



Cite this: *RSC Chem. Biol.*, 2025, 6, 307

## Uncovering the substrate of olefin synthase loading domains in cyanobacteria *Picosynechococcus* sp. strain PCC 7002<sup>†</sup>

Audrey E. Yñiguez-Gutierrez,<sup>ID</sup><sup>a</sup> Erin Conley,<sup>b</sup> Michael G. Thomas<sup>b</sup> and Brian F. Pfleger<sup>\*a</sup>

Cyanobacteria are widespread, photosynthetic, gram-negative bacteria that generate numerous bioactive secondary metabolites *via* complex biosynthetic enzymatic machinery. The model cyanobacterium *Picosynechococcus* sp. strain PCC 7002, hereafter referred to as PCC 7002, contains a type I polyketide synthase (PKS), termed olefin synthase (OlsWT), that synthesizes 1-nonadecene and 1,14-nonadecadiene:  $\alpha$ -olefins that are important for growth at low temperatures. The putative biochemistry encoded by the PKS domains suggests that OlsWT will create an olefin with one additional carbon relative to the original substrate (+1 mechanism). The first domain in the multi-module OlsWT protein has homology to fatty acyl-AMP ligases (FAALs) that typically activate free fatty acids prior to creating novel thioester linkages. Paradoxically, unmodified wildtype PCC 7002 is not known to maintain a substantial pool of free fatty acids, and prior work demonstrated conversion of exogenous pentadecanoic acid to 1-octadecene instead of the expected 1-hexadecene. In this study, we developed PCC 7002 as a heterologous host to facilitate the expression and study of Ols proteins in effort to discover their true substrates. Here, we report the successful expression of two Ols homologs from *Geminocystis* sp. NIES-3709 and *Xenococcus* sp. PCC 7305 in PCC 7002 that generated 1-heptadecene and 1-pentadecene, respectively. Through the additional deletion of a gene encoding an acyl-acyl carrier protein (ACP) synthetase (Aas) responsible for activation of exogenous free fatty acids, we demonstrated the expected conversion of exogenously provided odd-chain fatty acids to  $\alpha$ -olefins containing one additional carbon. These data suggest that short-lived fatty acids liberated from lipid membranes are the Ols substrate. We subsequently confirmed OlsWT activity on octadecanoic acid *via* *in vitro* chrome azurol S assay using a purified FAAL module. Collectively, this work clarifies the *in vivo* substrate of Ols FAAL domains and identifies the FAAL module as a target for future bioengineering to allow access to desired  $\alpha$ -olefins.

Received 29th September 2024,  
Accepted 31st December 2024

DOI: 10.1039/d4cb00234b

[rsc.li/rsc-chembio](http://rsc.li/rsc-chembio)

## Introduction

Cyanobacteria synthesize and maintain cell and thylakoid membranes that differ significantly from other bacteria. Cyanobacterial membranes consist mostly of glycolipids with the major phospholipid found in the thylakoid membrane being phosphatidylglycerol.<sup>1</sup> In addition to these lipid components, cyanobacterial membranes contain hydrocarbons in either alkane or alkene forms. Cyanobacteria hydrocarbons – including alkanes, internal olefins, and terminal olefins – are

implicated in maintaining cell size, growth, and membrane properties.<sup>2</sup> These alkanes/alkenes also play an important role in cyanobacterial halotolerance and temperature adaptation, through modulating cyclic electron flow.<sup>3–5</sup> The more widely distributed pathway for alkane biosynthesis is the fatty acyl-ACP reductase/aldehyde deformylating oxygenase (Aar/Ado) pathway.<sup>6</sup> The Aar reductively cleaves acyl-ACPs to produce a fatty aldehyde that undergoes deformylation catalyzed by Ado to produce an alkane with the net loss of one carbon.<sup>7</sup> Cyanobacterial terminal alkenes ( $\alpha$ -olefins) are produced by a type I polyketide synthase (PKS) termed olefin synthase (OlsWT). OlsWT was first identified in *Picosynechococcus* sp. strain PCC 7002, hereafter referred to as PCC 7002, and has been previously evaluated by our group and others.<sup>5,8–10</sup> The primary hydrocarbons produced by OlsWT in PCC 7002 are 1-nonadecene and 1,14-nonadecadiene. OlsWT is a multidomain enzyme consisting of a fatty acyl-AMP ligase (FAAL) domain hypothesized to initiate

<sup>a</sup> Department of Chemical and Biological Engineering,  
University of Wisconsin – Madison, Madison, Wisconsin, 53706, USA.

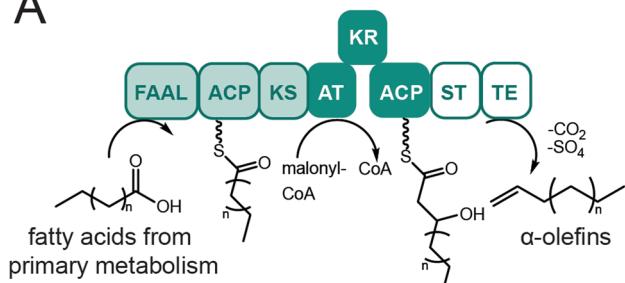
E-mail: [brian.pfleger@wisc.edu](mailto:brian.pfleger@wisc.edu)

<sup>b</sup> Department of Bacteriology, University of Wisconsin – Madison, Madison, Wisconsin, 53706, USA

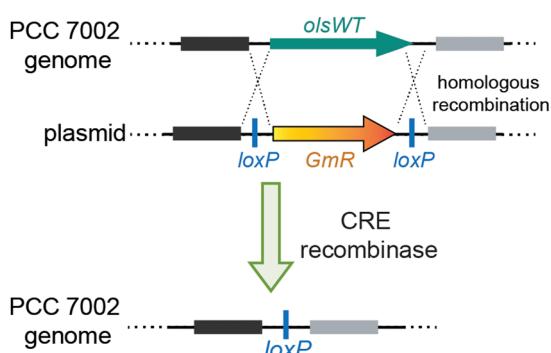
<sup>†</sup> Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4cb00234b>



A



B



C

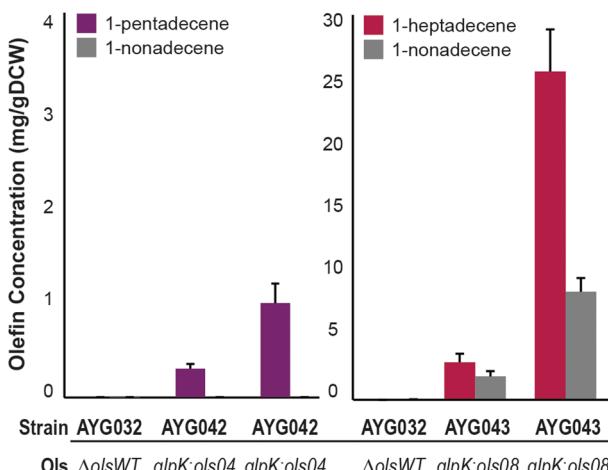


Fig. 1 (A) The OlsWT FAAL domain selects fatty acyl chains from primary metabolism for conversion to  $\alpha$ -olefin hydrocarbons. (B) Optimized genetic manipulations in *Picosynechococcus* sp. PCC 7002 using homologous recombination with a selective marker followed by removal via a CRE recombinase to provide a markerless deletion. (C)  $\alpha$ -Olefin production in PCC 7002 strains expressing Ols homologs. All data represent the mean  $\pm$  s.d. of biological triplicates.

