

# Genetic and lipidomic analyses suggest that *Nostoc punctiforme*, a plant-symbiotic cyanobacterium, does not produce sphingolipids

Samuel Belton<sup>1,\*</sup>, †, Nadia Lamari<sup>1,2,‡</sup>, Lars S. Jeremiin<sup>1,2,3</sup>, Vicente Mariscal<sup>4</sup>, Enrique Flores<sup>4</sup>, Paul F. McCabe<sup>1,2,5</sup> and Carl K. Y. Ng<sup>1,2,5,\*</sup>

## Abstract

Sphingolipids, a class of amino-alcohol-based lipids, are well characterized in eukaryotes and in some anaerobic bacteria. However, the only sphingolipids so far identified in cyanobacteria are two ceramides (i.e., an acetylsphingomyelin and a cerebroside), both based on unbranched, long-chain base (LCB) sphingolipids in *Scytonema julianum* and *Moorea producens*, respectively. The first step in *de novo* sphingolipid biosynthesis is the condensation of L-serine with palmitoyl-CoA to produce 3-keto-dihydrosphingosine (KDS). This reaction is catalyzed by serine palmitoyltransferase (SPT), which belongs to a small family of pyridoxal phosphate-dependent  $\alpha$ -oxoamine synthase (AOS) enzymes. Based on sequence similarity to molecularly characterized bacterial SPT peptides, we identified a putative SPT (Npun\_R3567) from the model nitrogen-fixing, plant-symbiotic cyanobacterium, *Nostoc punctiforme* strain PCC 73102 (ATCC 29133). Gene expression analysis revealed that *Npun\_R3567* is induced during late-stage diazotrophic growth in *N. punctiforme*. However, *Npun\_R3567* could not produce the SPT reaction product, 3-keto-dihydrosphingosine (KDS), when heterologously expressed in *Escherichia coli*. This agreed with a sphingolipidomic analysis of *N. punctiforme* cells, which revealed that no LCBs or ceramides were present. To gain a better understanding of *Npun\_R3567*, we inferred the phylogenetic position of *Npun\_R3567* relative to other bacterial AOS peptides. Rather than clustering with other bacterial SPTs, *Npun\_R3567* and the other cyanobacterial BioF homologues formed a separate, monophyletic group. Given that *N. punctiforme* does not appear to possess any other gene encoding an AOS enzyme, it is altogether unlikely that *N. punctiforme* is capable of synthesizing sphingolipids. In the context of cross-kingdom symbiosis signalling in which sphingolipids are emerging as important regulators, it appears unlikely that sphingolipids from *N. punctiforme* play a regulatory role during its symbiotic association with plants.

## INTRODUCTION

Sphingolipids are the third most abundant class of lipids to be found in cell membranes after glycerophospholipids and sterols [1]. Sphingolipids have numerous cellular functions, including lipid ordering [2], signal transduction [3], protein sorting [4] and outer membrane formation (in some gram-negative bacteria) [5]. Additionally, important biological processes mediated by sphingolipids include stress tolerance [6, 7], development [8] and, more recently described, cross-species signalling [9]. In their simplest form, sphingolipids are termed 'long-chain bases' (LCBs), which are C14–C20 fatty amino alcohols. Apart from chain length, LCBs can differ by the degree of hydroxylation, desaturation, and the presence or absence of a

Received 09 October 2021; Accepted 23 November 2021; Published 21 January 2022

**Author affiliations:** <sup>1</sup>UCD School of Biology and Environmental Science, University College Dublin, Belfield, Dublin D4, Ireland; <sup>2</sup>UCD Earth Institute, O'Brien Centre for Science, University College Dublin, Belfield, Dublin D4, Ireland; <sup>3</sup>Research School of Biology, Australian National University, Canberra, ACT 2600, Australia; <sup>4</sup>Instituto de Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, cicCartuja, Avda. Américo Vespucio 49, 41092 Seville, Spain; <sup>5</sup>UCD Centre for Plant Science, University College Dublin, Belfield, Dublin D4, Ireland.

\*Correspondence: Carl K. Y. Ng, carl.ng@ucd.ie; Samuel Belton, samuel.belton@ucdconnect.ie

**Keywords:** sphingolipids; long chain base; serine palmitoyltransferase; *Nostoc*; cyanobacteria; symbiosis.

**Abbreviations:** AIC, Akaike's information criterion; AOS,  $\alpha$ -oxoamine synthase; BIC, Bayesian information criterion; BioF, 8-amino-7-oxononanoate synthase; dpi, days post-induction; EIC, extracted ion chromatogram; gDNA, genomic DNA; HemaA, 5-aminolevulinic acid synthase; KBL, 2-amino-3-ketobutyrate coenzyme A ligase; KDS, 3-keto-dihydrosphingosine; LCB, long-chain base; MSA, multiple sequence alignment; PCD, programmed cell death; PLP, pyridoxal phosphate; Q-TOF, quadrupole time-of-flight; SE, sequence evolution; SPT, serine palmitoyltransferase; UHPLC, ultra-HPLC.

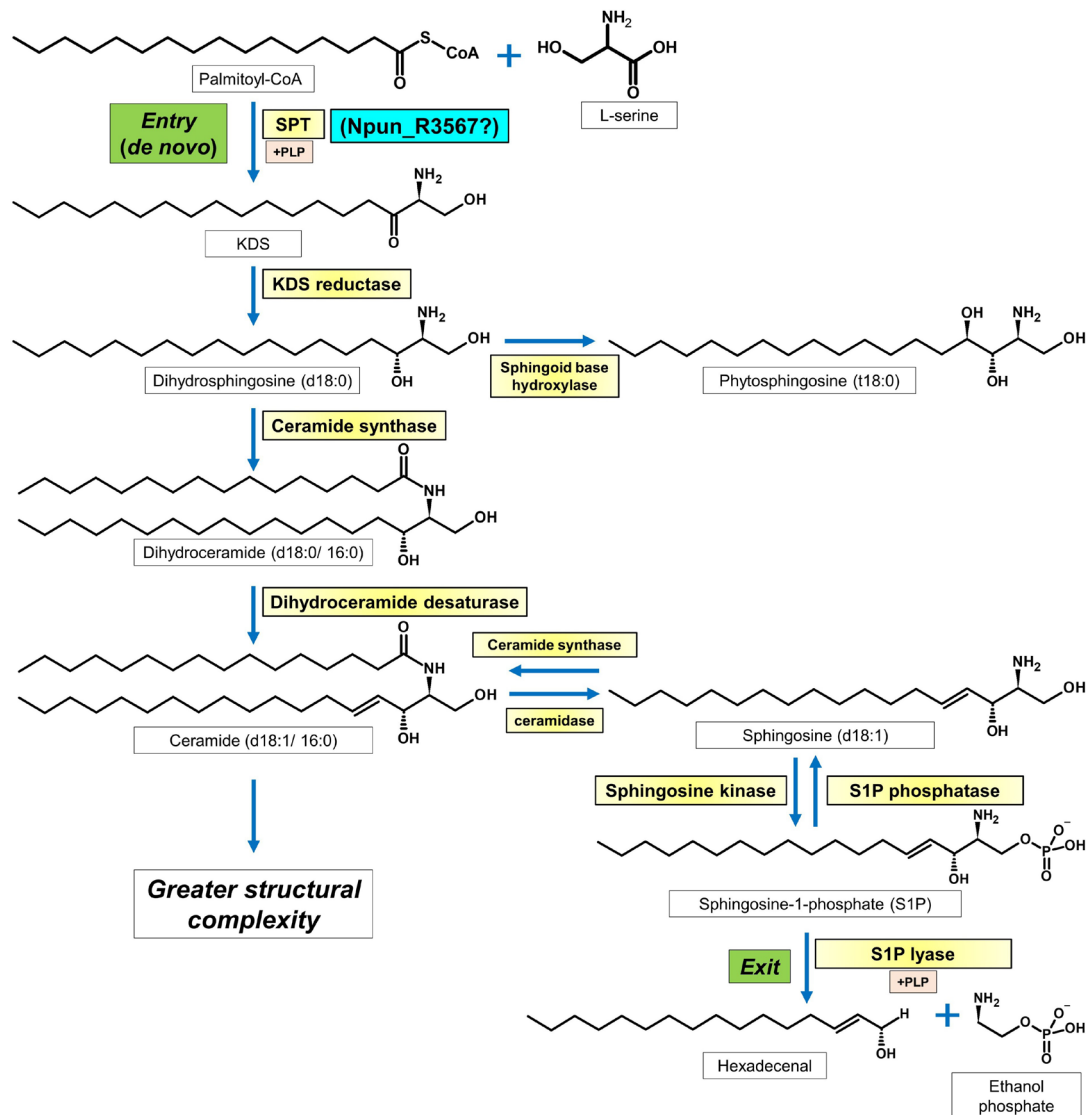
†Present address: DBN Plant Molecular Biology Lab, National Botanic Gardens of Ireland, Dublin, Ireland

‡Present address: Philip Morris International, Quai Jeanrenaud 3, 2000, Neuchâtel, Switzerland.

Supplementary material is available with the online version of this article.

000306 © 2022 The Authors





**Fig. 1.** Canonical sphingolipid metabolic pathway. The *de novo* step (entry point) in the pathway, which is catalyzed by serine palmitoyltransferase, is conserved across all sphingolipid-producing organisms. Ceramide (d18:1/16:0) is a core sphingolipid from which a wide diversity of complex sphingolipids are produced. Alternatively, ceramide may be catabolized into sphingosine (d18:1). Phosphorylation of d18:1 produces the well-characterized signalling sphingolipid, sphingosine-1-phosphate (S1P). The exit point of the pathway is marked by another pyridoxal phosphate-dependent enzyme, S1P lyase.

C1 phosphate and *iso-* or *anteiso-*methyl branches [10]. The most important LCB modification is *N*-acylation with long- or very-long-chain fatty acids to produce ceramides, which are the 'core' sphingolipids from which greater structural complexity is built [11] (Fig. 1). C1 head groups, such as sugars and phosphates, can be further added, producing the more soluble 'complex sphingolipids' [10, 12, 13].

LCB production begins with the decarboxylative condensation reaction between L-serine and palmitoyl-CoA to produce 3-keto-dihydrosphingosine (KDS), which rapidly reduces into dihydrosphingosine (d18:0). The former reaction is catalyzed by serine palmitoyltransferase (SPT), a member of the small, although widespread family of pyridoxal phosphate (PLP)-dependent  $\alpha$ -oxoamine synthase (AOS) enzymes [14–16] (Fig. 1). In eukaryotes, most SPTs are insoluble heterodimers comprising at least two subunits: LCB1 and LCB2 [14, 17]. In contrast, bacterial SPTs are soluble homodimers encoded by the *Spt* gene [18, 19]. This solubility has led to the characterization of several bacterial SPT enzymes [18–20], which are used as models for their eukaryotic counterparts [16].

Despite being ubiquitous in eukaryotes, sphingolipids have only been characterized in a few prokaryotes [12]. The majority of information about sphingolipids in prokaryotes comes from the gram-negative genera *Sphingomonas*, *Sphingobacterium* and *Bacteroides* [12, 21]. Although an acetylsphingomyelin and a cerebroside have been identified in two species of cyanobacteria – *Scytonema julianum* [22] and *Moorea producens* [23], respectively – a sphingolipidome or genetic basis has yet to be reported. To date, no free cyanobacterial LCB (i.e., LCBs not incorporated as part of a ceramide molecule) or sphingolipid enzymes have been characterized.

In eukaryotes, the most common LCBs are dihydrosphingosine (d18:0), sphingosine (d18:1) and phytosphingosine (t18:1), which is a tri-hydroxylated LCB present only in plants and fungi [10, 24] (Fig. 1). Whilst C14, C16, C20 and even odd-chain LCBs have been found [25–27], they appear to be less prevalent. Much less is known about the LCB composition of bacteria. Past research has uncovered both odd- and even-chain LCBs (C16–C19) in prokaryotes, along with unusual branched-chain variants similar to those found in the nematode *Caenorhabditis elegans* [28–31]. The ceramides found in *S. julianum* and *M. producens* are composed of C18 and C17 unbranched LCB backbones, respectively [22, 23].

