forward direction and spans 12.8 kb. Operon 4 begins by encoding a protein of unknown function (but conserved in many sxt pathways), Orf24, followed by several proteins involved in PST biosynthesis, resulting in the polycistronic operon 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 (Kingsford et al., 2007, Jacobs et al., 2009). Rho-independent transcription termination
sites were identified in the non-coding region of three out of four mRNA sxt transcripts (Table S1). Rho-dependent or Rho-independent termination sites could not be identified in the sequence of mRNA encoding operon 1.

Transcription start sites and promoter regions of the sxt operons. Once the transcription units of the sxt cluster were confirmed, the TSS of each operon was experimentally identified via 5’ RACE (Table 1; Supporting Information). 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 (Imamura & Asayama, 2009). All promoters identified in this study displayed sequence similarity to the consensus prokaryote -10 hexamer while there was sporadic presence of the -35 hexamer binding site (Table 1). This indicates that the sxt promoters of *D. circinale* AWQC131C are activated by an RNA polymerase core enzyme in conjunction with a group 1 or group 2 sigma factor. For the -10 promoter sites identified, a search was conducted for an extended -10 binding site and the upstream (UP) element. The 5´ untranslated region (UTR) region of each operon was also bioinformatically screened for the presence of ribosomal binding sites (RBS), which are periodically present in cyanobacteria (Mutsuda & Sugiura, 2006). Based on the 5´ RACE and bioinformatics data, the sxt gene cluster consists of a total of five TSSs under standard culture conditions (Figure 2).

Operon 1 begins with the transcription of sxtD and contains a short 5´ UTR of -32 bp in regards to the translation start site and a promoter site showing high similarity to the *E. coli* σ⁷₀ -10 and -35 hexamers. PsxtP initiates the transcription of operon 2 and was observed to possess a short 5´ UTR spanning 34 bp and containing both -10 and -35 regions. The transcript initiated by PsxtP also displayed a likely RBS with the sequence AAGA six nucleotides upstream of the sxtP translation start site. Expression of the putative transporter sxtPER occurred from two points, PsxtPER and PsxtPER2. PsxtPER begins -91bp upstream from the annotated translation start site of sxtPER and contains a highly conserved -10 and -35 RNA polymerase binding site. PsxtPER2 is situated +94 bp after the same translational start site. PsxtPER2 contains a highly conserved -10 site, including a single nucleotide seen in extended -10 promoters in addition to a RBS (AAAGAAG). A conserved -35 site was
identified 21 bp upstream of the extended -10, showing an unusually long distance between the two hexamers. Porf24, which begins transcription by coding for the hypothetical protein Orf24, has a promoter with a perfectly conserved -10 sequence, including the extended -10 TGn motif. The 5’ UTR for orf24 is 160 bp in length.

Heterologous promoter activity in E. coli

The five promoter regions of the *D. circinale* sxt cluster identified using 5’ RACE, PsxtP, PsxtD, PsxtPER, PsxtPER2 and Porf24, consisting of 1000 bp from the start codon and downstream, were amplified and cloned into the *E. coli* vector, pET28b, directly in front of a lux reporter operon replacing the T7 promoter. The plasmids were subsequently transformed into *E. coli* DH5α. Expression of luciferase from each of these promoters was measured and compared with negative controls; pET28-lux harbouring a non-promoter region from within the gene sxtO (pSXL6) and another consisting of the pET28-lux plasmid with no added promoter (Figure 3). Unpaired t-tests showed that all promoters exhibited significant levels of expression (Table S2) when compared to the negative control (pET28-lux). The heterologous PsxtD, PsxtP, and PsxtPER promoters (pSXL1–3) exhibited the highest levels of luciferase expression in *E. coli* (Figure 3A, Table S3). There was a statistically significant difference (*p*<0.0001) between the highest performing promoter PsxtD (pSXL1) and all the other promoters (pSXL2–5), as well as the controls (pSXL6 and pET28-lux) (Table S4).