transfer of a fatty acyl chain to the first acyl carrier protein (ACP<sub>0</sub>) domain. The acyl chain is subsequently extended by a two-carbon unit (*via* malonyl-CoA) by the ketosynthase (KS) and acyltransferase (AT) domains. The resulting ACP-domain-linked  $\beta$ -keto intermediate is reduced by the ketoreductase (KR) domain. Finally, the  $\beta$ -hydroxyl group of the acyl chain is activated *via* sulfation by the sulfotransferase (ST) domain. The sulfate provides a leaving group for simultaneous

decarboxylation and sulfate elimination catalyzed by the thioesterase (TE) domain. Overall, the combined enzymology releases an odd-numbered  $n + 1$   $\alpha$ -olefin relative to the  $n$ -carbon substrate (Fig. 1A). The exact substrate for the initial fatty acyl chain activated by the FAAL domain of OlsWT remains unclear. Canonically, FAALs activate free fatty acids *via* a two-step adenylation mechanism utilizing adenosine triphosphate (ATP) to first form the acyl adenylate (acyl-AMP) and eject pyrophosphate.<sup>11</sup> The activated acyl chain is then transferred to the 4'-phosphopantetheinyl arm of the downstream ACP domain. This second step distinguishes FAALs from fatty acyl-CoA ligases (FACLS) that transfer the acyl-AMP to coenzyme A.<sup>12,13</sup> FAAL domains in cyanobacterial secondary metabolite biosynthetic gene clusters have been described and evaluated for substrate usage previously. For example, the incorporation of free fatty acids into the scaffolds of nocuolin A, the hapalosins, and the chlorosphaerolactylates has been demonstrated using *in vivo* labeling and *in vitro* assays.<sup>14–16</sup> However, cyanobacteria have not been found to actively produce or accumulate a large pool of free fatty acids directly from either acyl-CoAs or acyl-ACPs.<sup>17,18</sup> Likewise, the presence of triacylglycerols (TAGs) has been observed in only a small minority of cyanobacteria, namely the Nostocales order.<sup>19,20</sup> Like other bacteria, cyanobacteria generate fatty acids and lipids through fatty acid biosynthesis with pathway intermediates covalently linked to ACPS as acyl-ACP thioesters.<sup>21</sup> To date, no native acyl-ACP thioesterases (TEs) have been identified in cyanobacteria that would hydrolyze free fatty acids directly from ACPS involved in fatty acid biosynthesis. In contrast, native plant acyl-ACP TEs, such as FATA or FATB, directly hydrolyze acyl-ACPS from fatty acid biosynthesis to produce free fatty acids.<sup>22,23</sup>

In general, cyanobacteria also lack complete, active  $\beta$ -oxidation pathways,<sup>24,25</sup> as they do not contain the genes encoding for key  $\beta$ -oxidation enzymes such as 3-hydroxyacyl-CoA dehydrogenase (FadB), acyl-CoA dehydrogenase (FadE), and 3-ketoacyl-CoA thiolase (FadI).<sup>16</sup> FadD acyl-CoA ligase homologs that would activate long chain fatty acids to the CoA species have been annotated for several cyanobacterial species within the KEGG (Kyoto Encyclopedia of Genes and Genomes) database.<sup>26</sup> However, these are proposed to be mis-annotations of acyl-ACP synthetases (Aas), which are also adenylate-forming enzymes that cyanobacteria use to incorporate exogenous fatty acids into lipid and fatty acid biosynthesis *via* the formation of a thioester to an ACP.<sup>21</sup> These acyl-ACP Aas products are then elongated *via* fatty acid biosynthesis and ultimately condensed into lipids.

In prior work, we detected elevated production of 1-octadecene in PCC 7002 strains fed pentadecanoic acid, but not the expected  $n + 1$  product 1-hexadecene.<sup>8</sup> Therefore, in this study, we developed PCC 7002 as a heterologous host to facilitate the expression and study of Ols proteins in effort to discover the source of their true substrates. Our findings indicate that free fatty acids liberated from lipid membranes are the native substrates of Ols, raising the possibility of engineering Ols variants that could access shorter chain  $\alpha$ -olefins for use as fuel additives.



## Results and discussion

### Heterologous host PCC 7002 synthesizes variable chain length $\alpha$ -olefins encoded by Ols homologs

This work was initially focused on the production of shorter chain  $\alpha$ -olefins *via* expression of Ols homologs known to generate products of varying length. The genes encoding Ols homologs are present in only a small subset of cyanobacterial species ( $\sim 10\%$ ), none of which have been evaluated in a heterologous host.<sup>27,28</sup> In order to introduce Ols homologs into the PCC 7002 genome, the endogenous *olsWT* gene was first deleted by introducing a gentamicin resistance cassette *via* homologous recombination directed by long homology arms. The CRE-lox recombinase was then employed to remove the antibiotic resistance cassette, allowing the generation of a markerless deletion and recycling the selective marker for future genetic manipulations (Fig. 1B).<sup>29</sup> Subsequently, genes coding for individual Ols homologs were integrated downstream to the *glpK* gene in the unmarked  $\Delta olsWT$  PCC 7002 mutant strain under an inducible synthetic promoter cLac143.<sup>30</sup> Homologs Ols04 from *Xenococcus* sp. PCC 7305 and Ols08 from *Geminocystis* sp. NIES-3709 were chosen since they should not produce the endogenous 1-nonadecene. Hydrocarbon products of Ols04 were previously found to be 1-pentadecene and 2-pentadecene, as extracted directly from *Xenococcus* sp. PCC 7305. While the hydrocarbon profile for Ols08 was not directly quantified, it was hypothesized that it would produce 1-heptadecene based on the previous phylogenetic analysis.<sup>27</sup> The cultures containing Ols homologs were induced with 1 mM IPTG and subsequently extracted and analyzed *via* gas chromatography/mass spectrometry (GC/MS). Strain AYG042 encoded for the homolog Ols04 from *Xenococcus* sp. PCC 7305; initial results showed only trace amounts of the expected 1-pentadecene (Fig. 1C). PCC 7002 does not accumulate high levels of the 1-pentadecene precursor, tetradecanoic acid, which we hypothesized led to lower titers of this  $\alpha$ -olefin product. This was confirmed by a total lipid extraction and fatty acid methyl ester (FAME) analysis of the culture, which showed only small amounts of tetradecanoic species (Fig. S1, ESI<sup>†</sup>). For strain AYG043, integration of homolog Ols08 from *Geminocystis* sp. NIES-3709 resulted in the production of 1-heptadecene and a small amount of 1-nonadecene (Fig. 1C). Ols08 was expected to only produce the 1-heptadecene due to its homology to previously characterized olefin synthases.<sup>27</sup> The ability for Ols08 to accept both hexadecanoic and octadecanoic acyl chains was not completely unexpected as *olsWT* has been shown to accept both the octadecanoic and heptadecanoic substrates, indicating that these enzymes are moderately flexible in which acyl chain they will load to the initial ACP domain. There was also a small amount of these  $\alpha$ -olefins present in the non-induced cultures indicating leaky gene expression from our genetic cassette. The precursor strain to both of these Ols homolog strains, AYG032 ( $\Delta olsWT$ :*GmR*), lacked any olefin production confirming the *olsWT* deletion. These results established that Ols homologs can be successfully expressed in PCC 7002 to generate shorter chain olefins. These results, however,

did not answer the fundamental question of the *in vivo* substrate of the Ols FAAL domain. Even though FAAL domains canonically utilize free fatty acids as substrates, cyanobacteria do not accumulate free fatty acids, and PCC 7002 has no homologs for FATA/B acyl-ACP TE enzymes responsible for the direct hydrolysis of fatty acyl-ACPs (Table S1, ESI<sup>†</sup>).<sup>31</sup>