In this study, efforts were made to characterize sphingolipids in *Nostoc punctiforme* strain PCC 73102 (ATCC 29133), a nitrogen-fixing, plant-symbiotic and cyanobacterial model organism [32–35]. This choice of species was motivated by a rapidly growing body of evidence implicating sphingolipids as important mediators of cross-species signalling in a diverse range of pathogenic and symbiotic relationships [9]. Examples of symbioses in which sphingolipids appear important include those formed between gut *Bacteroides* species and the mammalian gut [6, 36], dinoflagellates and cnidarians [7], Rhizobia and legumes [37], and *Flectobacillus major* (*Bacteroidetes*) and rainbow trout [38]. *Nostoc* species such as *N. punctiforme* can form nitrogen-fixing symbioses with a wide diversity of non-vascular and vascular plant species [34]. *N. punctiforme* strain PCC 73102 was originally isolated from the coralloid roots of a cycad belonging to the genus *Macrozamia* [39], and notably is capable of intracellularly colonizing *Gunnera* species and *Oryza sativa* [40–42]. Moreover, we recently demonstrated that this strain could protect against programmed cell death (PCD) in *Arabidopsis thaliana*, an effect that appeared to be dependent on secreted lipophilic and/or amphiphilic compounds [43]. This suggests that *N. punctiforme* lipids may play a role in regulating PCD, a trait that is hypothesized to be important in symbiosis establishment [7, 44, 45]. If cyanobacteria-derived sphingolipids are important during symbiosis establishment and maintenance, it could be that *N. punctiforme* is capable of sphingolipid production.

Here, we used a sphingolipidomic approach and found that *N. punctiforme* cells do not possess canonical sphingolipids (i.e., unbranched LCBs and ceramides). Keyword and BLAST sequence analysis of the well-annotated *N. punctiforme* genome [35, 46] using previously characterized bacterial AOS peptide sequences resulted in the identification of only one AOS-family protein (Npun\_R3567), which was annotated as an 8-amino-7-oxononanoate synthase (BioF). No phylogenetic evidence supporting an SPT identity for Npun\_R3567 could be found, and the protein was unable to produce KDS when heterologously expressed in *Escherichia coli*. This work is the first to attempt a genetic and lipidomic analysis of sphingolipid production in a cyanobacterium. The results provided should inform future avenues of experimental investigation to elucidate sphingolipid production in cyanobacteria.

## METHODS

### Strains and growth conditions

*N. punctiforme* strain PCC 73102 (ATCC 29133) cells were maintained under 24 h of continuous cool white fluorescent light (approx.  $50 \mu\text{mol m}^{-2} \text{s}^{-1}$ ) on BG11<sub>0</sub> medium [39] containing ammonium (4 mM NH<sub>4</sub>Cl, 8 mM TES, pH 7.5) and solidified with 1.5% BACTO agarose. For experiments, a volume of cells equivalent to 60  $\mu\text{g}$  of chlorophyll *a* was added to 100 ml of fresh growth medium and left to grow for 21 days under a photoperiod of 16 h light (approx.  $165 \mu\text{mol m}^{-2} \text{s}^{-1}$ ) and 8 h dark with constant shaking at 120 r.p.m. The growth conditions for cells used in sphingolipidomic and gene expression analyses were the same, with the exception that gene expression was analysed at different timepoints post-induction of 21-day-old cells (see below). Liquid cell suspension cultures of *A. thaliana* var. Ler-0 were grown under the same conditions but in Murashige and Skoog (MS) medium containing 3% sucrose, 0.5 mg l<sup>-1</sup> naphthaleneacetic acid (NAA) and 0.05 mg l<sup>-1</sup> kinetin (pH adjusted to 5.8 using KOH). Every 7 days, *A. thaliana* cells were sub-cultured by transferring 10 ml of medium with mature cells into 90 ml of fresh MS growth medium.

### Lipid extraction from *N. punctiforme* cells

Four protocols were tested in this work and are outlined in detail in Data Sheet S1 (available in the online version of this article). ‘Method IV’ developed by Markham *et al.* [47] proved most effective at enriching for sphingolipids (Fig. S1, Data Sheet S1). For *N. punctiforme*, lipids were extracted from 500 mg of snap frozen, 21-day-old cells grown in either BG11<sub>0</sub>(NH<sub>4</sub>) or BG11<sub>0</sub> medium, which is BG11 medium lacking any source of combined nitrogen. As a positive control, lipids were extracted from 120 mg of 7-day-old, lyophilized *A. thaliana* Ler-0 cells. As a negative control, a blank extraction was performed without adding biological material.

All standards were purchased from Avanti Polar Lipids and are listed in Table S2 (Data Sheet S1). As internal standards, d17:0, d17:1 and d17:1-P LCBs were added during each extraction. Where cells were being screened for C17 LCBs, d16:1, d18:1 and d18:1-P were instead added. Standards also allowed for calculation of sphingolipid recovery ( $R$ ), sample matrix effects ( $M$ ) and overall process efficiency ( $P$ ), which are explained in Data Sheet S1. Additional standards used for calibration and standard curve generation included d18:0, d18:0-P, CerP(d18:1/12:0-P), GlcCer(d18:1/12:0), LacCer(d18:1/12:0) and Cer(d18:1/25:0). A list of standard retention times, product ions and limit of detection is given in Table S3 (Data Sheet S1).

### LC-ESI-MS/MS analysis of lipid extracts

Chromatographic analyses were performed using an Agilent 1290 Infinity II Ultra-HPLC (UHPLC) system. Lipid extract analysis was carried out on a reversed-phase C8 column (Kinetex 2.6  $\mu\text{m}$ , 100  $\text{\AA}$ , 100 $\times$ 2.1 mm, Phenomenex; ref. 00A-4497-AN) maintained at 50  $^{\circ}\text{C}$  with a flow rate of 400  $\mu\text{l min}^{-1}$ . The mobile phases, consisting of (A) nanopure water with 5 mM ammonium formate and (B) methanol/acetonitrile/isopropanol (10:40:50, by vol.), 5 mM ammonium formate with 0.1% formic acid, were pumped into the UHPLC system with the following elution programme: 0–2 min, 40% B; 2–22 min, 40–90% B; 22–25 min, 90% B. Subsequently, the B content was rapidly (1 min) decreased to the initial conditions and the column was re-equilibrated for 5 min before the next acquisition (total run time 31 min). Prior to sample injection, raw extracts were re-dissolved in 500  $\mu\text{l}$  of isopropanol/methanol (2:1, v/v) and 500  $\mu\text{l}$  of  $\text{H}_2\text{O}$ /methanol (2:1, v/v).

Mass spectral analyses were performed using an Agilent 6550A iFunnel quadrupole time-of-flight (Q-TOF) mass spectrometer equipped with a Dual Agilent Jet Stream electrospray ionization source (Dual AJS ESI). Mass spectra were acquired in positive ion polarity full scan mode, with a mass-to-charge ratio ( $m/z$ ) detection range of 100–1400 selected. The electrospray source parameters were operated as follows: capillary voltage was 3.5 kV with 0 V of collision energy; and gas temperature was 150  $^{\circ}\text{C}$  with a dry gas flow of 14.01  $\text{min}^{-1}$ . The nebulizer was set at 40 psi, nozzle voltage was powered at 1.5 kV and fragmentor was 365 V. For internal mass calibration during the MS analysis, reference masses 121.0509 (purine [ $\text{C}_5\text{H}_4\text{N}_4 + \text{H}$ ] $^+$ ) and 922.0098 (HP-0921 [ $\text{C}_{18}\text{H}_{18}\text{O}_6\text{N}_3\text{P}_3\text{F}_{24} + \text{H}$ ] $^+$ ) were used.

MS/MS data were generated using a data-dependent strategy, selecting targeted precursor ions listed in Table S4, Data Sheet S1. The mass range detection for collision-induced dissociation (CID) was fixed between 30 and 630  $m/z$  for LCBs, 50 and 850  $m/z$  for ceramides, and 100 and 1000  $m/z$  for neutral glycosphingolipids, in which three microscans at an isolation width of 4 atomic mass units was selected. For each precursor ion, four fixed collision energies were selected (5, 12, 18 and 22 eV).

### Qualitative and quantitative analysis of sphingolipids

UHPLC/Q-TOF-MS/MS data were analysed using the Agilent MassHunter software for qualitative analysis (V. B.07.00). Sphingolipid identification was achieved by comparison with the generated MS-related sphingolipid standard reference information (Table S5, Data Sheet S1) and other MS-related information reported in previous works [48–51]. Putative molecular ions identified within 5 p.p.m. were selected for verification by MS/MS analysis. Sphingolipid concentrations were measured using standard curves generated using sphingolipid standards (Fig. S2, Data Sheet S1). Actual concentrations were estimated after factoring in internal standard recovery levels. However, as cautioned previously [47], absolute concentrations cannot be accurately determined in this way as recovery levels are not equal across all classes of compounds unless standards for each compound analyte are added prior to extraction, which often is not practical. Therefore, concentrations were relative to the mean recovery levels of the three C17 internal standards.

### Differential gene expression analysis

To induce diazotrophy, cells from a 100 ml, 21-day-old (stationary growth; Fig. S7, Data Sheet S2) liquid *N. punctiforme* culture were washed with BG11<sub>0</sub> medium (-N) to remove ammonium. Then, the culture was split into two 50 ml volumes. To one, fresh ammonium (as above) was added back [BG11<sub>0</sub>(NH<sub>4</sub>)], whereas nothing was added to the other (BG11<sub>0</sub>). Cultures were then transferred to standard liquid culture growth conditions. See Fig. S8 (Data Sheet S2) for an induction workflow.

RNA was isolated using a method developed by Pinto *et al.* [52] and cleaned using an RNeasy plant mini kit (Qiagen). Total RNA (500 ng) was treated with DNase I (Thermo Scientific), and cDNA was synthesized using SuperScript II reverse transcriptase (Invitrogen). For real-time quantitative PCR (RT-qPCR), *Npun\_F5020* and *Npun\_F5466* (unchanged expression between treatments) were used as housekeeping genes (homologues from *Anabaena* sp. strain PCC 7120 [53], genes *all5167* and *alr0599*, respectively). As a diazotrophy marker, *Npun\_F1722* (*hetR*) was used [54]. As a marker of hormogonia, *Npun\_F2507* (*pilT*) was used [32]. RT-qPCR was performed using FAST SYBR Green master mix (Applied Biosystems) on a ViiA RT-qPCR instrument (Applied Biosystems). All reactions were performed in triplicate. Relative changes in expression were calculated according to Pfaffl [55]. The primer sequences used in this section are detailed in Table S6 (Data Sheet S2).

### Cloning and heterologous expression of *Npun\_R3567* and *SmSpt*

The ORF *Npun\_R3567* was PCR-amplified from genomic DNA (gDNA) and cloned into pET21a+, which is part of a bacterial expression plasmid series [pET21a-d(+)] used before to clone bacterial SPTs [18, 56]. This was done as follows: gDNA was



isolated according to Tamagnini *et al.* [57]. Gene-length primers were designed to PCR-introduce *Bam*HI (5') and *Xho*I (3') restriction sites for insertion upstream of a polyhistidine tag (removing TAA stop codon). The *Spt* gene from *Sphingobacterium multivorum* (*SmSpt*) was restriction-cut from a custom-ordered plasmid (pEX-K4; Eurofins Genomics) harbouring the synthesized gene (*AB259214.1*). Ligation reactions were transformed into *E. coli* strain DH5 $\alpha$  before positive selection on ampicillin-containing growth medium (50  $\mu$ g ml<sup>-1</sup>). The resulting expression plasmids were pUDSB1 (pET21a+w/*Npun\_R3567*) and pUDSB2 (pET21a+w/*SmSpt*), which were verified by sequencing (Eurofins Genomics). The primer sequences and plasmids used in this section are listed in Tables S8 and S9 (Data Sheet S3), respectively.