When the expression of each promoter was compared to the expression of the non-promoter control sxtO (pSXL6), the strongest promoters (pSXL1–3) had a 1,000 to 10,000-fold increase of luciferase expression (Figure 3C). A trend was observed in the TSS nucleotide of each operon, with PsxtD, PsxtP, and, PsxtPER having a TSS of adenosine (Table 1). Nucleotide discrimination has been shown to play a role in transcription, with an ATP > GTP > UTP > CTP initiation sequence preference in both *E. coli* and *Synechocystis* PCC6803 (Walker & Osuna, 2002, Mitschke et al., 2011, Kim et al., 2012).

However, the variable transcription of the lux operon cannot be entirely explained by the TSS sequence. The distance of the TSS from -10 region (Pribnow box) has also been shown to be important in transcriptional regulation, with studies using sequence mutations at several positions downstream resulting in various changes in lacZ expression in *E. coli* (Liu & Turnbough, 1994, Han &
Turnbough, 2014).

The expression of PsxtP (pSXL2) 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 displayed high expression in *E. coli*. It is known that the distance between the –10 and –35 effects transcription in cyanobacteria and the –35 hexamer is not always required (Shibato *et al.*, 2002, Subudhi *et al.*, 2008, Dutheil *et al.*, 2012). 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.

The promoter responsible for the transcription of *orf24* and a second promoter of sxtPER2 (pSXL4 and pSXL5, respectively) expressed lux at levels lower than the other promoters, but still significantly different compared to the controls (Table S4). When compared to sxtO (pSXL6), both promoters showed 12–27-fold increase in expression, as well as 810- and 1770-fold higher expression than the pET28-lux control, respectively (Figure 3B; Table S3). These results indicated that the promoters appear to be active, though at lower levels than the other three promoters (Figure 3B).

Both bidirectional promoters of the *D. circinale* AWQC131C sxt pathway were active and able to drive the expression of the luciferase operon in *E. coli*. However, while the three promoters PsxtD, PsxtP and PsxtPER are expressed in *E. coli*, low levels of expression from Porf24 could be a potential bottleneck in the heterologous expression of operon 4 which encodes the majority of genes required for PST biosynthesis, making it a prime target for promoter switching for optimising STX production.

The presence of two TSSs, PsxtPER and PsxtPER2, located up- and down-stream of the annotated translation start sites is unusual and may result in two forms of SxtPER. While PsxtPER2 was able to induce expression of the luciferase operon, it was at significantly lower levels than PsxtPER. A putative NtcA-binding site was found to overlap the –35 position of the PsxtPER2 promoter, and may have affected *in vivo* activation of the promoter in *E. coli*. Therefore, we cannot conclusively
determine if this is in fact a real promoter and two isoforms of SxtPER are translated. Interestingly, BLASTp analysis revealed that a complete RhaT super domain of the drug/metabolite transporter family is present within the shorter isoform. In the event that sxtPER is translated as two isoforms, it is unclear what functionality this would provide, but may alter substrate specificity and enable the export of different PST analogues, which may additionally be regulated by NtcA binding. Further experimental work in the host, *D. circinale* will be required to determine the role of the gene regulated by PsxtPER2.

**Heterologous promoter activity in *Synechocystis* PCC6803**

The bidirectional promoters were also active in *Synechocystis* PCC6803 (Figure 4). The levels of luciferase expression in *Synechocystis* were 1.3-fold of PsxtD (SXL7), 8.45-fold of PsxtP (SXL8), 14.89-fold of PsxtPER (SXL9), and 4.4-fold of Porf24 (SXL10) higher than the control strain carrying only the lux operon (Figure 4B, Table S6). Unpaired t-tests showed that the PsxtP (SXL8), PsxtPER (SXL9), and Porf24 (SXL10) promoters were significantly different to the control strain (Table S5), while PsxtD (SXL7) expression was not statistically different to the control strain (*P*<0.05). By One-way ANOVA here was no statistical difference between PsxtD (SXL7) and Porf24 (SXL10) (*p* = 0.069) in *Synechocystis* PCC6803. All other promoters had a statistically significant difference from the control (Figure 4A, Table S7).