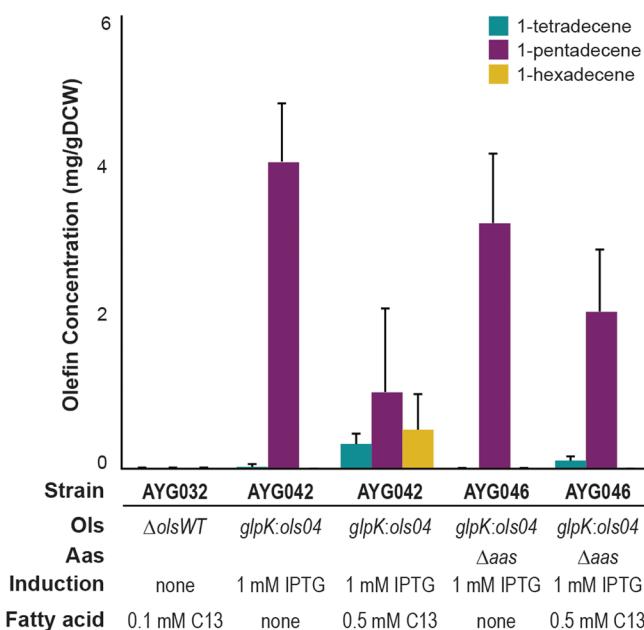
### Ols homologs utilize free fatty acid to generate $\alpha$ -olefins

In order to further investigate the FAAL domain substrate in Ols homologs, cultures were supplemented with non-native odd-chain free fatty acids to track the incorporation of these species into the lipid profile. If free fatty acids are the substrate of the FAAL domain, we hypothesized that the Ols homologs would be able to directly utilize supplemented odd-chain fatty acids. Therefore, strain AYG042 encoding Ols04 was supplied with 0.5 mM tridecanoic acid. Analysis of these lipid products did not contain the expected *n* + 1  $\alpha$ -olefin product, 1-tetradecene. Interestingly, a new 1-hexadecene product was observed, consistent with the addition of *n* + 3 carbons, albeit in similarly low titers (Fig. 2A). Strain AYG043 encoding homolog Ols08 was also supplemented with 0.1 mM pentadecanoic acid, which was expected to yield 1-hexadecene. Again, upon extraction of olefins following cell growth, we did not observe the *n* + 1  $\alpha$ -olefin but instead substantial amounts of the *n* + 3 product, 1-octadecene (Fig. 2B). Based on these results, we hypothesized that the fed odd-chain fatty acids were internalized and activated by the acyl-ACP synthetase (Aas) to produce a bound acyl-ACP species that could then enter the fatty acid synthesis (FAS) cycle for an elongation by two carbons before being intercepted by the Ols homologs. This hypothesis was bolstered by analyses of the entire lipid environment to demonstrate that tridecanoic acid was being activated and elongated to pentadecanoic and heptadecanoic acids in AYG042 (Fig. S2A, ESI<sup>†</sup>). A corroborating result was observed for strains AYG043 when fed pentadecanoic acid; the levels of heptadecanoic acid increased in the fed cultures as compared to control (Fig. S3A, ESI<sup>†</sup>). Therefore, the odd-chain free fatty acids are likely first activated by Aas to produce the ACP-bound acyl chain followed by FAS elongation prior to being intercepted by Ols. While these data indicate that ACP-bound acyl species may be the Ols substrate, there is no precedent for the FAAL domain to transfer an ACP-bound acyl chain to a *trans* ACP. Rather, FAAL domains activate free fatty acids to the acyl-AMP for subsequent transfer to the 4'-phosphopantetheinyl arm of a *holo*-ACP.<sup>11,32</sup> Therefore, we sought to identify a pathway that allows liberation of the elongated ACP-bound acyl species to free fatty acids.

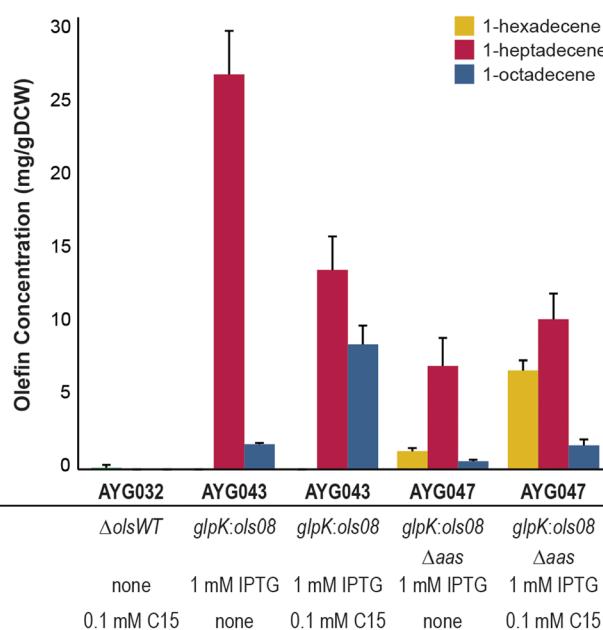
Since no acyl-ACP thioesterases have been identified in cyanobacteria to date, we hypothesized that the Ols free fatty acid substrate may be a result of lipases/hydrolases liberating free fatty acids from the lipid or thylakoid membranes.<sup>1</sup> These free fatty acids would eventually be re-activated to ACP-bound acyl chains by Aas, but may also be briefly available to Ols for  $\alpha$ -olefin synthesis (Fig. 2C). To test this, the gene encoding for Aas was deleted and replaced with a gene encoding kanamycin resistance in strains AYG042 and AYG043 to generate strains AYG046 and AYG047, respectively. The same odd-chain feeding