For heterologous expression, plasmids were transformed into the *E. coli* lysogen BLD21(DE3) and induced by adding 0.4 mM IPTG to liquid cultures with an OD<sub>600</sub> of 0.6 and subsequent incubation for 18 h. A workflow depicting the heterologous expression approach is shown in Fig. S11, Data Sheet S3. Protein synthesis was checked by immunoblotting for the presence of the polyhistidine tag in the soluble protein fraction. To do this, 2 ml of induced cells was centrifuged, flash frozen and resuspended in 600  $\mu$ l of ice-cold potassium phosphate buffer, as done by Ikushiro *et al.* [56], except that no AEBSF protein inhibitor was used. Cells were lysed by shaking cells for 5 min at 30 Hz in the presence of 300 mg of glass beads (0.1 mm  $\varnothing$ ), after which samples were centrifuged at 20000 g for 30 min at 4°C. Immunoblotting was performed by dry-transferring protein from SDS-PAGE gels onto nitrocellulose membranes (Bio-Rad) and incubating in primary [6 $\times$ -His Epitope Tag mouse monoclonal antibody (ThermoScientific; ref. MA1-21315)] and then secondary [rabbit anti-mouse IgG (H+L) secondary antibody, HRP (ThermoScientific; ref. 61–6520)] antibody. Luminescence of recombinant protein was measured at 428 nm after incubation with enhanced chemiluminescence substrate (Amersham ECL Select).

### Lipid extraction from *E. coli* cells

For *E. coli* cells which were induced for the heterologous expression of *Npun\_R3567* and *SmSpt* [18], lipids were extracted using the same protocol as described above for *N. punctiforme* cells. Unlike *N. punctiforme* cells, however, lipids were extracted from snap frozen (rather than lyophilized) cells.

### Sequence analysis of *Npun\_R3567*

For phylogenetic analysis, a suitable set of 28 homologous, non-cyanobacterial, AOS-family peptides were obtained from Geiger *et al.* [58]. These included a number of enzymatically characterized peptides, which were used as BLAST-search queries to identify 29 putative cyanobacterial AOS-family proteins on GenBank. Finally, an additional four non-cyanobacterial sequences were identified from keyword searches on PubMed, focusing on post-2010 (i.e., post-Geiger *et al.* [58]) enzyme characterization studies. All sequence accession numbers are listed in Table S10 (Data Sheet S4). The sequences varied in length from 313 to 409 residues but features and domains that characterize the sequences as members of the AOS family of enzymes were found in all 60 polypeptides.

A multiple sequence alignment (MSA) of the data was inferred using MAFFT v7.471, using the G-INS-i option, which assumes that all sites can be globally aligned [59, 60]. For functional residue comparisons of *Npun\_R3567* with enzymatically characterized SPTs, a box-shaded sub-MSA was made using Jalview2 [61]. Before phylogenetic analysis, the MSA was examined. First, AliStat v1.12 [62] was used to assess completeness. This revealed that the MSA was sparse (i.e. containing many alignment gaps and/or ambiguous letters). Because missing data in MSAs may interfere with phylogenetic inference, the sites of the MSA were masked. Using a threshold of  $C_c=1.0$ , all sites which contained alignment gaps and/or ambiguous letters from the original MSA with 527 sites were masked. This produced a sub-MSA with 291 sites. Next, the sub-MSA was surveyed for evidence that the data have evolved under stationary and reversible conditions (most phylogenetic methods assume evolution under such conditions and might return biased results if the data actually had evolved under non-stationary and/or non-reversible conditions [63, 64]). In practice, Homo 2.0 (<http://github.com/lsermiin/Homo.v2.0>) [65] was employed to assess whether it was sensible to assume that the sequences evolved under stationary and reversible conditions. This appeared to be the case (i.e., the 5% family-wise error rate of 0.05/1770 tests was less than the smallest probability of getting a result from the matched-pairs test of symmetry by chance, which was 0.0046).

Having surveyed the data as outlined above, the phylogenetic inference was started. Unlike most phylogenetic studies, the task was not split into a step aimed at finding the optimal model of sequence evolution (SE) for the data, and a step aimed at finding the optimal phylogeny, conditional on the optimal model of SE. Instead, the advanced version of ModelFinder [66] was employed, which allowed both tasks to be performed simultaneously while considering a greater-than-normal set of models of rate-heterogeneity across sites. In practice, the '-mrate E,I,G,I+G,R,I+R,H,\*H -mfreq FU,FFO' options were used. To accommodate both sensitivity and specificity during model selection [67], both the Bayesian information criterion (BIC) [68] and Akaike's information criterion (AIC) [69] optimality criteria were used. Under BIC, the optimal model of SE was LG+FO+I+G4 (138 parameters), and under AIC, it was LG+FO\*H5 (684 parameters). Both trees were inferred using IQ-TREE 2 [70] while the consistency of the phylogenetic estimates was inferred using the UFboot method with 10000 replicates [71]. The most likely trees were visualized using FigTree [72].

## RESULTS

### Identification of an AOS-family homologue in *N. punctiforme*

In a step towards identifying a SPT homologue in *N. punctiforme*, eleven enzymatically characterized AOS-family peptides (Table 1), including eight SPTs, were used as BLAST query sequences to identify an *N. punctiforme* homologous peptide (Npun\_R3567) comprising 394 amino acids (Fig. 2). The protein is predicted to function as a BioF, which catalyses the first-step reaction in biotin biosynthesis [73]. However, this is a common annotation for bacterial SPTs [19, 74] and the predicted protein sequence was most similar (38.75% identity) to the SPT from the gram-negative bacillus *Sphingobacterium multivorum* [18] (Table 1). Alignment of these, along with two other molecularly characterized bacterial SPTs, revealed that Npun\_R3567 possesses several conserved residues which, in the other proteins, are associated with the SPT enzymatic mechanism (Fig. 2). Of these, eight are shared by all four proteins, whereas two are shared by two and one other protein, respectively. In the characterized proteins, five of these ten residues are involved in the formation of a Schiff base with the PLP cofactor (internal aldimine), three are involved in the formation of the PLP:L-serine quinonoid intermediate (external aldimine), and two interact with ribose and adenine moieties of acyl-CoA molecules (Fig. 2) [16, 20, 75, 76].

### Sphingolipidomic profiling of *N. punctiforme* cells

As no sphingolipid extraction protocol was available for cyanobacterial cells, we tested four lipid extraction protocols and selected one which is based on a method developed by Markham *et al.* [47] to enrich for both polar and non-polar sphingolipids from plant cells (Fig. S1, Data Sheet S1). This was due to its superior levels of internal LCB standard recovery, which ranged between 10 and 20% (Table S1, Fig. S3, Data Sheet S1). Using a UHPLC/Q-TOF-MS/MS method, internal standards eluted between 4 and 8 min on a C8 reversed-phase column, as shown for the C17 internal standards in Fig. 3. An overall extraction process efficiency of  $10.3 \pm 1.6\%$  and an instrument loss on detection of  $\sim 1 \text{ ng ml}^{-1}$  for commercial C16, C17 and C18 LCBs (Table S3, Data Sheet S1) indicated that we would only have failed to detect LCBs if they naturally occurred at a concentration of less than  $9.7 \text{ ng ml}^{-1}$ , equivalent to an  $\sim 32 \text{ fM}$  concentration of d18:0. As a positive control for natural sphingolipid recovery, undifferentiated *A. thaliana* Ler-0 cells grown in liquid suspension were screened. In these cells, t18:0, which was the most abundant LCB detected, was estimated to occur at an  $\sim 3 \text{ pM}$  concentration (Figs S4 and S6, Data Sheet S1). Indeed, the method allowed for the successful identification of several C18 LCBs and ceramides in *A. thaliana* cells (Table S5, Fig. S5, Data Sheet S1).

To screen for sphingolipids in *N. punctiforme* cells, lipids were extracted from both vegetative and diazotrophic (involving heterocyst differentiation) cultures that were 21 days old, which was approximately the beginning of the stationary growth phase for cells grown with and without a source of combined nitrogen (Fig. S7, Data Sheet S2). However, no sphingolipids could be detected under either growth condition. To control for the possibility that *N. punctiforme* LCBs could be C17-based, extractions were then repeated using C16 and C18 as internal standards, but none were detected. Finally, using the METLIN database integrated into the Agilent MassHunter data analysis software, we used an algorithm to search for more atypical sphingolipid classes. Considering all possible charge carriers (i.e. sodiation and protonation), dimers, in-source fragmentation losses ( $\text{H}_2\text{O}$  and  $\text{H}_3\text{PO}_4$ ) and charge ranges yielded no sphingolipid matches.

**Table 1.** Results of a BLAST protein comparison of Npun\_R3567 with different enzymatically characterized AOS-family enzymes from bacteria