The promoters that regulate saxitoxin biosynthesis, showed lowered expression levels in *Synechocystis* PCC6803 compared to *E. coli*. The surprising nature of this result has previously been observed by the zinc inducible promoter Psmt from *Synechococcus* sp. PCC7002. Psmt resulted in stronger induction of protein synthesis for ethylene production in *E. coli* compared to the residual levels observed in *Synechocystis* PCC6803 (Guerrero et al., 2012). It is possible that the higher level of transcription inhibits biosynthesis in the heterologous host. For example, heterologous expression of lyngbyatoxin (*ltxA-D*) in *E. coli* was only successful under the weaker *PtetO* promoter compared to a T7 promoter (Ongley et al., 2013). The low level expression by the *PtetO* promoter system has since
been successful for the heterologous expression of multiple cyanobacterial biosynthetic gene clusters in \textit{E. coli} (Liu et al., 2017, D’Agostino & Gulder, 2018, Greunke et al., 2018).

The \textit{sxtD} promoter is responsible for the transcription of operon 1 of the \textit{D. circinale} \textit{sxt} cluster, which carries the core biosynthetic genes including the polyketide-like enzyme \textit{sxtA}. \textit{PsxtD} (SXL7) showed very low expression levels that were only 1.3-fold higher than the control strain. The lack of statistically significant expression of \textit{PsxtD} (SXL7) has adverse implications for heterologous expression of PSTs using \textit{Synechocystis} PCC6803 as a heterologous host. Similar to \textit{Porf24} in \textit{E. coli}, \textit{PsxtD} is a candidate for promoter exchange for optimised heterologous expression.

\textit{PsxtP} (SXL8) and \textit{Porf 24} (SXL10) showed consistent level of luciferase expression per OD$_{730}$ throughout the experiment (Figure 4C). The promoter responsible for permease expression (\textit{PsxtPER}) had initial levels that were up to three-fold higher than \textit{PsxtP} (SXL8) but decreased over the course of growth (Figure 4C). This is consistent with previous studies that indicated both intracellular and extracellular saxitoxin levels in \textit{D. circinalis} AWQC131C were consistent throughout exponential growth phase (Ongley et al., 2016). Higher expression of SxtPER during early stages of growth prevents the internal toxin accumulation leading to the consistent levels of intracellular and extracellular saxitoxin.

**Conclusion**

PSTs have a range of biotechnology applications and manipulation of the gene cluster for expression in a heterologous host will improve availability of compounds for these uses. This study experimentally revealed \textit{D. circinale} AWQC131C \textit{sxt} promoters and tested their activity in \textit{E. coli} and \textit{Synechocystis} PCC6803 hosts. In \textit{E. coli}, the \textit{PsxtD}, \textit{PsxtP}, \textit{PsxtPER} promoters exhibited higher levels of luciferase expression while the \textit{Porf24} promoter expression was at a lower level. Therefore, future design a \textit{D. circinale} AWQC131C \textit{sxt} gene cluster construct to be cloned directly into \textit{E. coli} with the \textit{Porf24} promoter being the target of promoter exchange with an alternative promoter. Similarly, \textit{PsxtD} is the candidate for promoter exchange in \textit{Synechocystis} PCC6803 in future cloning experiments. In \textit{E. coli}, the \textit{stx} promoters appeared to result in a greater transcriptional increase compared with the control, however, further analysis is required to determine if this organism will be the better
heterologous host.

Promoter characterisation and analysis of the *D. circinale* AWQC 131C sxt gene cluster has identified heterologous host dependent bottlenecks in PST biosynthesis. This study demonstrates the potential use of native promoters for the heterologous expression of STX and other pharmaceutically important cyanobacterial metabolites.