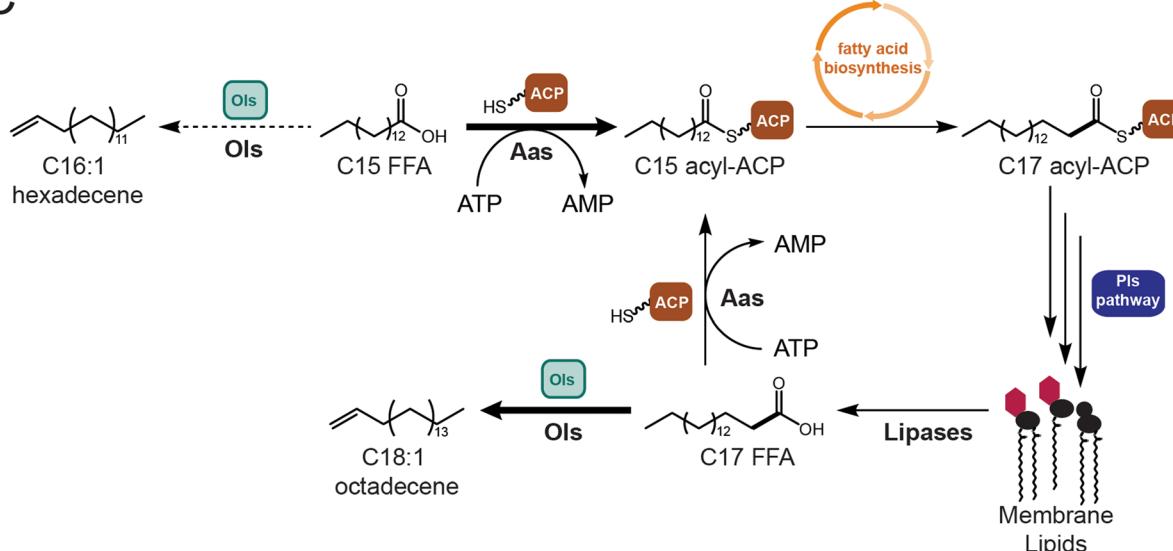
A



B



C



**Fig. 2** Ols homologs utilize free fatty acids as *in vivo* substrates to generate  $\alpha$ -olefins (A) evaluation of  $\alpha$ -olefins extracted from heterologous expression of Ols04 in PCC 7002. Deletion of *aas* results in the loss of the  $n + 3$  product 1-hexadecene from fed tridecanoic acid. (B) Evaluation of  $\alpha$ -olefins extracted from heterologous expression of Ols08 in PCC 7002. Deletion of *aas* results in a decrease of the  $n + 3$  product 1-octadecene from fed pentadecanoic acid. (C) Clarified  $\alpha$ -olefins biosynthesis in cyanobacteria. Free fatty acids are quickly activated by *Aas* to acyl-ACPs, which are extended in fatty acid biosynthesis and/or integrated into membrane lipids. Lipases liberate free fatty acids in lipid remodelling to provide short-lived substrates for Ols. All data represent the mean  $\pm$  s.d. of biological triplicates.

strategy was employed, and olefins were extracted for analysis. Supplementation of AYG046 with tridecanoic acid resulted in only a small amount of direct conversion to the  $n + 1$  olefin 1-tetradecene. Conversely, the production of 1-hexadecene was completely abolished in AYG046, indicating that the deletion of *aas* prevented extension to the pentadecanoate product (Fig. 2A). In strain AYG047, the exogenously provided pentadecanoic acid was converted to the  $n + 1$  olefin 1-hexadecene, with only a small amount of the  $n + 3$  olefin 1-octadecene being

observed (Fig. 2B). A total lipid extraction of AYG047 showed an increase in the amount of accumulated pentadecanoate species as compared to the fed AYG043 strain, while simultaneously showing a decrease in the amount of conversion to the extended  $n + 2$  heptadecanoate acyl chains (Fig. S3A and B, ESI<sup>†</sup>). The lack of significant production of 1-tetradecene in AYG046 supplemented with tridecanoic acid may be a result of poor internalization of the tridecanoic acid following deletion of the gene encoding *Aas*. The total lipid extraction of AYG042



shows a dramatic increase in the  $n + 2$  pentadecanoate acyl chains when tridecanoic acid is fed (Fig. S2A, ESI<sup>†</sup>). However, upon deletion of *aas*, strain AYG046 does not accumulate either tridecanoate nor pentadecanoate species in its lipid environment, indicating that very little of the tridecanoic acid is entering the cell to act as substrate for Ols04 to produce 1-tetradecene (Fig. S2B, ESI<sup>†</sup>). Together, however, these results indicate that deletion of *aas* significantly diminished the sequestration of the fed free fatty acids to the ACP-bound forms to allow the Ols homologs to directly access these substrates for  $\alpha$ -olefin production.

### Deciphering OlsWT acyl-chain loading *in vitro*

To provide further support that the free fatty acid was the substrate of the FAAL domain of OlsWT, an *in vitro* analysis was completed. The region of *olsWT* that encodes for the di-domain containing the FAAL domain and the initial ACP<sub>0</sub> was cloned into a protein expression vector that introduced a C-terminal histidine tag and overproduced in *E. coli* BAP1 containing the promiscuous 4'-phosphopantetheinyl transferase Sfp. This truncated OlsWT enzyme was evaluated for fatty acyl activation *in vitro* using the colorimetric assay chrome azurol S (CAS) assay (Fig. 3A).<sup>33</sup> In order to make the forward adenylation reaction to form the acyl-AMP by the FAAL domain irreversible, hydroxylamine was added to the *in vitro* assay. The pyrophosphate released from this reaction then sequesters iron from the CAS-Fe<sup>3+</sup> complex (blue in colour) to produce free CAS (yellow in colour) resulting in a decrease in the absorbance measured at 630 nm.

The CAS assay additionally afforded an assessment of selectivity for specific fatty acyl chain lengths by the FAAL domain of Ols. From these results, we observed that the FAAL-ACP di-domain of OlsWT shows a significant preference in fatty acyl chain loading based on chain length. Octadecanoic acid was demonstrated to be the preferred substrate for OlsWT (Fig. 3B). To confirm these results, we also analysed related fatty acyl compounds including fatty alcohols and fatty methyl esters, which should not be activated by the FAAL domain of OlsWT. As expected, these fatty acyl compounds did not result in a significant change in the absorbance of the CAS reagent (Fig. S4, ESI<sup>†</sup>). We next aimed to assess free fatty acid activation by the full-length OlsWT. Unfortunately, we were not successful in purifying a soluble preparation of this large (300 kDa) enzyme (data not shown). This may be a result of improperly folded protein lacking key secondary and tertiary structures required for function. Together, these data further establish free fatty acids as the substrate for the FAAL domain of Ols and demonstrate chain length selectivity with octadecanoic acid as the preferred substrate. These results are also consistent with our data that deletion of the gene encoding *aas* prevented the sequestration of the fed free fatty acids to the ACP bound forms to allow Ols to directly intercept the free fatty acid substrate.

### Olefin synthases are widely distributed in cyanobacteria

Previous groups have shown that olefin synthases are found across different morphotypes of cyanobacteria, including