Species	Enzyme	Length (aa)	E-value	Percentage identity
<i>Sphingobacterium multivorum</i>	SPT [18]	399	7.6E-92	38.75%
<i>Porphyromonas gingivalis</i> str. W84	SPT [92]	395	1.9E-79	36.12%
<i>Bacteriovorax stolpii</i>	SPT [18]	420	2.4E-79	35.78%
<i>Bacteroides thetaiotaomicron</i>	SPT [74]	394	2.1E-78	37.40%
<i>Bacteroides fragilis</i> str. S23L17	SPT [6]	394	1.8E-76	33.88%
<i>Escherichia coli</i>	KBL [93]	398	2.9E-67	33.08%
<i>Caulobacter crescentus</i>	SPT [94]	404	5.8E-61	32.38%
<i>Sphingomonas wittichii</i>	SPT [19]	400	1.1E-58	30.65%
<i>Sphingomonas paucimobilis</i>	SPT [56]	420	3.3E-57	31.05%
<i>Rhodobacter capsulatus</i>	HemA [95]	401	2.1E-54	33.16%
<i>Escherichia coli</i>	BioF [73]	384	2.4E-48	32.97%

```

Npun_R3567 1 MNFEF-----GAKRRDPYAWLEQSLAT IHRADWYRSVQP INGR--PGATVVLAGQEV
SmSPT      1 MSKGG-----LGEK I SQFK I VEE----LKAKGLYAYFRP I QSK--QDTEVK I DGRRV
SpSPT      1 MTEAAAQPHALPADAPD I APERDLLSKFDGL I AERQKLLDSGVTDPFA I VMEQVKSPTEAV I RGKDT
SwSPT      1 -----MADLLSKFDPL I AEREALLATGVRDPYAI VMDKVLSPTEAMINGRKT

Npun_R3567 51 INFASNDYLGLAGDKRLMAAATAA I AEFGTSGTGSRL LSGHRE LHRELEKA I ASTKQTEDALVFSSG
SmSPT      47 LMFGSNSYLG LTTDTR I I KAAQDALEKYGTGCAGSRFLNGTLD I HVELEEKLSAYVGKEAA I LFSGT
SpSPT      68 ILLGTYN YMGMTFDPDV I AAGKEALEKFGSGTNGSRMLNGTFHDHMEVEQALRDFYGTGA I VFSGT
SwSPT      48 ILLGTYN YMGMTFDPDV I AAGKQALDEFGSGTTGSRV LNGTYQGHKACEDALKEFYGTEHA I VFSGT

Npun_R3567 118 YLANLGA I TALVGRDL I LSDQYNHSS LKNGA I LSGAAVVEYPHCDVAVLKTQLSQQRQNYRRC I L
SmSPT      114 FQSNLGPLSCLMGRNDY I L LDERDHAS I I DGSRLSFSKVI KYGHNNMEDLRAKLSRLPEDSAK - L I C
SpSPT      135 YMANLGI I STLAGKGEYV I LDADSHAS I YDGCQQGNAE I VRFHNSVEDL DKRLGRLPKPAK - LVV
SwSPT      115 YQANLGM I STLAGKGDY I I LDADSHAS I YDGCWLGDAE I VRFHNSVEDL DKRLGRLPAEAGK - LVV

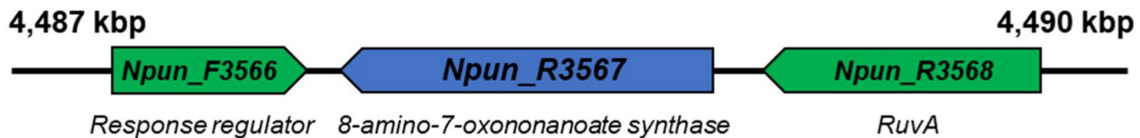
Npun_R3567 185 TDTVFSMDGDLCP LPA LFDLAD EFCMLL VDEAHATGV LKGTGAGCV EYFGCTGKQL I QI GTLSKAL
SmSPT      180 TDGI FMEGD I VNLPEL TS I ANEFDAAVMDDAHS LGVI GHKGAGTASHFGLNDDVDL I MGTFSKSL
SpSPT      201 LEGVYSMLGD I AP LKEMVA VAKKHGAMVL VDEAHSMGF FGPNGRGVYE AQGLEQ I D FVVGTFSKSV
SwSPT      181 LEGVYSMMD I AP LQEMVA VSKKHGAMI LVDEAHGMGF FGEHGRGVFEEAGVEADVDFVVGTFSKSV

Npun_R3567 252 GSLGGYVAGSSAL I DFLRNRAPSWI YTTGLSPADTAAALAAK I VQQEPQH LAQLWRN I HYLKTL LQ
SmSPT      247 ASLGGFVAGDADV I DFLKHNARS VMFASMTFASVASTLKALE I IQNEPEH I EKLWNKTDYAKAQL L
SpSPT      268 GTVGGFVVSNHPKFEAVRLACRPY I FTASLPSPVVATAATS I RKLMTAHEKRERLWSNARALHGGLK
SwSPT      248 GTVGGFCVSNHPKFEVLR LVCRPYVFTASLPSPVVATAATS I RKLMTAHEKRERLWSNARALHGGLK

Npun_R3567 319 QLPNLKLL--PSESP I LCFQLPNATDALKAGKQLRDAG I FAPA I RPP TVPT--SR I R I SVMATHKVA
SmSPT      314 D-HGFDLG--ATESP I LP I F I RSNKETFWTKMLQDDGVFVNPV VSPAVPAEESL I RFSLMATHYD
SpSPT      335 A-MGFR LGTETCDSA I VAVMLDQEQAAMMQAL LDGGL YVNMARPPATPAGTFLLRCS I CAEHTPA
SwSPT      315 D-MGYK LGTETAQSA I I AVI L TDMAQAVLVQGLLEAGL YVNTARPPATPAGMFLLRCS I LCAEHSDE

Npun_R3567 382 HIEKLV-----A--ALKDVI
SmSPT      378 QIDEA I EKMVKVFKQA--EVE TLI
SpSPT      401 QIQTVL----GMFQAAGRAVGVI G
SwSPT      381 QVEQIL----GMFESA GRATGVI P
    
```

Annotation	Residue (NpSPT)	Putative function:
a	Asn56	Binds with L-serine
b	Asp138	Interacts with CoA moieties
c	His142	Anchoring site for L-serine and interacts with CoA moieties
d	Asp215	Facilitates PLP bind to the active site lysine
e	His218	Facilitates PLP bind to the active site lysine
f	Thr246	Facilitates PLP bind to the active site lysine
g	Ser248	Facilitates PLP bind to the active site lysine
h	Lys249	Active site, forms Schiff base with PLP cofactor
i	Arg361	Stabilises PLP:L-serine external aldimine intermediate
j	Arg371	Involved in PLP:L-serine quinonoid intermediate formation

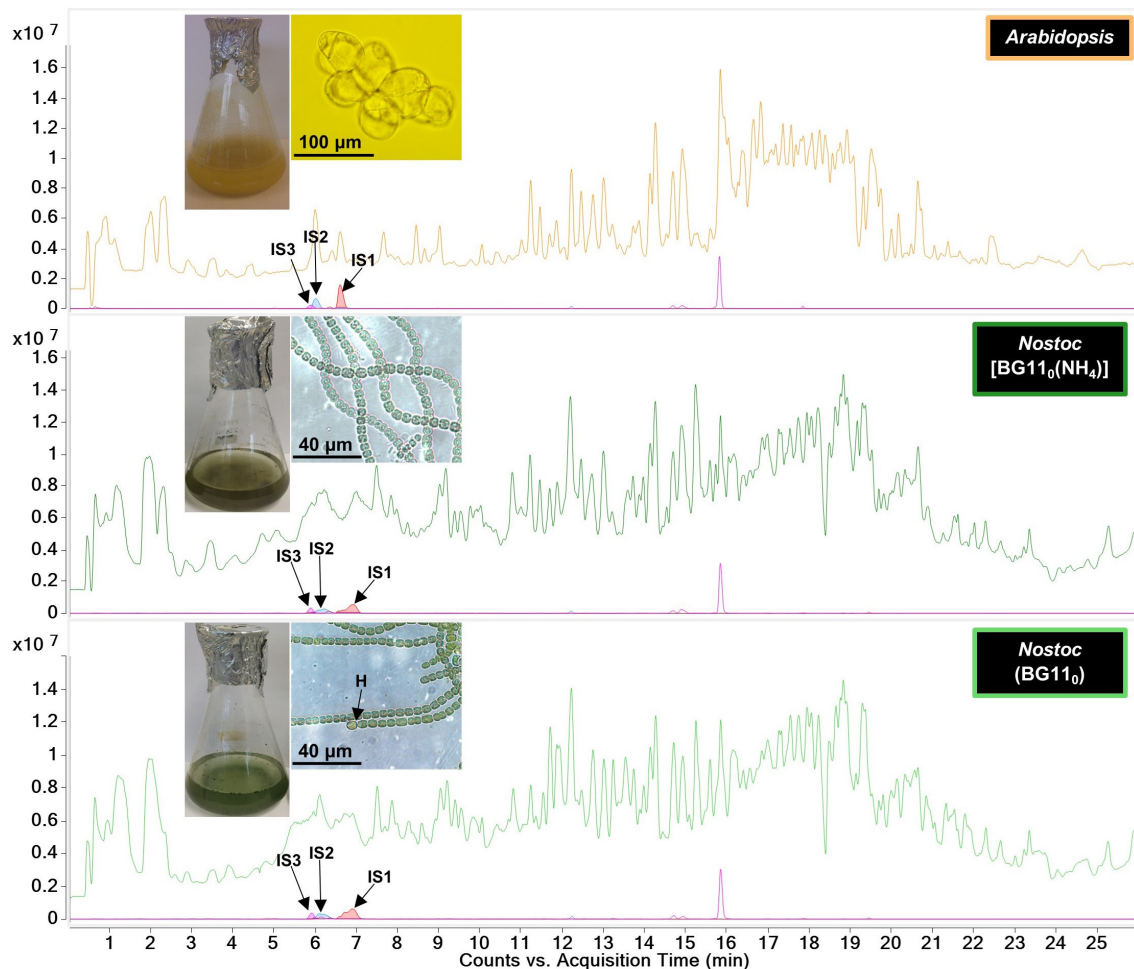


**Fig. 2.** (Top) Alignment of the *N. punctiforme* protein, Npun\_R3567, with the peptide sequences of three characterized SPTs from *Spingobacterium multivorum* (SmSPT), *Spingomonas paucimobilis* (SpSPT) and *Spingomonas wittichii* (SwSPT). Indicated are residues (green) in Npun\_R3567 which correspond to catalytically important residues in the characterized SPTs (see the table). (Bottom) Genomic context of Npun\_R3567, which is annotated as an enzyme which encodes an 8-amino-7-oxonanoate synthase (BioF).

### Gene expression analysis of *Npun\_R3567*

The expression of *Npun\_R3567* was measured at 1, 3, 5, 10 and 20 days post-induction (dpi) for diazotrophy (Fig. 4). In response to nitrogen stepdown, hormogonia cells transiently differentiated (3–5 dpi), followed by heterocyst formation, which were observed at 10 and 20 dpi (Fig. 4, Table S7, Data Sheet S2). These responses were mirrored in an initial, transient up-regulation of *pilT* at 1 dpi, and a subsequent up-regulation of *hetR* at 10 and 20 dpi. No significant changes in the expression of *Npun\_R3567* were observed between 1 and 5 dpi, but at 10 and 20 dpi the gene was significantly up-regulated. The diazotrophy-induced changes in expression of *Npun\_R3567* therefore mirrored that of *hetR* (Fig. 4). These results indicate that *Npun\_R3567* may be important under conditions of diazotrophic growth.





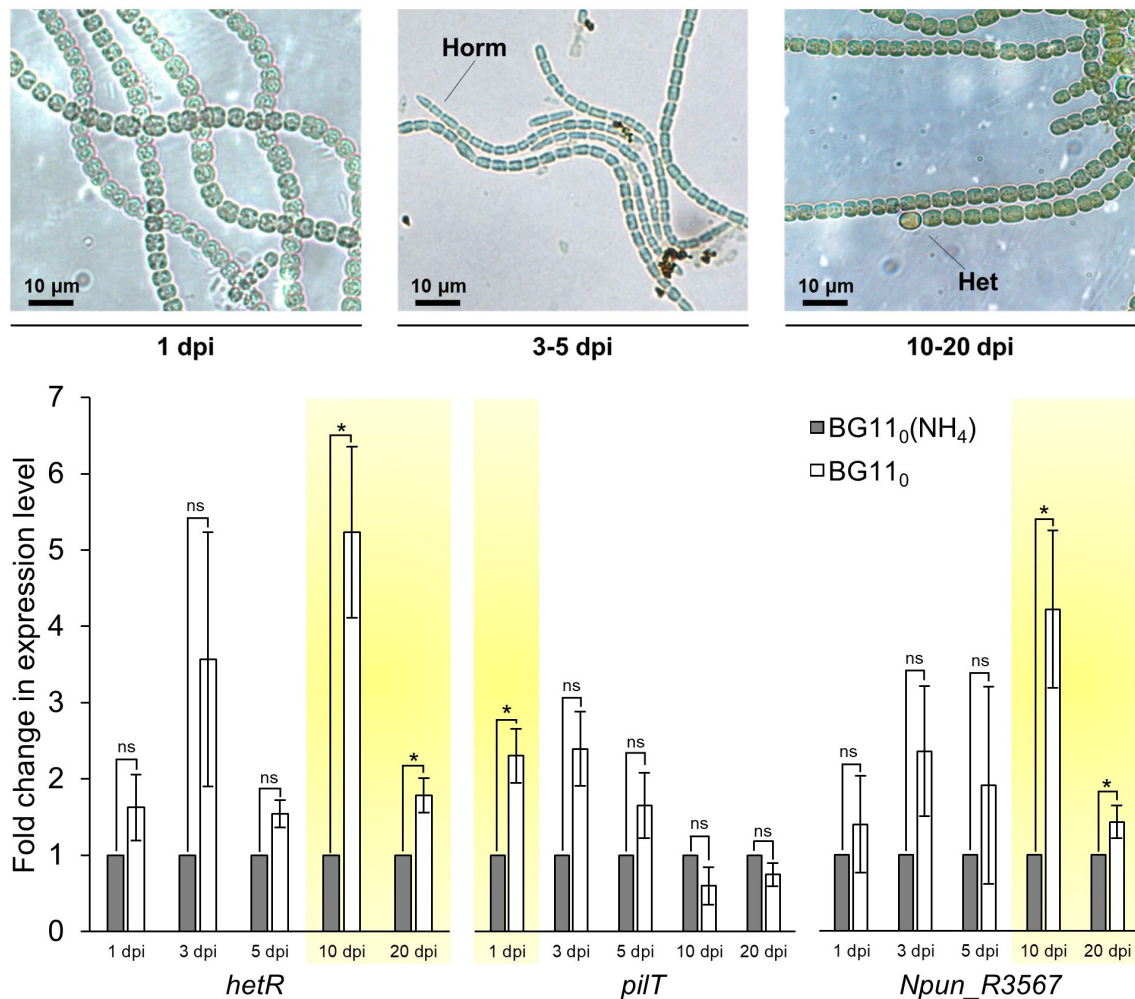
**Fig. 3.** Total ion chromatograms generated by UHPLC/Q-TOF-MS analysis of lipid extracts from *A. thaliana* and *N. punctiforme* cells. The latter were cultured with and without a source of combined nitrogen. Diazotrophic growth is marked by the formation of heterocyst (H) cells. Indicated are the extracted ion chromatograms for the three LCB internal standards (IS), which are d17:0 (IS1), d17:1 (IS2) and d17:1 P (IS3).  $n=4$ .