**Methods**

- **Strains and culture conditions**
  
  *D. circinale* AWQC131C was maintained in Jaworski’s Medium (JM) (Thompson *et al.*, 1988) at 24°C ±1°C and illuminated with 11 μmol m⁻² s⁻¹ of photons on a 12:12 h light/dark cycle. Unless otherwise specified, *E. coli* strains (Table S8) 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. *Synechocystis* sp. PCC6803 was maintained in BG-11 media supplemented with 100 μg mL⁻¹ spectinomycin when required, at 30°C under constant illumination.

- **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 gDNA was removed from total RNA samples using TURBO DNA-free™ DNase as described by the manufacturer (Ambion). Removal of contaminating gDNA was validated via PCR with the 27F/809R PCR primer set targeting the cyanobacterial 16S rRNA gene (Jungblut *et al.*, 2005). RNA quality was also checked by 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 S9), each designed to target an adjacent gene. Amplification was observed if the two adjacent genes were located on the same mRNA transcript. A 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 D. circinale AWQC131C sxt gene cluster transcription start sites and promoters using 5′RACE

To isolate the promoter of each transcriptional unit, transcription start sites (TSSs) were localised with the FirstChoice® RLM-RACE kit for 5′ rapid amplification of cDNA ends (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 S1, Table S10). Reactions containing amplified products from the first round PCR became the template for second round nested PCR 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 DNA Sequencing Analysis Facility, 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 S8). 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 x 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. Colonies were selected by blue and white screening and the presence of an 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.

Promoter expression in E. coli
Five promoters (PsxtD, PsxtP, PsxtPER, Porf24, PsxtPER2) and a non-promoter region within the sxtO open reading frame were cloned into the pET28b expression vector along with the luciferase reporter operon from *Photobacterium luminescens* (Figure S4A). The luciferase operon (*luxCDABE*) was amplified from the pLuxNSII plasmid (Woelfle et al., 2007) (denaturation at 98°C for 3 min, followed by thirty cycles of denaturation at 98°C for 15 s, annealing step at 60°C for 20 s, extension at 72°C for 30 s/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 S11). Double-stranded PCR fragments were amplified using the KAPA HiFi Hotstart DNA polymerase (KAPA biosystems). The pET28b backbone, sxt promoter (Psxt), and luciferase (*lux*) fragments were assembled using the Gibson assembly master mix (NEB) (Gibson et al., 2009), and incubated at 50°C, for 1 h. The reaction was transformed into chemically competent *E. coli* DH5α and positive colonies selected.

Promoter expression in *Synechocystis PCC6803*
The *Psxt-lux* integration vector was constructed through classical restriction/ligation cloning using the restriction enzymes NotI and KpnI (NEB). The *Psxt-lux* fragments were amplified from the pET28b-*Psxt-lux* vectors using the lacI-P-lux_NotI_F, and lacI-P-lux_KpnI_R primers (Table S11). 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 PCC6803* genome (Figure S4C) was
achieved via the natural competence of *Synechocystis* PCC6803 (Eaton-Rye, 2011). *Synechocystis* PCC6803 was grown at 30°C, shaking at 100 rpm, under constant light until exponential phase, and used to inoculate a 50 mL flask to an initial OD$_{730}$ of 0.05. After ~4 days of photoautotrophic growth, the cells reached an OD$_{730}$ of 0.5, and cells harvested in a by centrifugation at 2750 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, spread with 200 µL of the transformation mixture and incubated for 12 h under light 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* 6803 colonies were picked and streaked on BG-11 agar plates supplemented with 100 µg mL$^{-1}$ spectinomycin for three rounds to achieve full segregation of the integration. Transformants were confirmed using the PhaCaF and PhaCbR PCR primers (Table S10). Psxt activity in E. coli DH5α Luciferase 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. Luminescence and optical density measurements were taken at one-hour intervals until the OD$_{600}$ reached 0.8. A final measurement was taken at 24 h (PsxtD $n = 6$, PsxtP $n = 6$, PsxtPER $n = 6$, Porf24 $n = 3$, PsxtPER2 $n = 6$, sxtO $n = 6$, pET28-lux $n = 9$). The strength of each promoter was measured as the highest rate of luminescence change, normalised to a change in OD$_{600}$. 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. Psxt activity in Synechocystis PCC6803 Transformed *Synechocystis* PCC6803 strains were inoculated at an OD$_{730}$ of 0.05 and grown in BG-11 media with the 100 µg mL$^{-1}$ spectinomycin at 30°C, shaking in constant light. OD$_{730}$ and relative light
units were measured every 24 h for 400 h (PsxtD n = 3, PsxtP n = 3, PsxtPER n = 3, Porf24 n = 3, Syn6803-lux n = 3). Promoter strength was measured by determining the highest luminescence per OD_{730} and statistically assessed for expression in *Synechocystis* PCC6803. 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.