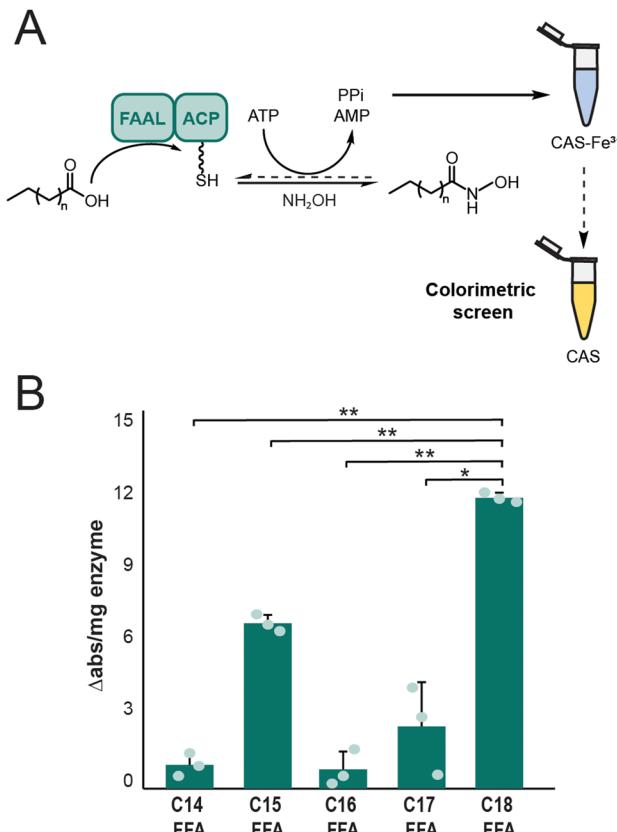
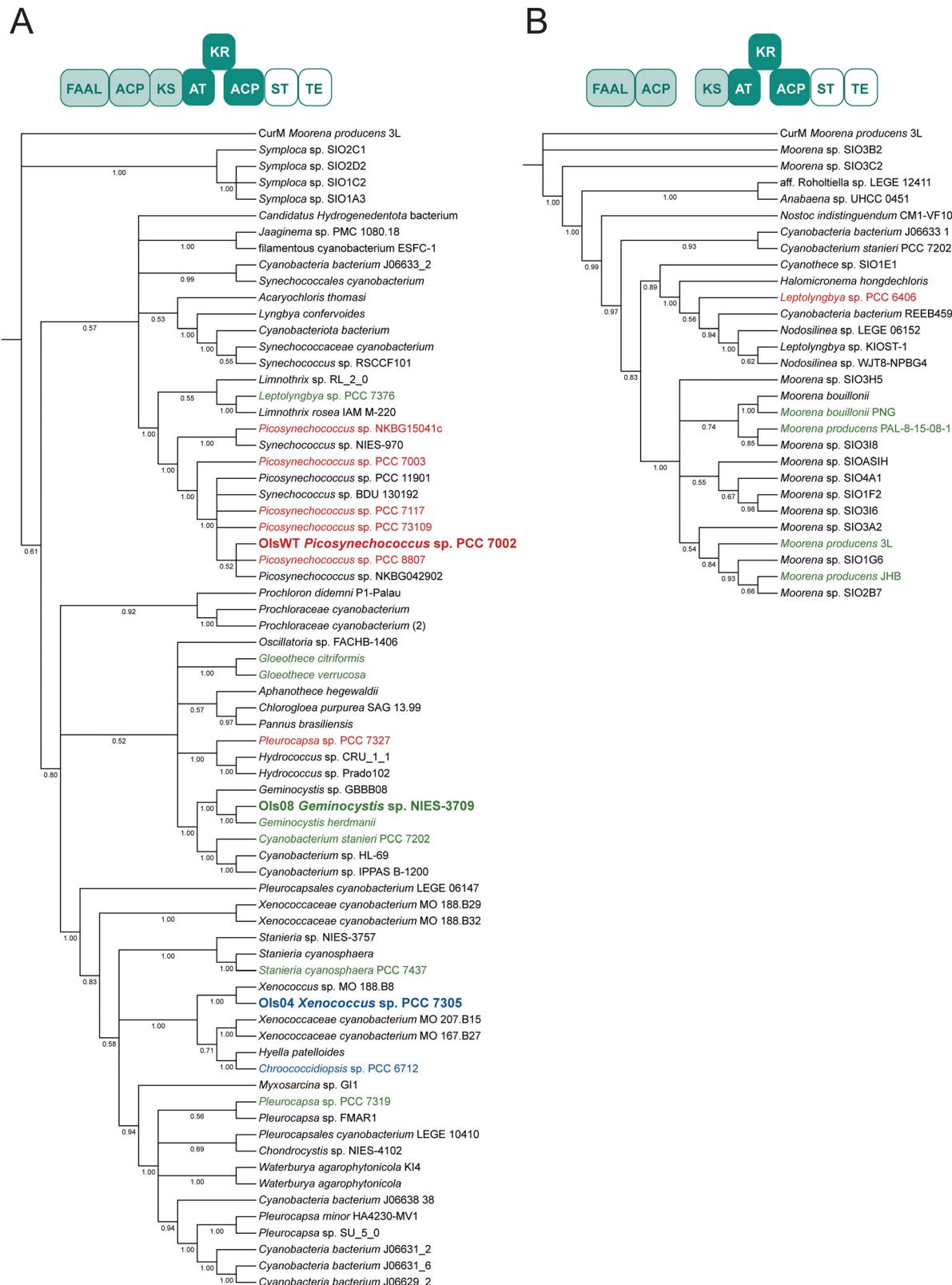


Fig. 3 The FAAL domain of OlsWT demonstrates significant preference for stearic acid *in vitro*. (A) Overview of CAS colorimetric assay *in vitro*. The hydroxylamine reagent drives the FAAL adenylation reaction forward by capturing the acyl-AMP intermediate as a hydroxamic acid. Accumulated pyrophosphate (PP<sub>i</sub>) chelates Fe<sup>3+</sup> from the CAS reagent to decrease absorbance at 630 nm. (B) *In vitro* FAAL activation for Ols LD-ACP di-domain demonstrating a significant preference for octadecanoic acid over other fatty acids. All data represent the mean  $\pm$  s.d. of biological triplicates. *P* values were analysed based on student two-tailed *t* test assuming equal variances. \**P* = 0.0126, \*\**P* < 0.0001.

unicellular, colonial, and filamentous.<sup>27,28</sup> We updated the list of predicted Ols homologs by confirming previous sequences and utilizing them for blastp searches against cyanobacterial genomes from GenBank (Tables S3 and S4, ESI<sup>†</sup>). New homologs were confirmed to contain the sulfotransferase (ST) domain responsible for sulfonation of the  $\beta$ -hydroxy acyl intermediate. Any homologs that lacked the FAAL domain responsible for loading were excluded. Homologs that do not contain the FAAL domain were often found to be encoded upstream or downstream of genes coding for other PKS or NRPS modules, indicating that they belonged to non-olefin biosynthetic gene clusters such as the marine cyanobacterial natural product curacin A.<sup>34,35</sup> Finally, homologs were confirmed to contain a terminal TE domain responsible for the termination and release of the product to allow for subsequent decarboxylation and desulfonation. Protein homologs containing the ST domain but lacking a TE domain are likely to result in sulfated metabolites instead of  $\alpha$ -olefins.<sup>36,37</sup>

Publicly available cyanobacteria genomes were also reviewed for both major types of Ols architectures. The open reading





**Fig. 4** (A) Phylogeny of 71 Ols homologs with fused, *cis* modules identified from available cyanobacterial genomes. (B) Phylogeny of 29 Ols homologs with decoupled, *trans* modules with FAAL-ACP on a separate ORF. Ols homologs with confirmed  $\alpha$ -olefins production are highlighted as follows: red, 1-nonadecene; green, 1-heptadecene; blue, 1-pentadecene. Homologs directly studied in this work in bold. Trees were generated by the maximum likelihood method with 500 bootstrap replicates. Bootstrap values  $> 0.5$  are shown. CurM from the curacin pathway was used to root both trees.

frame (ORF) for OlsWT from PCC 7002 encodes for all domains of the enzyme (FAAL-ACP-KS-AT-KR-ACP-ST-TE) in a single protein. Other Ols homologs are arranged in two contiguous ORFs with the first encoding for the first two domains (FAAL-ACP) in *trans* and the second encoding for the remaining domains (KS-AT-KR-ACP-ST-TE). In total, we identified 48 new Ols homologs that contain all domains in a single ORF and 22 new Ols homologs with the FAAL-ACP domains on separate ORFs. A previously identified Ols homolog with the FAAL encoded by the first ORF and the subsequent domains (ACP-KS-AT-KR-ACP-ST-TE) on a second ORF from *Cyanobacterium stanieri* PCC 7202 was also evaluated. No homologs containing this third, rare architecture were identified.

Two phylogenetic trees of both types of Ols architectures were constructed using the CurM domain from the curacin A pathway as the outgroup (Fig. 4). Previously reported hydrocarbon profiles along with our observed heterologous hydrocarbon production demonstrate that related Ols pathways produce similar  $\alpha$ -olefins. However, these hydrocarbon profiles are not exclusive to either Ols architecture. For example, the production of 1-heptadecene is widespread in both types of Ols organization. The production of shorter chain 1-pentadecene appears to be the rarest with only two confirmed strains capable of its biosynthesis. While this may be due to substrate selection *via* the FAAL domain, it could also be attributed to substrate availability. *Chroococcidiopsis* sp. PCC 6712 and *Xenococcus* sp. PCC 7305, which contain Ols homologs capable of synthesizing 1-pentadecene, both produce tetradecanoyl lipid chains. Conversely, strains that contain Ols homologs capable of producing 1-nonadecene tend to accumulate longer hexadecanoyl and octadecanoyl lipids.