### **Npun\_R3567 does not produce KDS when heterologously expressed in *E. coli***

Although *Npun\_R3567* would have been expressed at the timepoint of lipid extraction in our sphingolipidomic screen, silencing at the post-transcriptional level could not be ruled out. To get around this, we used the UHPLC/Q-TOF-MS workflow to test if the SPT reaction product, KDS, would be produced in *E. coli* cells heterologously expressing the *Npun\_R3567* protein, as most *E. coli* strains do not possess an *Spt* gene [58]. *SmSpt* was expressed as a positive control and both genes were cloned such that they could be expressed as recombinant proteins bearing respective C-terminal polyhistidine tags. This allowed for their immunoblot detection in the soluble protein fraction of induced cells. Chemiluminescence was similar for both proteins in the presence of IPTG (Fig. 5a). However, signals were also detected in the absence of IPTG, indicating promoter leakiness. In this case, a stronger signal was detected in cells harbouring *Npun\_R3567*. This may be explained by the formation of insoluble inclusion bodies under inducing conditions leading to a weaker signal. This was supported by immunoblotting the total crude cellular fraction, in which case a significantly stronger signal was detected in the induced cells (Fig. S12, Data Sheet S3).

Nonetheless, the level of solubilization achieved under inducing conditions allowed for the detection of a large, split chromatographic peak in the soluble protein fraction of cells expressing *SmSpt* (Fig. 5b). The first shoulder peaked at ~7.8 min, which closely matches the retention time of the commercial d18:0 standard, suggesting that KDS may have been reduced by endogenous *E. coli* reductases. Inspection of the mass spectra for this peak revealed product ions from neutral water and formaldehyde loss, as well as a molecular ion ( $[M+H]^+$ ) corresponding to d18:0 (28 p.p.m. mass error). We infer the second shoulder, which peaked at ~8.1 min, to represent KDS. Although we did not have access to a KDS commercial standard, sphingosine (d18:1), which is isomeric with KDS, eluted at 7.16 min. However, d18:1 is C4-desaturated, which probably reduces affinity with the C8 column compared to the fully saturated KDS. Therefore, we can expect KDS to elute later than the d18:1 standard. Observed product and



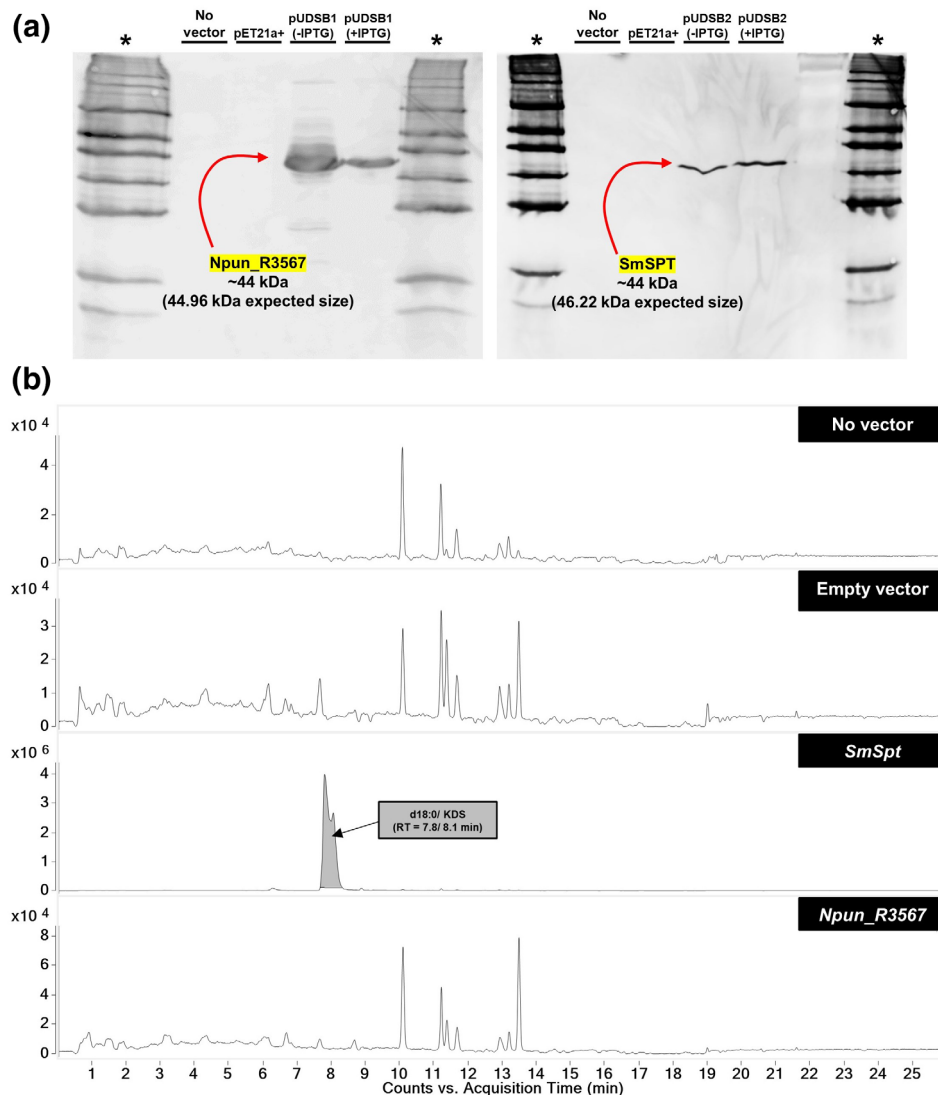


**Fig. 4.** (Top) Representative *N. punctiforme* cells imaged at various days post-induction (dpi) for diazotrophy. At 1 dpi, only vegetative filaments are observable. Between 3 and 5 dpi, hormogonia cells (Horm) are observable. Between 10 and 20 dpi, mature heterocyst (Het) cells are differentiated. (Bottom) Relative change in expression levels of *hetR* (marker for heterocyst formation), *pilT* (marker for hormogonia formation) and *Npun\_R3567* at different days post-induction (dpi). Significant differences (yellow shading) between treatments were determined using a Student's *t*-test (\* $P \leq 0.05$ ). Values are means  $\pm$  SE,  $n=5$ .

molecular (3.7 p.p.m. mass error) ions in the mass spectra of the extracted ion chromatogram (EIC) peak at 8.1 min confirmed that KDS was indeed produced by SmSPT (Fig. 6). We estimated that the KDS/d18:0 mix was produced at an  $\sim 2$  nM level (19.6  $\mu\text{mol g dry weight}^{-1}$ ). By contrast, neither KDS nor d18:0 could be detected in cells expressing *Npun\_R3567*, which suggests that *Npun\_R3567* is not an SPT.

### Phylogenetic position of *Npun\_R3567*

To determine the phylogenetic position of *Npun\_R3567*, we performed a maximum-likelihood-based phylogenetic analysis of the SPT and non-SPT AOS-family enzymes (Fig. 7). Together, this included 60 sequences predicted to function as either SPT, BioF, 2-amino-3-ketobutyrate coenzyme A ligase (KBL) or 5-aminolevulinic synthase (HemA) enzymes, which are the best-characterized members of the AOS enzyme family [12]. Optimal phylogenetic trees were inferred under two models of SE (i.e., identified using the AIC and BIC optimality criteria in order to accommodate sensitivity and specificity during model selection). Both models identified the LG amino acid substitution model [77] and optimized state frequency by maximum likelihood (FO). AIC identified a heterotachous model of rate-heterogeneity across sites [78] with five classes of sites (H5) (Table S11, Fig. S13, Data Sheet S4), whereas BIC assumed a proportion of invariable sites ( $I=4.96\%$ ) and a discrete  $\Gamma$  distribution of variable sites with four equally probable classes (G4) and a shape parameter of 1.2128. The most likely trees inferred under these models of SE differed at 10 internal edges, being present in one tree and absent in the other (Fig. 7). Both models revealed three main peptide groups, which can be seen in the tree shown in Fig. 7 (estimated under the LG+FO\*H5 model of SE). The first group consisted of



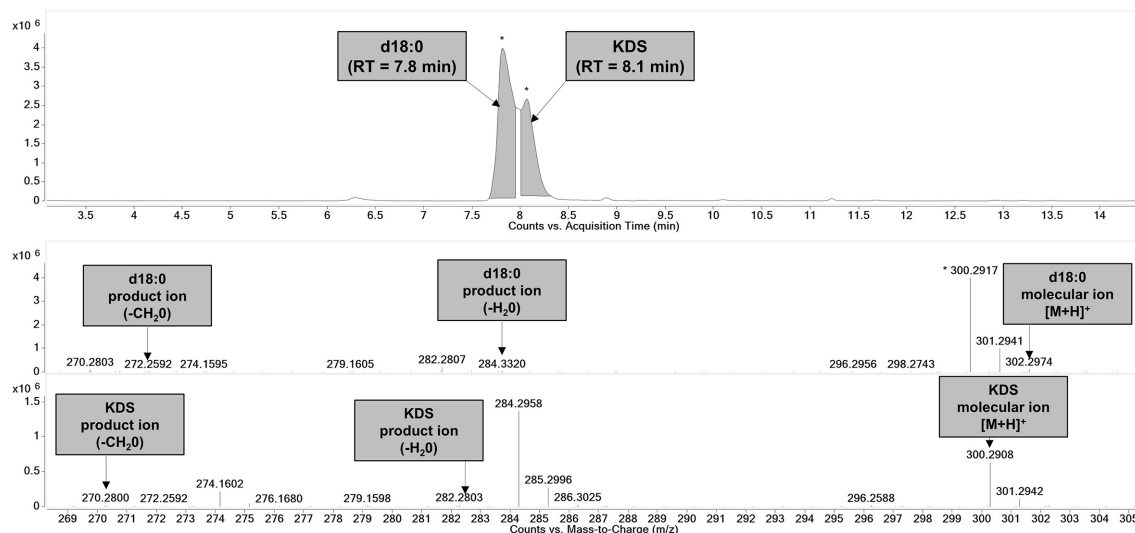
**Fig. 5.** (a) Immunoblot detection of recombinant Npun\_R3567 (left) and SmSPT (right) in the soluble protein fraction of induced *E. coli* BL21(DE3) harbouring pUDSB1 and pUDSB2, respectively. Negative controls include protein from cells harbouring no vector and an empty pET21a+vector, respectively. \*A BenchMark His-tagged protein standard ladder (ThermoScientific) was included to estimate protein size. (b) Extracted ion chromatogram generated by UHPLC/Q-TOF-MS analysis of lipid extracts from *E. coli* cells induced for heterologous expression of SmSpt and Npun\_R3567. No vector and empty vector controls are also shown. A split peak identified at a retention time (RT) of 7.8 and 8.1 min correspond to d18:0 and KDS, respectively.  $n=3$ .

the non-cyanobacterial BioF peptides and the HemaA peptides. This group was related to the second and third groups. Although these latter two were inferred to be a monophyletic group, this was not well supported (i.e., false positive rate for the group is more than 5% [71]). The second group was a well-supported monophyletic group of cyanobacterial BioF peptides, whereas the third group comprised two well-supported groups of KBL and SPT peptides, respectively. These groups were estimated to be monophyletic, but this was also not well supported.