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of data and material**

All data generated or analysed during this study are included in this published article and its supplementary information files.

**Competing interests**

The authors declare that they have no competing interests

**Funding**

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

**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, BA MCM, and BAN 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.
We would also like to thank Dr. Tamsyn Garby, Amanda Ling for help with cloning the sxt promoters. The authors would also like to thank Anna Glöckle 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* PCC6803 strain.

**References**

Adams HJ, Blair MR, Jr. & Takman BH (1976) The local anesthetic activity of saxitoxin alone and with vasoconstrictor and local anesthetic agents. *Arch Int Pharmacodyn Ther* 224: 275–282.

Barnet CS, Tse JY & Kohane DS (2004) Site 1 sodium channel blockers prolong the duration of sciatic nerve blockade from tricyclic antidepressants. *Pain* 110: 432–438.

Cestèle S & Catterall WA (2000) Molecular mechanisms of neurotoxin action on voltage-gated sodium channels. *Biochimie* 82: 883–892.

Chorny M & Levy RJ (2009) Site-specific analgesia with sustained release liposomes. *Proc Natl Acad Sci U S A* 106: 6891–6892.

Cullen A, D’Agostino PM, Mazmouz R, Pickford R, Wood SA & Neilan BA (2018) Insertions within the saxitoxin biosynthetic gene cluster result in differential toxin profiles. *ACS Chem Biol* 13: 3107–3114.

Curtis SE & Martin JA (1994) The transcription apparatus and the regulation of transcription initiation. *The Molecular Biology of Cyanobacteria*, Vol. 1 (Bryant DA, ed.) pp. 613–639. Kluwer Academic Publishers, Dordrecht.

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

D’Agostino PM & Gulder TAM (2018) Direct pathway cloning combined with sequence- and ligation-independent cloning for fast biosynthetic gene cluster refactoring and heterologous expression. *ACS Syn Biol* 7: 1702–1708.

Dutheil J, Saenkham P, Sakr S, Leplat C, Ortega-Ramos M, Bottin H, Cournac L, Cassier-Chauvat C & Chauvat F (2012) The AbrB2 autorepressor, expressed from an atypical promoter, represses the hydrogenase operon to regulate hydrogen production in *Synechocystis* strain PCC6803. *J Bacteriol*
Eaton-Rye JJ (2011) Construction of gene interruptions and gene deletions in the cyanobacterium Synechocystis sp. strain PCC 6803. *Photosynthesis Research Protocols*, (Carpentier R, ed.) p. pp. 295–312. Humana Press, Totowa, NJ.

Epstein-Barash H, Shichor I, Kwon AH, Hall S, Lawlor MW, Langer R & Kohane DS (2009) Prolonged duration local anesthesia with minimal toxicity. *Proc Natl Acad Sci U S A 106*: 7125–7130.

Etheridge SM (2010) Paralytic shellfish poisoning: Seafood safety and human health perspectives. *Toxicon 56*: 108–122.