## Conclusions

It was previously hypothesized that the Ols substrates were the abundant acyl-ACPs similar to the Aar/Ado pathway. However, the initiating FAAL domain of Ols has homology (29.9% similarity) to characterized FAAL domains that activate free fatty acids.<sup>38,39</sup> Therefore, we sought to confirm free fatty acids as the Ols substrate while also investigating the heterologous expression of Ols homologs capable of synthesizing shorter chain  $\alpha$ -olefins. Through deletion of *olsWT* in PCC 7002 and subsequent genomic integration of *ols04* and *ols08*, we were able to access 1-pentadecene and 1-heptadecene, respectively. To further confirm free fatty acids as the Ols substrates, feeding studies utilizing odd-chain fatty acids were performed. This led to the production of  $\alpha$ -olefins with a net addition of three carbons, indicating that fed free fatty acids are quickly incorporated into fatty acid biosynthesis *via* Aas, which is consistent with Aas being involved in fatty acid intake and activation to acyl-ACPs.<sup>24,40</sup> The deletion of *aas* was performed and the direct incorporation of odd-chain free fatty as  $n + 1$   $\alpha$ -olefins was observed. Additionally, *in vitro* activation of free fatty acids by the OlsWT FAAL-ACP shows that these FAAL domains may also have a substrate selectivity mechanism to preferentially select

specific chain lengths. Similar results were observed in the production of secondary metabolite nocuolin A where a FAAL domain fused to an ACP domain preferentially loads hexanoic or octanoic acid to initiate biosynthesis.<sup>14</sup> In this case, the authors were able to observe incorporation of exogenously provided free fatty acids in producing strain *Nodularia* sp. LEGE 06071, which may be due to lower incorporation of these short-chain fatty acids into the fatty acyl-ACP pool by Aas.

Overall, this work indicates that free fatty acids are the substrate of Ols enzymes. Given that free fatty acids do not accumulate in cyanobacteria, these substrates are likely a result of free fatty acids being liberated from the lipid and thylakoid membranes by lipases prior to their reactivation by Aas (Fig. 2C). Unfortunately, only a few putative lipases have been characterized in cyanobacteria. One example is LipA in *Synechocystis* sp. PCC 6803, which hydrolyzes glycolipids at the *sn*-1 position to increase fatty acyl recycling and photosystem repair under high light conditions.<sup>41,42</sup> However, the LipA homolog in PCC 7002 (SYNPCC7002\_A1441) has not been investigated for its involvement in these processes nor as the source of the free fatty acid substrate for OlsWT. Another ORF from *Synechocystis* sp. PCC 6803, sll0482, is suspected to be responsible for diacylation at the *sn*-2 position of glycolipids, but there is no homolog of this lipase in PCC 7002 to perform an analogous function.<sup>43</sup> During this work, we identified a putative hydro-lase encoded directly upstream of *olsWT* in PCC 7002 (SYNPCC7002\_A1174), which contained conserved domains typically associated with alpha/beta hydrolases. We hypothesized that this uncharacterized enzyme could be at least partially responsible for liberation of these fatty acids for Ols incorporation. Deletion of this hydrolase produced no changes in  $\alpha$ -olefin production (Fig. S5, ESI†). Although, this may be a result of a redundant lipase capable of functional complementation. A more complete analysis of the lipase environment in PCC 7002 will facilitate future engineering efforts to increase  $\alpha$ -olefin production.

## Experimental

### Chemicals, reagents, and media

Wildtype *Picosynechococcus* sp. PCC 7002 is from the Pasteur Culture Collection. All PCC 7002 strains were grown on Media A+ with 1.5% (w/v) Bacto-Agar (Fisher), when necessary.<sup>44</sup> Mutated strains containing antibiotic markers were grown on selective media (kanamycin, 50–100  $\mu$ g mL<sup>-1</sup>; gentamicin, 10–30  $\mu$ g mL<sup>-1</sup>). Starter cultures were grown in glass culture tubes (2  $\times$  15 cm) with 10–20 mL media A+ bubbled with air (0.04% CO<sub>2</sub>) at 37 °C with a light intensity of 200  $\mu$ E. For larger scale cultures, PCC 7002 was grown in 250 mL baffled flasks in 50 mL of Media A+ with continuous light from LED strips mounted to the top of a shaking incubator to provide  $\sim$  400  $\mu$ E. Cultures were maintained at 37 °C and supplemented with 1% CO<sub>2</sub>. Evaporation of cultures was corrected daily with water based on initial weight. *Escherichia coli* cultures were grown in LB media (Bacto) at 37 °C.



## Strain construction

Plasmids used are listed in Table S6 (ESI<sup>†</sup>). *E. coli* DH5 $\alpha$  was used to clone all plasmids. Plasmids contained  $\sim$ 500 base pair homology arms targeted to the integration site. Homology arms were amplified directly from wildtype PCC 7002 to construct plasmids *via* Gibson assembly; primers used are listed in Table S7 (ESI<sup>†</sup>). Genetic alterations were performed using homologous recombination and natural competency of PCC 7002, as previously described.<sup>45</sup> Briefly, DNA was transformed into PCC 7002 by adding 1–2  $\mu$ g of purified plasmid to 1 mL of concentrated overnight culture ( $OD_{730} \leq 1$ ). Cultures were placed in direct light at 37 °C for 4 hours. The cells were plated on selective media and grown in direct light to select for transformants, which were subsequently patched onto fresh selection plates. Segregation was analysed by colony PCR (see Table S7, ESI<sup>†</sup>). For recombinase assays, strain  $\Delta$ ols::lox66-aaC1-lox71 was transformed as described above using pCJ111 and plated on selective Media A+ plates containing kanamycin to select for the integration of the Cre-lox recombinase at the *rbcLXS* locus.<sup>29</sup> Following incubation for 4 days, individual transformants were patched onto non-selective Media A+ plates for an additional 4–5 days. From these patch plates, colonies were repatched onto both non-selective Media A+ plates and plates supplemented with gentamicin to confirm the loss of the resistance cassette. Colony PCR was used to confirm the recombinase activity. Genomic DNA was also extracted from constructed strains (Zymo Quick-DNA Fungal/Bacterial Mini-prep Kit) and submitted for full genome sequencing (Plasmid-saurus) to confirm deletions and integrations.

## Analysis of olefin or fatty acyl production

Strains of PCC 7002, including mutant Ols strains (Table S5, ESI<sup>†</sup>) were inoculated into 50 mL Media A+ containing antibiotics, if necessary, at an initial  $OD_{730} = 0.05$  from seed cultures.  $OD_{730}$  was monitored daily for 5 days. After 120 hours, 2.5 mL of culture was transferred to a glass hungate tube for analysis. For  $\alpha$ -olefin quantification, the culture was centrifuged, and the supernatant was removed. The resulting pellet was resuspended in 2.5 mL water prior to extraction. Hexadecane in ethanol was added to each sample (final concentration 50 or 100 mg L<sup>-1</sup>) as an internal standard. External, recovery samples containing  $\alpha$ -olefins C12–C19 were prepared from 1 mg L<sup>-1</sup> to 100 mg L<sup>-1</sup>. These samples were extracted identically to experimental samples, depending on chain length of olefins.