## DISCUSSION

### *N. punctiforme* does not produce unbranched sphingolipids

To the best of our knowledge, the biosynthesis of sphingolipids, which are the third most common class of lipids in eukaryotic cells, has not yet been studied in a cyanobacterium. In this work, we investigated whether *N. punctiforme* strain PCC 73012 (ATCC 29133), a nitrogen-fixing, plant-symbiotic and cyanobacterial model organism, produces sphingolipids. A preliminary BLAST analysis suggested that *N. punctiforme* may possess the first-step enzyme in the *de novo* sphingolipid biosynthesis pathway,

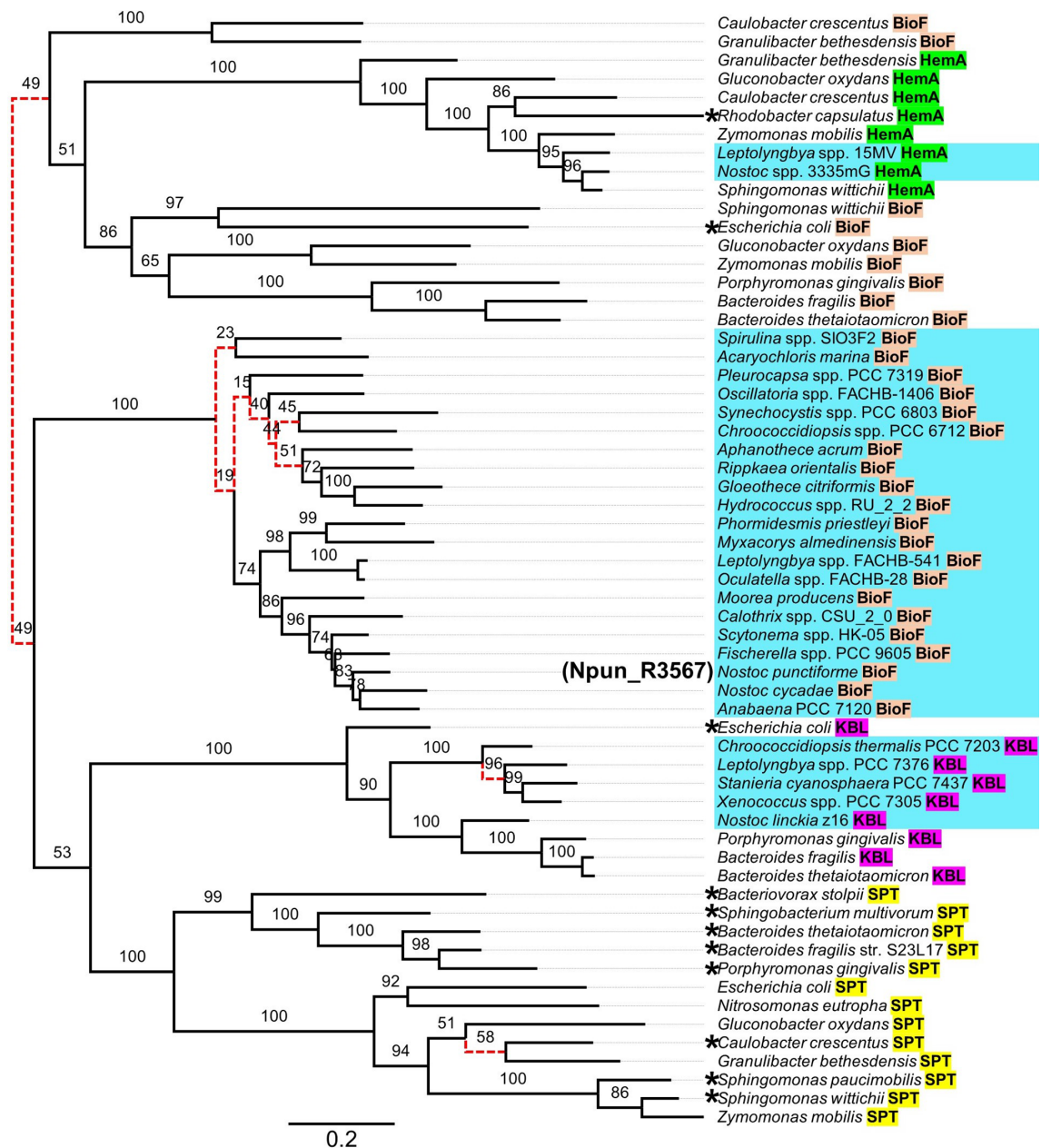


**Fig. 6.** Identification of KDS in lipids extracted from *E. coli* heterologously expressing the previously characterized bacterial SPT from *Sphingobacterium multivorum* (SmSpt). (Top) EIC split peak corresponding to two coeluted compounds, d18:0 and the SPT reaction product, KDS ( $n=3$ ). (Bottom) Extracted mass spectra from each peak show the presence of the diagnostic product and molecular ions derived from in-source fragmentation during full scan analysis.

SPT (Fig. 2). This served as the basis to develop and implement a sphingolipidomic screening method to search for sphingolipids in *N. punctiforme* cells. The method, which consisted of a sphingolipid enrichment protocol developed by Markham *et al.* [47] coupled with a UHPLC/Q-TOF-MS/MS experimental design, enabled the successful identification of LCBs and ceramides in *A. thaliana* but not *N. punctiforme* cells. This is despite the fact that *Npun\_R3567*, which we had hypothesized to encode an SPT, was expressed at the time of lipid extraction (Fig. 4). Moreover, it was also up-regulated in diazotrophic cultures, which agrees with earlier transcriptomic studies in which the gene was shown to be regulated by combined nitrogen depletion [79–82].

It is worth mentioning that our sphingolipidomic analysis focused on the more widespread, unbranched sphingolipids which are based on C14–C20 LCBs. This was based on *a priori* knowledge that unbranched C17 and C18 LCBs form the backbone of ceramides identified in the cyanobacteria, *M. producens* and *S. julianum*, respectively [22, 23]. Methyl-branched LCBs were not manually mined for in the LC/MS data. However, although they appear to be much less common than unbranched LCBs, branched sphingolipids can still be found across kingdoms. For example, C17 methyl-branched LCBs are produced by *Sphingobacterium* species [31] and *Caenorhabditis elegans* [83, 84], whereas a C19 methyl-branched LCB is produced by a marine strain of the fungus *Aspergillus niger* [85]. Iso- and anteiso-branched fatty acids do occur in *Nostoc* and *Anabaena* species [86]. If *N. punctiforme* cells naturally produce branched LCBs, which are often isomeric with unbranched LCBs, then our sphingolipidomic screen would not have detected them because they would have displayed different retention times on the C8 column.

Application of our sphingolipidomic method to *E. coli* cells heterologously expressing the previously characterised *SmSpt* [18] revealed that KDS was produced at a high level (Fig. 7). In contrast, the absence of KDS in *E. coli* cells expressing *Npun\_R3567* suggests that *Npun\_R3567* is not an SPT. This agrees with our sphingolipidomic analysis of *N. punctiforme* cells, which revealed that no sphingolipids were present. However, whilst *Npun\_R3567* is probably not an SPT, it is worth considering two possible reasons which may have precluded it from producing an LCB precursor molecule in *E. coli*. First, unique *N. punctiforme* interacting proteins not present in *E. coli* may be required for any SPT activity by *Npun\_R3567*. Raman *et al.* [19] observed that the SPT from *Sphingomonas wittichii* (SwSPT) co-expressed with an acyl carrier protein (ACP), which they hypothesized may function as an alternative acyl thioester to CoASH. Whilst such proteins may be required by *Npun\_R3567*, the *Npun\_R3567* gene does not appear to be part of an operon containing any gene coding for an ACP. Moreover, of all the bacterial SPT homodimers so far characterized, all can still produce a KDS product in the absence of any co-expressed proteins. This suggests that even in the absence of interacting proteins, if *Npun\_R3567* were a functional SPT, a KDS product should still have been detectable. Second, because we cannot rule out that *N. punctiforme* cells contain the rarer, methyl-branched LCBs, we also cannot rule out that *Npun\_R3567* can only produce an LCB precursor molecule using branched acyl-CoAs. *E. coli* does not produce branched-chain fatty acids [87], indicating that supplementation of cells with branched acyl-CoAs may have been necessary to observe a product. Therefore, we can only report that *N. punctiforme* does



**Fig. 7.** Unrooted phylogeny of prokaryotic AOS-family enzymes inferred using maximum-likelihood under the LG+FO\*H5 model of SE (optimal under AIC). Estimates of consistency (%) for internal edges were inferred using the ultrafast bootstrap method ( $n=10000$ ). Using the LG+FO+I+G4 model of SE (optimal under the AIC), the phylogeny differed from the tree shown here at the 10 red, dashed edges. Cyanobacterial sequences are coloured in cyan. Sequences which have been enzymatically characterized are marked with asterisks. Bar, 0.2 substitutions per site. Branch lengths are weighted average over the different heterotachy classes (Table S11, Data Sheet S4).

not produce the more common, unbranched sphingolipids and that future works focusing on characterizing sphingolipids in *N. punctiforme* and closely related taxa should instead prioritize screening for branched-chain LCBs and ceramides.

### What is the enzymatic identity of Npun\_R3567?

The first-step enzyme in the *de novo* sphingolipid pathway, SPT, has been well-characterized in several species belonging to the phyla *Bacteroidetes* and *Proteobacteria* [18, 19, 56, 74]. The SPT peptide sequences from these species in addition to other bacterial AOS-family peptides allowed us to conduct a homology-based search of GenBank. This led to the identification of what we hypothesized to be an *N. punctiforme* SPT (Npun\_R3567) based on its possession of amino acid residues which in



the query sequences are critical for the SPT reaction mechanism (Fig. 2). For example, His159 in the SPT from *Sphingomonas paucimobilis* (SpSPT) was shown previously to be the anchoring site for the PLP-L-serine external aldimine and a residue which interacts with the carbonyl groups of palmitoyl-CoA and KDS [88]. This residue is probably functionally homologous with His142 in Npun\_R3567. Also present is a residue (Lys249) corresponding to Lys265, which is the active site in SpSPT. However, other AOS-family enzymes (BioF, KBL and HemA) shared a similar level of homology with Npun\_R3567 (Table 1), which was the only identifiable AOS protein in the *N. punctiforme* genome.

The annotation of Npun\_R3567 as a BioF coupled with the observation that BioF homologues could also be BLAST-identified in other cyanobacteria prompted us to include them in a phylogenetic analysis of bacterial AOS-family enzymes. Phylogenies inferred under the optimal AIC- and BIC-predicted models of SE indicate that the cyanobacterial BioF homologues are a well-conserved monophyletic group, even more so than the non-cyanobacterial BioF peptides. The reliability of the phylogenetic estimate is evident in the well-supported HemA, KBL and SPT subgroups, which clustered based on enzyme identity (predicted or characterized) and not based on taxonomic relatedness. By contrast, the non-cyanobacterial BioF peptides were not as closely related, and under BIC, two sequences from *Caulobacter crescentus* and *Granulibacter bethesdensis*, respectively, were instead positioned as part of a monophyletic group with the cyanobacterial sequences. This greater variation in the BioF sequences may be explained by a greater freedom of the BioF peptides to evolve within their respective taxonomic lineages without losing BioF functionality. In this case, stronger taxon-specific signals may impose a masking effect on the functionally important residues. This, in turn, could explain why the cyanobacterial BioF peptides do not cluster with the non-cyanobacterial sequences.