Garrido R, Lagos N, Lattes K, Azolas CG, Bocic G, Cuneo A, Chiong H, Jensen C, Henriquez Al & Fernandez C (2004) The gonyautoxin 2/3 epimers reduces anal tone when injected in the anal sphincter of healthy adults. *Biol Res 37*: 395–403.

Garrido R, Lagos N, Lattes K, et al. (2005) Gonyautoxin: New treatment for healing acute and chronic anal fissures. *Dis Colon Rectum 48*: 335–343.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA & Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth 6*: 343–345.

Ginn HP, Pearson LA & Neilan BA (2010) NtcA from *Microcystis aeruginosa* PCC 7806 is autoregulatory and binds to the microcystin promoter. *Appl Environ Microbiol 76*: 4362–4368.

Greunke C, Duell ER, D’Agostino PM, Glöckle A, Lamm K & Gulder TAM (2018) Direct pathway cloning (DiPaC) to unlock natural product biosynthetic potential. *Metab Eng 47*: 334–345.

Guerrero F, Carbonell V, Cossu M, Correddu D & Jones PR (2012) Ethylene synthesis and regulated expression of recombinant protein in *Synechocystis* sp. PCC 6803. *PLoS ONE 7*: e50470.

Han X & Turnbough CL (2014) Transcription start site sequence and spacing between the −10 region and the start site affect reiterative transcription-mediated regulation of gene expression in *Escherichia coli*. *J Bacteriol 196*: 2912–2920.

Huang H-H, Camsund D, Lindblad P & Heidorn T (2010) Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res 38*: 2577–2593.
Imamura S & Asayama M (2009) Sigma factors for cyanobacterial transcription. *Gene Regulation and Systems Biology* 3: 65–87.

Imamura S, Yoshihara S, Nakano S, Shiozaki N, Yamada A, Tanaka K, Takahashi H, Asayama M & Shirai M (2003) Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium. *J Mol Biol* 325: 857–872.

Jacobs GH, Chen A, Stevens SG, Stockwell PA, Black MA, Tate WP & Brown CM (2009) Transterm: A database to aid the analysis of regulatory sequences in mRNAs. *Nucleic Acids Res* 37: D72-D76.

Jones A, Gerwick L, Gonzalez D, Dorrestein P & Gerwick W (2009) Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins. *BMC Microbiol* 9: 247.

Jungblut A, Hawes I, Mountfort D, Hitzfeld B, Dietrich DR, Burns BP & Neilan BA (2005) Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ Microbiol* 7: 519–529.

Kaebernick M, Neilan BA, Börner T & Dittmann E (2000) Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Appl Environ Microbiol* 66: 3387–3392.

Kaebernick M, Dittmann E, Börner T & Neilan BA (2002) Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide. *Appl Environ Microbiol* 68: 449–455.

Kellmann R, Mihali TK, Jeon YJ, Pickford R, Pomati F & Neilan BA (2008) Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Appl Environ Microbiol* 74: 4044–4053.

Kim D, Hong JS-J, Qiu Y, Nagarajan H, Seo J-H, Cho B-K, Tsai S-F & Palsson BØ (2012) Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet* 8: e1002867.

Kingsford C, Ayanbule K & Salzberg S (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8: R22.
Lagos N (2014) Clinical applications of paralytic shellfish poisoning toxins. *Toxins and Biologically Active Compounds from Microalgae*, Vol. 2 (Rossini GP, ed.) pp. 309–329. CRC Press.

Lattes K, Venegas P, Lagos N, Lagos M, Pedraza L, Rodriguez-Navarro AJ & Garcia C (2009) Local infiltration of gonyautoxin is safe and effective in treatment of chronic tension-type headache. *Neurol Res* 31: 228–233.

Laycock MV, Thibault P, Ayer SW & Walter JA (1994) Isolation and purification procedures for the preparation of paralytic shellfish poisoning toxin standards. *Nat Toxins* 2: 175–183.

Liu J & Turnbough CL (1994) Effects of transcriptional start site sequence and position on nucleotide-sensitive selection of alternative start sites at the pyrC promoter in *Escherichia coli*. *J Bacteriol* 176: 2938–2945.