A direct hexane extraction was used for short chain  $\alpha$ -olefin analysis (C12:1–C15:1). Glacial acetic acid (100  $\mu$ L) was added to each sample followed by 1 mL of *n*-hexane. Samples were vortexed vigorously for 20 minutes at 1500 rpm. The layers were separated *via* centrifugation and the organic layer was removed for analysis.

For long chain  $\alpha$ -olefin analysis (C16:1–C19:1), a modified fatty acid extraction method was used. Glacial acetic acid (100  $\mu$ L) was added to each sample and a 1:1 solution of chloroform:methanol (5 mL) was added. Samples were

vortexed vigorously for 10 minutes at 1500 rpm. The layers were separated *via* centrifugation and the aqueous layer was removed by aspiration. The organic layer was removed under reduced pressure, and the dried residue was resuspended in 1 mL *n*-hexane for analysis.

For fatty acid analysis, nonanoic acid and nonadecanoic acid were used for internal standards (100 mg L<sup>-1</sup>). Extraction and methylation of fatty acyl species were completed as previously described.<sup>46</sup>

## Gas chromatography/mass spectrometry (GC/MS) analysis of olefins

Gas chromatography/mass spectrometry (GC/MS) was performed on a coupled Shimadzu GCMS-QP2010S. 1  $\mu$ L of olefin extracts were injected into an DB-Fatwax Ultra Inert column (polyethylene glycol-type; 30 m length; 0.25 mm inner diameter; 0.25  $\mu$ m film thickness; Agilent) with a split ratio of 1:5. The extracts were separated using the following temperature program: hold at 75 °C (2 min), ramp at 10 °C min<sup>-1</sup> to 180 °C, hold at 180 °C (2 min), ramp at 50 °C min<sup>-1</sup> to 250 °C, hold at 250 °C (2 min). The column flow rate was 2.14 mL min<sup>-1</sup> with an injection temperature of 250 °C. Helium was used as the carrier gas with an inlet pressure of 1.4 bar. The mass spectra were recorded at an ion source and interface temperature of 230 °C and 250 °C, respectively. The spectra were collected from *m/z* range of 30–350.

## Gas chromatography/mass spectrometry (GC/MS) analysis of fatty acid methyl esters (FAMEs)

Gas chromatography/mass spectrometry (GC/MS) was performed on a coupled Shimadzu GCMS-QP2010S. 1  $\mu$ L of FAME extracts were injected into an HP-5 ms Ultra Inert column (5% phenyl-methylpolysiloxane; 30 m length; 0.25 mm inner diameter; 0.25  $\mu$ m film thickness; Agilent) with a split ratio of 1:20. The extracts were separated using the following temperature program: hold at 100 °C (2 min), ramp at 7 °C min<sup>-1</sup> to 125 °C, ramp at 10 °C min<sup>-1</sup> to 235 °C, ramp at 40 °C min<sup>-1</sup> to 300 °C, hold at 300 °C (2.5 min). The column flow rate was .09 mL min<sup>-1</sup> with an injection temperature of 250 °C. Helium was used as the carrier gas with an inlet pressure of 1.5 bar. The mass spectra were recorded at an ion source and interface temperature of 230 °C and 250 °C, respectively. The spectra were collected from *m/z* range of 30–350.

## Heterologous protein expression and purification

Heterologous proteins were expressed in *E. coli* BAP1 cells. A single colony was inoculated into 10 mL LB with appropriate antibiotic selection for overnight incubation at 37 °C. This culture was used to inoculate 500 mL terrific broth to an initial  $OD_{600}$  of 0.05. This culture was allowed to grow to  $OD_{600} \sim 0.4$ –0.6 before induction with isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG; 0.01 mM). The cultures were grown overnight at 18 °C with shaking. Cells were collected *via* centrifugation, resuspended in binding buffer (20 mM Tris-Cl, 0.5 M NaCl, 20 mM imidazole, 10% glycerol, pH 7.4), and lysed *via* sonication or French press. The cleared lysate was loaded onto a



His-Trap column (GE) and eluted using a gradient to 100% elution buffer (20 mM Tris-Cl, 0.5 M NaCl, 300 mM imidazole, 10% glycerol, pH 7.4). The relevant fractions based on UV and SDS-PAGE gel were combined and concentrated using Amicon filter units (Fig. S4A and Table S2, ESI†). The buffer was exchanged to a storage buffer (20 mM Tris-Cl, 0.5 M NaCl, 5% glycerol, pH 7.4) and stored at  $-80^{\circ}\text{C}$ .

### Colorimetric assay of FAAL domain for chain specificity

The chrome azurol S (CAS)-Fe<sup>3+</sup> solution was prepared as described previously.<sup>33</sup> Loading studies were performed as 100  $\mu\text{L}$  assays in triplicate with error bars showing standard deviation. Samples contained 50 mM Tris HCl, 2.5 mM MgCl<sub>2</sub>, 2.5 mM ATP, 150 mM hydroxylamine, 0.2 mg mL<sup>-1</sup> bovine serum albumin, and 1 mM of specified substrate. To start the reaction, 3.6  $\mu\text{M}$  of the OlsWT FAAL-ACP didomain was added to the assay. For all negative controls, the OlsWT didomain was boiled at  $100^{\circ}\text{C}$  for 10 minutes. Insoluble protein was removed *via* centrifugation and the supernatant was used as the boiled enzyme control. Samples were incubated for 18 hours at  $30^{\circ}\text{C}$  with shaking. The reactions were quenched with 100  $\mu\text{L}$  of the CAS-Fe<sup>3+</sup> solution and incubated at room temperature for 15 minutes. The absorbance of the CAS assay was read on a Tecan plate reader at 630 nm.

### Phylogenetic analysis

The amino acid sequences from identified Ols homologs were downloaded from GenBank. The multiple sequence alignment of the amino acid sequences and CurM was performed using MUSCLE and visualized *via* MEGA11.<sup>47,48</sup> The phylogenetic trees were generated *via* maximum likelihood method and JTT matrix-based model with 500 bootstraps.

### Data availability

Additional data underlying this article are available in the ESI.† This includes additional *in vitro* loading CAS data as well as a list of Ols homologs and protein accession numbers. Nucleotide sequences for all relevant portions of wildtype genomes and mutated strains can also be found in Tables S8–S17 (ESI†).

### Conflicts of interest

The authors declare no competing financial interest.

### Acknowledgements

This work was supported by the National Science Foundation (EFMA-2132036) and in part by the National Institutes for Health. A. Y. G. was a trainee of the NIH Genomic Sciences Training Program (NHGRI 5-T32-HG002760) and supported by an NIH NRSA fellowship (NIGMS 1-F32-GM149146-01).