The extent to which this phylogeny is correct depends on how likely it is that we have accurately accounted for the evolution of these sequences. The evolutionary processes inferred to have acted on the sites in the sequences differ in an important and informative manner. The selected model of SE by BIC was LG+FO+I+G4, which was significantly better than the LG+FO model ( $\Delta\text{BIC}=1927.83$ ), indicating that rate-heterogeneity across sites is likely to be a factor that has played a significant role in the evolution of these sequences, as would be expected from a biochemical perspective. However, the LG+FO+I+G4 model assumes homogeneity of site-specific rates of change across lineages, which is unrealistic from a biochemical perspective. To accommodate this, we used the AIC, which is a more sensitive optimality criterion. According to the estimates obtained using AIC, the best model of rate-heterogeneity across sites is heterotachous with five classes of sites, implying a combination of rate-heterogeneity across sites and across lineages. Furthermore, parameters for the five classes of sites are unlinked, implying a large increase in the number of parameters to optimize (including marginal frequencies and edge lengths). Shown in Table S11 (Data Sheet S4) are the marginal frequencies for the five site classes inferred under the LG+FO\*H5 model along with the corresponding frequencies inferred under the LG+FO+I+G4 model. The table reveals large differences across the classes of sites in the relative frequencies of some amino acids (e.g. Gly) and small differences for other amino acids (e.g. His). This is consistent with the notion that the amino acids are distributed non-uniformly across the sites and that the majority (possibly all) of the peptides serve a functional purpose. Fig. S13 (Data Sheet S4) shows the edge lengths for the five site classes inferred under the LG+FO\*H5 model. By comparing the lengths of identical edges from the five subfigures, it is clear the rate of evolution over a given edge varies considerably for many of the edges, so rate-heterogeneity across lineages is likely to be a factor that has played a major role in the evolution of these sequences. This is a far more realistic evolutionary scenario, even from a biochemical perspective.

In support of the large cyanobacterial group being a conserved cluster of BioF enzymes is the recent characterization by Sakaki *et al.* [89] of the second-step enzyme in the *de novo* biotin biosynthesis pathway in the model cyanobacterium, *Synechocystis* sp. strain PCC 6803. This finding gave good reason for the authors to identify the *Synechocystis* homologue for Npun\_R3567 as a *bona fide* BioF, which is the first-step enzyme in the *de novo* pathway [15]. Therefore, based on sequence homology, there is no reason to suspect that the BioF annotation of Npun\_R3567 is incorrect. Given that we could find no other AOS-family enzyme in the *N. punctiforme* genome, it suggests that there is no genetic basis for *de novo* sphingolipid biosynthesis in *N. punctiforme*. Finally, it is worth mentioning that unlike *N. punctiforme* strain PCC 73102, the two cyanobacteria in which ceramide-type sphingolipids have been identified – *M. producens* and *S. julianum* – appear to possess more than one AOS sequence. For example, on CyanoBase [90], BLAST analysis of *M. producens* using SmSPT retrieves a second BioF sequence (LYNGBM3L\_27370) of 571 aa in length. It should be of interest in the future to test whether this and other AOS homologues in cyanobacteria can produce KDS when heterologously expressed in *E. coli*.

## CONCLUSIONS

This work provides evidence which strongly suggests that sphingolipid production does not occur in the model plant-symbiotic cyanobacterium, *N. punctiforme* strain PCC 73102 (ATCC 29133). Our data show that unlike the cyanobacteria *S. julianum* and *M. producens*, *N. punctiforme* does not possess the more common, unbranched sphingolipids. In addition, sphingolipid production in this species seems altogether unlikely given that the only AOS-family protein annotated in the *N. punctiforme* genome, Npun\_R3567, is phylogenetically distinct from bacterial SPTs and cannot produce the SPT reaction product, KDS, when expressed in *E. coli*. Therefore, despite the emerging body of research which is implicating sphingolipids

as key signalling molecules in cross-species communication [7, 9, 91], this work indicates that sphingolipid signalling in *N. punctiforme* is not an aspect of the symbiotic associations which this species forms with multiple plant taxa. However, we also must point out that our approach would not have been able to identify the less common, methyl-branched LCBs and their ceramide derivatives. Therefore, although it seems unlikely that *N. punctiforme* produces sphingolipids, we suggest that future workers instead focus their efforts on these structurally rare sphingolipids in *Nostoc* species. Characterizing the AOS homologues in *S. julianum* and *M. producens* should also be of interest, as these are the only cyanobacteria in which sphingolipids have so far been identified.

#### Funding information

This work was financially supported by a Government of Ireland Postgraduate Scholarship from the Irish Research Council (GOIPG/2015/2695) and a Short-Term Fellowship (ASTF 291-2016) granted by the European Molecular Biology Organisation.

#### Author contributions

S.B.: Conceptualization, Methodology, Formal Analysis, Investigation, Writing – Original Draft, Visualisation, Funding Acquisition. N.L.: Methodology, Formal Analysis, Investigation, Writing - Review & Editing. L.S.J.: Methodology, Software, Formal Analysis, Writing - Review & Editing. V.M.: Methodology, Resources, Writing - Review & Editing, Supervision. E.F.: Validation, Resources, Writing - Review & Editing, Supervision. P.F.M.: Conceptualization, Validation, Resources, Writing - Review & Editing, Funding acquisition. C.K.Y.N.: Conceptualization, Validation, Resources, Writing - Review & Editing, Project administration, Funding acquisition. P.F.M. and C.K.Y.N. conceived of the study and designed experiments. S.B. prepared the draft guided by P.F.M. and C.K.Y.N. N.L. designed, developed and performed the lipidomic experimentation workflow and analysis. L.S.J. performed the phylogenetic analysis. E.F. and V.M. aided in mutagenic and gene expression analysis. All authors read, edited and approved the manuscript.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- Harayama T, Riezman H. Author Correction: Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* 2019;20:715.
- Bieberich E. Sphingolipids and lipid rafts: Novel concepts and methods of analysis. *Chem Phys Lipids* 2018;216:114–131.
- Raben DM, Wattenberg BW. Signaling at the membrane interface by the DGK/SK enzyme family. *J Lipid Res* 2009;50 Suppl:S35-9.
- Ait Slimane T, Hoekstra D. Sphingolipid trafficking and protein sorting in epithelial cells. *FEBS Lett* 2002;529:54–59.
- Kawasaki S, Moriguchi R, Sekiya K, Nakai T, Ono E, et al. The cell envelope structure of the lipopolysaccharide-lacking gram-negative bacterium *Sphingomonas paucimobilis*. *J Bacteriol* 1994;176:284–290.
- An D, Na C, Bielawski J, Hannun YA, Kasper DL. Membrane sphingolipids as essential molecular signals for *Bacteroides* survival in the intestine. *Proc Natl Acad Sci U S A* 2010;108:4666–4671.
- Kitchen SA, Weis VM. The sphingosine rheostat is involved in the chidarian heat stress response but not necessarily in bleaching. *J Exp Biol* 2017;220:1709–1720.
- Tartaglio V, Rennie EA, Cahoon R, Wang G, Baidoo E, et al. Glycosylation of inositol phosphorylceramide sphingolipids is required for normal growth and reproduction in *Arabidopsis*. *Plant J* 2017;89:278–290.
- Heaver SL, Johnson EL, Ley RE. Sphingolipids in host-microbial interactions. *Curr Opin Microbiol* 2018;43:92–99.
- Pruett ST, Bushnev A, Hagedorn K, Adiga M, Haynes CA, et al. Biodiversity of sphingoid bases (“sphingosines”) and related amino alcohols. *J Lipid Res* 2008;49:1621–1639.
- Hannun YA, Obeid LM. Many ceramides. *J Biol Chem* 2011;286:27855–27862.
- Harrison PJ, Dunn TM, Campopiano DJ. Sphingolipid biosynthesis in man and microbes. *Nat Prod Rep* 2018;35:921–954.
- Merrill AH, Sullards MC, Allegood JC, Kelly S, Wang E. Sphingolipidomics: high-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry. *Methods* 2005;36:207–224.
- Hanada K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochim Biophys Acta* 2003;1632:16–30.
- Kerbarh O, Campopiano DJ, Baxter RL. Mechanism of  $\alpha$ -oxoamine synthases: identification of the intermediate Claisen product in the 8-amino-7-oxononanoate synthase reaction. *Chem Commun* 2006:60–62.
- Raman MCC, Johnson KA, Yard BA, Lowther J, Carter LG, et al. The external aldimine form of serine palmitoyltransferase: structural, kinetic, and spectroscopic analysis of the wild-type enzyme and HSN1 mutant mimics. *J Biol Chem* 2009;284:17328–17339.
- Hanada K, Hara T, Nishijima M. Purification of the serine palmitoyltransferase complex responsible for sphingoid base synthesis by using affinity peptide chromatography techniques. *J Biol Chem* 2000;275:8409–8415.
- Ikushiro H, Islam MM, Tojo H, Hayashi H. Molecular characterization of membrane-associated soluble serine palmitoyltransferases from *Sphingobacterium multivorum* and *Bdellovibrio stolpii*. *J Bacteriol* 2007;189:5749–5761.
- Raman MCC, Johnson KA, Clarke DJ, Naismith JH, Campopiano DJ. The serine palmitoyltransferase from *Sphingomonas wittichii* RW1: an interesting link to an unusual acyl carrier protein. *Biopolymers* 2010;93:811–822.
- Yard BA, Carter LG, Johnson KA, Overton IM, Dorward M, et al. The structure of serine palmitoyltransferase: gateway to sphingolipid biosynthesis. *J Mol Biol* 2007;370:870–886.
- Olsen I, Jantzen E. Sphingolipids in bacteria and fungi. *Anaerobe* 2001;7:103–112.
- Antonopoulou S, Oikonomou A, Karantonis HC, Fragopoulou E, Pantazidou A. Isolation and structural elucidation of biologically active phospholipids from *Scytonema julianum* (cyanobacteria). *Biochem J* 2002;367:287–293.
- Youssef DTA, Ibrahim SRM, Shaala LA, Mohamed GA, Banjar ZM. New Cerebroside and nucleoside derivatives from a red sea strain of the marine cyanobacterium *Moorea producens*. *Molecules* 2016;21:324.
- Pata MO, Hannun YA, Ng CK-Y. Plant sphingolipids: decoding the enigma of the Sphinx. *New Phytol* 2010;185:611–630.
- Li Y, Lou Y, Mu T, Ke A, Ran Z, et al. Sphingolipids in marine microalgae: development and application of a mass spectrometric method for global structural characterization of ceramides and glycosphingolipids in three major phyla. *Anal Chim Acta* 2017;986:82–94.
- Fyrst H, Herr DR, Harris GL, Saba JD. Characterization of free endogenous C14 and C16 sphingoid bases from *Drosophila melanogaster*. *J Lipid Res* 2004;45:54–62.