Liu T, Mazmouz R, Ongley SE, Chau R, Pickford R, Woodhouse JN & Neilan BA (2017) Directing the heterologous production of specific cyanobacterial toxin variants. *ACS chemical biology* 12: 2021–2029.

Long PF, Dunlap WC, Battershill CN & Jaspars M (2005) Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. *Chembiochem* 6: 1760–1765.

Martin-Luna B, Sevilla E, Hernandez JA, Bes MT, Fillat MF & Peleato ML (2006) Fur from *Microcystis aeruginosa* binds *in vitro* promoter regions of the microcystin biosynthesis gene cluster. *Phytochemistry* 67: 876–881.

Mihali TK, Kellmann R & Neilan BA (2009) Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in *Anabaena circinalis* AWQC131C and *Aphanizomenon* sp. NH–5. *BMC Biochem* 10: 8.

Mihali TK, Carmichael WW & Neilan BA (2011) A putative gene cluster from a *Lyngbya wollei* bloom that encodes paralytic shellfish toxin biosynthesis. *PLoS ONE* 6: e14657.

Mitschke J, Georg J, Scholz I, et al. (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proceedings of the National Academy of Sciences* 108: 2124–2129.

Mutsuda M & Sugiura M (2006) Translation initiation of cyanobacterial rbcS mRNAs requires the 38-
kDa ribosomal protein S1 but not the shine-dalgarano sequence. *J Biol Chem* 281: 38314-38321.

Neilan BA, Pearson LA, Muenchhoff J, Moffitt MC & Dittmann E (2012) Environmental conditions that influence toxin biosynthesis in cyanobacteria. *Environ Microbiol* 15: 1239-1253.

Ongley SE, Pengelly JJL & Neilan BA (2016) Elevated Na+ and pH influence the production and transport of saxitoxin in the cyanobacteria *Anabaena circinalis* AWQC131C and *Cylindrospermopsis raciborskii* T3. *Environmental microbiology* 18: 427-438.

Ongley SE, Bian X, Zhang Y, Chau R, Gerwick WH, Müller R & Neilan BA (2013) High-titer heterologous production in *E. coli* of lyngbyatoxin, a protein kinase C activator from an uncultured marine cyanobacterium. *ACS Chem Biol* 8: 1888-1893.

Rodriguez-Navarro Alberto J, Lagos N, Lagos M, *et al.* (2007) Neosaxitoxin as a local anesthetic preliminary observations from a first human trial. *Anesthesiology* 106: 339-345.

Shibato J, Agrawal G, Kato H, Asayama M & Shirai M (2002) The 5′-upstream *cis*-acting sequences of a cyanobacterial *psbA* gene: Analysis of their roles in basal, light-dependent and circadian transcription. *Mol Genet Genomics* 267: 684-694.

Stucken K, John U, Cembella A, Murillo AA, Soto-Liebe K, Fuentes-Valdés JJ, Friedel M, Plominsky AM, Vásquez M & Glöckner G (2010) The smallest known genomes of multicellular and toxic cyanobacteria: Comparison, minimal gene sets for linked traits and the evolutionary implications. *PloS one* 5: e9235.

Subudhi S, Kurdrid P, Hongsthong A, Sirijuntarut M, Cheevadhanarak S & Tanticharoen M (2008) Isolation and functional characterization of *Spirulina* D6D gene promoter: Role of a putative GntR transcription factor in transcriptional regulation of D6D gene expression. *Biochem Biophys Res Commun* 365: 643-649.

Thompson AS, Rhodes JC & Pettman I (1988) *Culture collection of algae and protozoa: Catalogue of strains.* Natural Environment Research Council. Freshwater Biological Association, Ambleside, United Kingdom.