## Notes and references

- 1 A. Kahn, P. Oliveira, M. Cuau and P. N. Leao, *FEMS Microbiol. Rev.*, 2023, **47**, fuad015.
- 2 D. J. Lea-Smith, M. L. Ortiz-Suarez, T. Lenn, D. J. Nurnberg, L. L. Baers, M. P. Davey, L. Parolini, R. G. Huber, C. A. Cotton, G. Mastroianni, P. Bombelli, P. Ungerer, T. J. Stevens, A. G. Smith, P. J. Bond, C. W. Mullineaux and C. J. Howe, *Plant Physiol.*, 2016, **172**, 1928–1940.
- 3 T. Yamamori, H. Kageyama, Y. Tanaka and T. Takabe, *Lett. Appl. Microbiol.*, 2018, **67**, 299–305.
- 4 B. M. Berla, R. Saha, C. D. Maranas and H. B. Pakrasi, *Sci. Rep.*, 2015, **5**, 14894.
- 5 D. Mendez-Perez, N. A. Herman and B. F. Pfleger, *Appl. Environ. Microbiol.*, 2014, **80**, 6073–6079.
- 6 A. Schirmer, M. A. Rude, X. Li, E. Popova and S. B. del Cardayre, *Science*, 2010, **329**, 559–562.
- 7 H. Parveen and S. S. Yazdani, *J. Ind. Microbiol. Biotechnol.*, 2022, **49**, kuab075.
- 8 D. Mendez-Perez, M. B. Begemann and B. F. Pfleger, *Appl. Environ. Microbiol.*, 2011, **77**, 4264–4267.
- 9 M. Ludwig and D. A. Bryant, *Front. Microbiol.*, 2012, **3**, 354.
- 10 M. S. Donia, W. F. Fricke, J. Ravel and E. W. Schmidt, *PLoS One*, 2011, **6**, e17897.
- 11 P. Arora, A. Goyal, V. T. Natarajan, E. Rajakumara, P. Verma, R. Gupta, M. Yousuf, O. A. Trivedi, D. Mohanty, A. Tyagi, R. Sankaranarayanan and R. S. Gokhale, *Nat. Chem. Biol.*, 2009, **5**, 166–173.
- 12 H. K. D'Ambrosio and E. R. Derbyshire, *ACS Chem. Biol.*, 2020, **15**, 17–27.
- 13 K. Priyadarshan and R. Sankaranarayanan, *J. Indian Inst. Sci.*, 2018, **98**, 261–272.
- 14 T. P. Martins, N. R. Glasser, D. J. Kountz, P. Oliveira, E. P. Balskus and P. N. Leao, *ACS Chem. Biol.*, 2022, **17**, 2528–2537.
- 15 K. Abt, R. Castelo-Branco and P. N. Leao, *J. Nat. Prod.*, 2021, **84**, 278–286.
- 16 S. A. C. Figueiredo, M. Preto, G. Moreira, T. P. Martins, K. Abt, A. Melo, V. M. Vasconcelos and P. N. Leao, *Angew. Chem., Int. Ed.*, 2021, **60**, 10064–10072.
- 17 X. Liu, J. Sheng and R. Curtiss, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6899–6904.
- 18 A. M. Ruffing, *Front. Bioeng. Biotechnol.*, 2014, **2**, 17.
- 19 A. Peramuna and M. L. Summers, *Arch. Microbiol.*, 2014, **196**, 881–890.
- 20 T. Rezanka, J. Lukavsky, L. Siristova and K. Sigler, *Phytochemistry*, 2012, **78**, 147–155.
- 21 D. Kaczmarzyk and M. Fulda, *Plant Physiol.*, 2010, **152**, 1598–1610.
- 22 C. Benning, *Annu. Rev. Cell Dev. Biol.*, 2009, **25**, 71–91.
- 23 R. S. Kalinger, I. P. Pulsifer, S. R. Hepworth and O. Rowland, *Lipids*, 2020, **55**, 435–455.
- 24 S. von Berlepsch, H. H. Kunz, S. Brodesser, P. Fink, K. Marin, U. I. Flugge and M. Gierth, *Plant Physiol.*, 2012, **159**, 606–617.
- 25 J. Beld, R. Abbriano, K. Finzel, M. Hildebrand and M. D. Burkart, *Mol. BioSyst.*, 2016, **12**, 1299–1312.



26 M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato and K. Morishima, *Nucleic Acids Res.*, 2017, **45**(D1), D353–D361.

27 T. Zhu, T. Scalvenzi, N. Sassoon, X. Lu and M. Gugger, *Appl. Environ. Microbiol.*, 2018, **84**, e00425-00418.

28 R. C. Coates, S. Podell, A. Korobeynikov, A. Lapidus, P. Pevzner, D. H. Sherman, E. E. Allen, L. Gerwick and W. H. Gerwick, *PLoS One*, 2014, **9**, e85140.

29 C. M. Jones, S. Parrish and D. R. Nielsen, *ACS Synth. Biol.*, 2021, **10**, 2371–2382.

30 A. L. Markley, M. B. Begemann, R. E. Clarke, G. C. Gordon and B. F. Pfleger, *ACS Synth. Biol.*, 2015, **4**, 595–603.

31 D. C. Cantu, Y. Chen, M. L. Lemons and P. J. Reilly, *Nucleic Acids Res.*, 2011, **39**, D342–D346.

32 O. A. Trivedi, P. Arora, V. Sridharan, R. Tickoo, D. Mohanty and R. S. Gokhale, *Nature*, 2004, **428**, 441–445.

33 L. Kahlert, M. S. Lichstrahl and C. A. Townsend, *ChemBioChem*, 2023, **24**, e202200668.

34 L. Gu, B. Wang, A. Kulkarni, J. J. Gehret, K. R. Lloyd, L. Gerwick, W. H. Gerwick, P. Wipf, K. Håkansson, J. L. Smith and D. H. Sherman, *J. Am. Chem. Soc.*, 2009, **131**, 16033–16035.

35 J. G. McCarthy, E. B. Eisman, S. Kulkarni, L. Gerwick, W. H. Gerwick, P. Wipf, D. H. Sherman and J. L. Smith, *ACS Chem. Biol.*, 2012, **7**, 1994–2003.

36 T. B. Rounge, T. Rohrlack, T. Kristensen and K. S. Jakobsen, *BMC Microbiol.*, 2008, **8**, 141.

37 T. B. Rounge, T. Rohrlack, A. Tooming-Klunderud, T. Kristensen and K. S. Jakobsen, *Appl. Environ. Microbiol.*, 2007, **73**, 7322–7330.

38 G. S. Patil, P. Kinatukara, S. Mondal, S. Shambhavi, K. D. Patel, S. Pramanik, N. Dubey, S. Narasimhan, M. K. Madduri, B. Pal, R. S. Gokhale and R. Sankaranarayanan, *eLife*, 2021, **10**, e70067.

39 J. Mares, J. Hajek, P. Urajova, J. Kopecky and P. Hrouzek, *PLoS One*, 2014, **9**, e111904.

40 Q. Gao, X. Tan and X. Lu, *J. Basic Microbiol.*, 2013, **53**, 848–855.

41 K. Eungrasamee, R. Miao, A. Incharoensakdi, P. Lindblad and S. Jantaro, *Biotechnol. Biofuels*, 2019, **12**, 8.

42 H. Jimbo and H. Wada, *Plant Physiol.*, 2023, **191**, 87–95.

43 K. Kojima, U. Matsumoto, S. Keta, K. Nakahigashi, K. Ikeda, N. Takatani, T. Omata and M. Aichi, *Plant Cell Physiol.*, 2022, **63**, 82–91.

44 S. E. Stevens Jr, C. O. P. Patterson and J. Myers, *J. Phycol.*, 1973, **9**, 427–430.

45 N.-U. Frigaard, Y. Sakuragi and D. A. Bryant, *Methods Mol. Biol.*, 2004, **274**, 325–340.

46 R. M. Lennen, D. J. Braden, R. A. West, J. A. Dumesic and B. F. Pfleger, *Biotechnol. Bioeng.*, 2010, **106**, 193–202.

47 R. C. Edgar, *Nucleic Acids Res.*, 2004, **32**, 1792–1797.

48 K. Tamura, G. Stecher and S. Kumar, *Mol. Biol. Evol.*, 2021, **38**, 3022–3027.