27. Sonnino S, Chigorno V. Ganglioside molecular species containing C18- and C20-sphingosine in mammalian nervous tissues and neuronal cell cultures. *Biochim Biophys Acta* 2000;1469:63–77.
28. Chitwood DJ, Lusby WR, Thompson MJ, Kochansky JP, Howarth OW. The glycosylceramides of the nematode *Caenorhabditis elegans* contain an unusual, branched-chain sphingoid base. *Lipids* 1995;30:567–573.
29. Naka T, Fujiwara N, Yano I, Maeda S, Doe M, et al. Structural analysis of sphingophospholipids derived from *Sphingobacterium spiritivorum*, the type species of genus *Sphingobacterium*. *Biochim Biophys Acta Mol Cell Biol Lipids* 2003;1635:83–92.
30. White DC, Tucker AN, Sweeley CC. Characterization of the iso-branched sphinganine from the ceramide phospholipids of *Bacteroides melaninogenicus*. *Biochim Biophys Acta* 1969;187:527–532.
31. Yano I, Tomiyasu I, Yabuuchi E. Long chain base composition of strains of three species of *Sphingobacterium* gen. nov. *FEMS Microbiol Lett* 1982;15:303–307.
32. Duggan PS, Gottardello P, Adams DG. Molecular analysis of genes in *Nostoc punctiforme* involved in pilus biogenesis and plant infection. *J Bacteriol* 2007;189:4547–4551.
33. Ekman M, Picossi S, Campbell EL, Meeks JC, Flores E. A *Nostoc punctiforme* sugar transporter necessary to establish a Cyanobacterium-plant symbiosis. *Plant Physiol* 2013;161:1984–1992.
34. Whitton BA. Ecology of Cyanobacteria II. In: *Ecology of Cyanobacteria II: Their Diversity in Space and Time*. Dordrecht: Springer, 2012
35. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, et al. An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynth Res* 2001;70:85–106.
36. Brown EM, Ke X, Hitchcock D, Jeanfavre S, Avila-Pacheco J, et al. *Bacteroides*-derived sphingolipids are critical for maintaining intestinal homeostasis and symbiotic Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and Symbiosis. *Cell Host Microbe* 2019;25:668–680.
37. Perotto S. Differential expression of a glycosyl inositol phospholipid antigen on the peribacteroid membrane during pea nodule development. *MPMI* 1995;8:560.
38. Sepahi A, Cordero H, Goldfine H, Esteban MÁ, Salinas I. Symbiont-derived sphingolipids modulate mucosal homeostasis and B cells in teleost fish. *Sci Rep* 2016;6:39054.
39. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* 1979;111:1–61.
40. Bergman B, Johansson C, Söderbäck E. The *Nostoc-Gunnera* symbiosis. *New Phytol* 1992;122:379–400.
41. Chiu W-L, Peters GA, Levieille G, Still PC, Cousins S, et al. Nitrogen deprivation stimulates symbiotic gland development in *Gunnera manicata*. *Plant Physiol* 2005;139:224–230.
42. Álvarez C, Navarro JA, Molina-Heredia FP, Mariscal V. Endophytic colonization of rice (*Oryza sativa* L.) by the symbiotic strain *Nostoc punctiforme* PCC 73102. *Mol Plant Microbe Interact* 2020;33:1040–1045.
43. Belton S, McCabe PF, Ng CKY. The cyanobacterium, *Nostoc punctiforme* can protect against programmed cell death and induce defence genes in *Arabidopsis thaliana*. *J Plant Interact* 2021;16:64–74.
44. Hernández-López A, Díaz M, Rodríguez-López J, Guillén G, Sánchez F, et al. Uncovering Bax inhibitor-1 dual role in the legume-rhizobia symbiosis in common bean roots. *J Exp Bot* 2019;70:1049–1061.
45. Liang Y, Cao Y, Tanaka K, Thibivilliers S, Wan J, et al. Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. *Science* 2013;341:1384–1387.
46. Moraes LE, Blow MJ, Hawley ER, Piao H, Kuo R, et al. Resequencing and annotation of the *Nostoc punctiforme* ATCC 29133 genome: facilitating biofuel and high-value chemical production. *AMB Express* 2017;7:42.
47. Markham JE, Li J, Cahoon EB, Jaworski JG. Separation and identification of major plant sphingolipid classes from leaves. *J Biol Chem* 2006;281:22684–22694.
48. Ishikawa T, Ito Y, Kawai-Yamada M. Molecular characterization and targeted quantitative profiling of the sphingolipidome in rice. *Plant J* 2016;88:681–693.
49. Nurul Islam M, Chambers JP, Ng C-Y. Lipid profiling of the model temperate grass, *Brachypodium distachyon*. *Metabolomics* 2011;8:598–613.
50. Qu F, Zhang H, Zhang M, Hu P. Sphingolipidomic profiling of rat serum by UPLC-Q-TOF-MS: application to rheumatoid arthritis study. *Molecules* 2018;23:E1324.
51. Wang J-R, Zhang H, Yau LF, Mi J-N, Lee S, et al. Improved sphingolipidomic approach based on ultra-high performance liquid chromatography and multiple mass spectrometries with application to cellular neurotoxicity. *Anal Chem* 2014;86:5688–5696.
52. Pinto FL, Thapper A, Sontheim W, Lindblad P. Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol Biol* 2009;10:79.
53. Ramos-León F, Mariscal V, Frías JE, Flores E, Herrero A. Divisome-dependent subcellular localization of cell-cell joining protein SepJ in the filamentous cyanobacterium *Anabaena*. *Mol Microbiol* 2015;96:566–580.
54. Wong FCY, Meeks JC. The hetF gene product is essential to heterocyst differentiation and affects HetR function in the cyanobacterium *Nostoc punctiforme*. *J Bacteriol* 2001;183:2654–2661.
55. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29:e45.
56. Ikushiro H, Hayashi H, Kagamiyama H. A water-soluble homodimeric serine palmitoyltransferase from *Sphingomonas paucimobilis* EY2395T strain. Purification, characterization, cloning, and overproduction. *J Biol Chem* 2001;276:18249–18256.
57. Tamagnini P, Troshina O, Oxelfelt F, Salema R, Lindblad P. Hydrogenases in *Nostoc* sp. strain PCC 73102, a strain lacking a bidirectional enzyme. *Appl Environ Microbiol* 1997;63:1801–1807.
58. Geiger O, González-Silva N, López-Lara IM, Sohlenkamp C. Amino acid-containing membrane lipids in bacteria. *Prog Lipid Res* 2010;49:46–60.
59. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;30:3059–3066.
60. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
61. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189–1191.
62. Wong TKF, Kalyaanamoorthy S, Meusemann K, Yeates DK, Misof B, et al. A minimum reporting standard for multiple sequence alignments. *NAR Genom Bioinform* 2020;2:lqaa024.
63. Ho SY, Jermini L. Tracing the decay of the historical signal in biological sequence data. *Syst Biol* 2004;53:623–637.
64. Jermini L, Ho SY, Ababneh F, Robinson J, Larkum AW. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Syst Biol* 2004;53:638–643.
65. Vera-Ruiz VA, Robinson J, Jermini LS, Holder M. A Likelihood-Ratio Test for Lumpability of Phylogenetic Data: Is the Markovian Property of an Evolutionary Process retained in Recoded DNA? *Syst Biol* 2021:syab074.
66. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;14:587–589.
67. Dziaik JJ, Coffman DL, Lanza ST, Li R, Jermini LS. Sensitivity and specificity of information criteria. *Brief Bioinform* 2020;21:553–565.
68. Schwarz G. Estimating the dimension of a model. *Ann Statist* 1978;6:461–464.

69. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K and Kitagawa G (eds). *Selected Papers of Hirotugu Akaike*. New York, NY: Springer New York; . pp. 199–213.
70. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;37:1530–1534.
71. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018;35:518–522.
72. Rambaut A. Figtree v1.4.4; 2018. <http://tree.bio.ed.ac.uk/software/figtree/>
73. Alexeev D, Alexeeva M, Baxter RL, Campopiano DJ, Webster SP, et al. The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *J Mol Biol* 1988;284:401–419.
74. Johnson EL, Heaver SL, Waters JL, Kim BI, Bretin A, et al. Sphingolipids produced by gut bacteria enter host metabolic pathways impacting ceramide levels. *Nat Commun* 2020;11:2471.
75. Ikushiro H, Islam MM, Okamoto A, Hoseki J, Murakawa T, et al. Structural insights into the enzymatic mechanism of serine palmitoyltransferase from *Sphingobacterium multivorum*. *J Biochem* 2009;146:549–562.
76. Lowther J, Charmier G, Raman MC, Ikushiro H, Hayashi H, et al. Role of a conserved arginine residue during catalysis in serine palmitoyltransferase. *FEBS Lett* 2011;585:1729–1734.
77. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol* 2008;25:1307–1320.
78. Crotty SM, Minh BQ, Bean NG, Holland BR, Tuke J, et al. GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Syst Biol* 2019;69:249–264.
79. Campbell EL, Christman H, Meeks JC. DNA microarray comparisons of plant factor- and nitrogen deprivation-induced hormogonia reveal decision-making transcriptional regulation patterns in *Nostoc punctiforme*. *J Bacteriol* 2008;190:7382–7391.
80. Christman HD, Campbell EL, Meeks JC. Global transcription profiles of the nitrogen stress response resulting in heterocyst or hormogonium development in *Nostoc punctiforme*. *J Bacteriol* 2011;193:6874–6886.
81. Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW. Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC genomics* 2011;12:332.
82. Risser DD, Meeks JC. Comparative transcriptomics with a motility-deficient mutant leads to identification of a novel polysaccharide secretion system in *Nostoc punctiforme*. *Mol Microbiol* 2013;87:884–893.
83. Hannich JT, Mellal D, Feng S, Zumbuehl A, Riezman H. Structure and conserved function of iso-branched sphingoid bases from the nematode *Caenorhabditis elegans*. *Chem Sci* 2017;8:3676–3686.
84. Cheng X, Jiang X, Tam KY, Li G, Zheng J, et al. Sphingolipidomic analysis of *C. elegans* reveals development- and environment-dependent metabolic features. *Int J Biol Sci* 2019;15:2897–2910.
85. Zhang Y, Wang S, Li X-M, Cui C-M, Feng C, et al. New sphingolipids with a previously unreported 9-methyl-C20-sphingosine moiety from a marine alga endophytic fungus *Aspergillus niger* EN-13. *Lipids* 2007;42:759–764.
86. Caudales R, Wells JM. Differentiation of free-living *Anabaena* and *Nostoc* cyanobacteria on the basis of fatty acid composition. *Int J Syst Bacteriol* 1992;42:246–251.
87. Choi KH, Heath RJ, Rock CO. beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* 2000;182:365–370.
88. Shiraiwa Y, Ikushiro H, Hayashi H. Multifunctional role of His159 in the catalytic reaction of serine palmitoyltransferase. *J Biol Chem* 2009;284:15487–15495.
89. Sakaki K, Ohishi K, Shimizu T, Kobayashi I, Mori N, et al. A suicide enzyme catalyzes multiple reactions for biotin biosynthesis in cyanobacteria. *Nat Chem Biol* 2020;16:415–422.
90. Fujisawa T, Narikawa R, Maeda S-I, Watanabe S, Kanesaki Y, et al. CyanoBase: a large-scale update on its 20th anniversary. *Nucleic Acids Res* 2017:D551–D554.
91. An D, Oh SF, Olszak T, Neves JF, Avci FY, et al. Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* 2014;156:123–133.
92. Moyer ZD, Valiuskyte K, Dewhurst FE, Nichols FC, Davey ME. Synthesis of sphingolipids impacts survival of *porphyromonas gingivalis* and the presentation of surface polysaccharides. *Front Microbiol* 2016;7:1919.
93. Schmidt A, Sivaraman J, Li Y, Larocque R, Barbosa JA, et al. Three-dimensional structure of 2-amino-3-ketobutyrate CoA ligase from *Escherichia coli* complexed with a plp-substrate intermediate: inferred reaction mechanism. *Biochemistry* 2001;40:5151–5160.
94. Stankeviciute G, Guan Z, Goldfine H, Klein EA. *Caulobacter crescentus* adapts to phosphate starvation by synthesizing anionic glycolipids and a novel glycosphingolipid. *mBio* 2019;10:e00107-19.
95. Astner I, Schulze JO, van den Heuvel J, Jahn D, Schubert W-D, et al. Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *EMBO J* 2005;24:3166–3177.

### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](http://microbiologyresearch.org).