Videau P, Wells KN, Singh AJ, Gerwick WH & Philmus B (2016) Assessment of *Anabaena* sp. strain PCC 7120 as a heterologous expression host for cyanobacterial natural products: Production of
lyngbyatoxin A. *ACS Syn Biol.*

Walker KA & Osuna R (2002) Factors affecting start site selection at the *Escherichia coli* fis promoter. *Journal of bacteriology* **184**: 4783-4791.

Wells KN, Nelson D, Eiting JE, Videau P & Philmus B (2018) The influence of sigma factors and ribosomal recognition elements on heterologous expression of cyanobacterial gene clusters in *Escherichia coli*. *FEMS Microbiology Letters* **365**.

Wiese M, D’Agostino PM, Mihali TK, Moffitt MC & Neilan BA (2010) Neurotoxic alkaloids: Saxitoxin and its analogs. *Mar Drugs* **8**: 2185-2211.

Woelfle MA, Xu Y, Qin X & Johnson CH (2007) Circadian rhythms of superhelical status of DNA in cyanobacteria. *Proceedings of the National Academy of Sciences* **104**: 18819-18824.

Ziemert N, Ishida K, Liaimer A, Hertweck C & Dittmann E (2008) Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria. *Angew Chem, Int Ed* **47**: 7756-7759.

Tables

**Table 1**: Conserved -10 and -35 hexamers and identification of the *D. circinale* AWQC131C Psxt TSSs

| Promoter      | Promoter sequence 5’ à 3’                                                                 | Po |
|---------------|------------------------------------------------------------------------------------------|----|
|               | .................................................................................................................................. |    |
| PsxtD         | tgtcTTGTTGG....(14bp).....GAgTATACTtgactagtA                                              |    |
| PsxtP         | gtatCTATCA....(12bp).....GTgTATACTagtcaagtA                                               |    |
| PsxtPER       | ttccTTGCAA....(15bp).....AG iTACAATtacatgA                                              |    |
| PsxtPER2      | tgagATGACA....(21bp)..... CGaTATAATttgggtG                                              |    |
| Porf24        | aaaaTTTCT....(15bp).....TGcTATAATgaaatcT                                                  |    |
| *E. coli* σ70 consenus | .................................................................................................................................. |    |
|               | ....TTGACA....(14bp).....TGnTATAAT.......N                                              |    |

N indicates the TSS of the respective operon

Ψ Ref

^ From translation start site

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 directed from two bidirectional promoters.
Figure 2

Sequence analysis of the two bidirectional promoters present within the D. circinale AWQC131C sxt gene cluster. Transcriptional elements including promoters, TSS, translation start site, and RBS are shown for the five sxt promoters including PsxtD (red), PsxtP (green), PsxtPER2 (light blue), PsxtPER (dark blue) and Porf24 (yellow).
Heterologous activation of sxt promoters in E. coli: Luciferase expression in E.coli DH5α. The lux expression was normalised to the optical density, OD600. The expression of five promoters was tested: PsxtD, PsxtP, PsxtPER, Porf24, and PsxtPER2. A random intergenic sxtO sequence, was used as the non-promoter control. (A) 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 an unpaired t-tests. (B) Expression relative to the pET28-lux control. The mean normalized luciferase expression for each promoter was divided by the mean normalized expression of pET28-lux. (C) Relative expression of promoters vs sxtO in E. coli. Expression relative to sxtO driven luminescence. The mean normalised luciferase expression for each promoter was divided by the mean normalised expression of sxtO.
Figure 4

Heterologous expression of luciferase from sxt promoters in Synechocystis sp. PCC 6803: Luciferase expression from the four main sxt promoters in Synechocystis sp. PCC 6803. (A) 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 an unpaired t-test. PsxtP, PsxtPER, and Porf24 were significantly different to the lux control. (B) Expression relative to Synechocystis sp. PCC 6803-lux. The mean normalised luciferase expression from each promoter was divided by the mean normalised expression from the promoter-less control. (C) Normalised luciferase expression over time in Synechocystis PCC6803. 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.

190916 - SupplInfo - BA_PD.docx