

## *In Vitro* Reconstitution of the First Steps of Anatoxin-a Biosynthesis in *Oscillatoria* PCC 6506: From Free L-Proline to Acyl Carrier Protein Bound Dehydroproline<sup>†</sup>

Annick Méjean,<sup>‡,§</sup> Stéphane Mann,<sup>‡</sup> Gaëlle Vassiliadis,<sup>‡,§</sup> Bérandère Lombard,<sup>||</sup> Damarys Loew,<sup>||</sup> and Olivier Ploux<sup>\*,‡</sup>

<sup>‡</sup>*Biochimie des micro-organismes, Laboratoire Charles Friedel, UMR CNRS 7223, ENSCP, 11 rue Pierre et Marie Curie, 75231 Paris Cedex 05, France,* <sup>§</sup>*Université Paris Diderot-Paris 7, 75013 Paris, France,* and <sup>||</sup>*Laboratory of Proteomic Mass Spectrometry, Centre de Recherche, Institut Curie, 26 rue d'Ulm, 75248 Paris, France*

Received November 3, 2009; Revised Manuscript Received December 2, 2009

**ABSTRACT:** Anatoxin-a and homoanatoxin-a are two potent cyanobacterial neurotoxins. We recently reported the identification of the gene cluster responsible for the biosynthesis of these toxins in cyanobacteria and proposed a biosynthetic scheme starting from L-proline and involving three polyketide synthases for which the starter would be (S)-1-pyrroline-5-carboxylate bound to an acyl carrier protein, AnaD. We now report the *in vitro* reconstitution of the first steps of this biosynthesis in *Oscillatoria* PCC 6506. We identified in PCC 6506 the gene coding for an Sfp-like phosphopantetheinyl transferase and purified the gene product, OsPPT, that catalyzed the transfer of the phosphopantetheinyl arm to the serine 41 of AnaD. The pure adenylation protein AnaC loaded L-proline on holo-AnaD and was specific for L-proline ( $K_m = 0.97$  mM,  $k_{cat} = 68$  min<sup>-1</sup>) among the 20 natural amino acids. Among six close structural analogues of L-proline, including (S)-1-pyrroline-5-carboxylate, we only found 3,4-dehydro-L-proline to be an alternate substrate for AnaC ( $K_m = 1.5$  mM,  $k_{cat} = 29$  min<sup>-1</sup>). The putative prolyl-AnaD dehydrogenase, AnaB, purified to homogeneity as a histidine-tagged protein, showed an absorption spectrum characteristic of FAD-containing proteins. It oxidized prolyl-AnaD to dehydroprolyl-AnaD as shown by tryptic digestion of the protein followed by liquid chromatography coupled to tandem mass spectrometry. Alignment of the amino acid sequence of this dehydrogenase with related enzymes showed that AnaB belongs to the acyl-CoA dehydrogenase superfamily and thus probably catalyzes an  $\alpha$ - $\beta$ -dehydrogenation of the thioester-bound proline followed by an aza-allylic isomerization to yield (S)-pyrroline-5-carboxyl-AnaD, the proposed starter for the subsequent polyketide synthase, AnaE.

Cyanobacteria are oxygenic photosynthetic prokaryotes that can be found in almost every terrestrial or aquatic environment. This group of bacteria is extremely diverse, and certain species produce a wide range of secondary metabolites, some of which are potent toxins for higher animals (1). Several classes of cyanobacterial toxins have so far been identified including hepatotoxins, neurotoxins, cytotoxins, and irritant toxins. Cases of animal death, due to cyanobacterial toxin exposure, are regularly reported in different places around the world (2–7). For instance, anatoxin-a and homoanatoxin-a, two cyanobacterial neurotoxins, provoke the rapid death of animals by acute asphyxia, when ingested, because these alkaloids are potent agonists of the nicotinic acetylcholine receptor (8). It is now recognized that the release of cyanobacterial toxins in waters and water supplies has major implications for public health and for the environment (9). Therefore, several countries have developed survey programs to assess the risk for humans of cyanobacterial toxin exposure in drinking and recreational waters. In that context, it is important to deepen our knowledge on toxic cyanobacteria and particularly to develop accurate methods to detect these toxic bacteria in the environment.

While a great deal of information is available on hepatotoxins, such as microcystins, and the cyanobacteria that produce these metabolites, our knowledge concerning the cyanobacterial neurotoxins, anatoxin-a, homoanatoxin-a, and saxitoxin, and the species that produce them is limited. For instance, the genomes of two strains of the genus *Microcystis*, that produce microcystins, have been sequenced (10, 11), and the biosynthetic gene cluster responsible for the formation of microcystins has been identified (12, 13). On the other hand, the sequencing of the genome of neurotoxic cyanobacteria has not yet been accomplished, and the biosynthetic genes responsible for the production of saxitoxin, anatoxin-a, and homoanatoxin-a have just been sequenced and annotated (14, 15). These data will certainly be helpful for the detection of cyanobacteria producing these neurotoxins, in water reservoir and in the environment, by using specific PCR-based methodologies.

To contribute to the knowledge on anatoxin-a-producing cyanobacteria, we have initiated a project on this interesting group of bacteria. We have reported the isolation of 15 new filamentous cyanobacteria from dog poisoning sites in the south of France (5) and showed that all of these isolated strains were producing anatoxin-a and homoanatoxin-a and belonged to the *Oscillatoria* genus. More recently, we have reported the identification of the anatoxin-a and homoanatoxin-a biosynthetic gene cluster, called *ana* cluster (15), and showed that a specific DNA sequence could serve as a genetic marker for this class of cyanobacteria (16).

<sup>†</sup>This work was supported in part by Aventis Pharma (Groupe Sanofi-Aventis) and Bayer Pharma, grant "Multiorganismes" for the ENSCP laboratory. The Laboratory of Proteomic Mass Spectrometry is supported by grants from Cancéropôle Ile-de-France and from the Institut National du Cancer, INCa.

\*To whom correspondence should be addressed. Tel/Fax: +(33) 1 44 27 67 01. E-mail: olivier-ploux@enscp.fr.

Annotation of the *ana* gene cluster, feeding experiments as well as preliminary *in vitro* data, led us to propose a biosynthetic route starting from L-proline that involved three modular type I polyketide synthases (PKS)<sup>1</sup> (15). The *ana* cluster is represented in Scheme 1 together with the putative function of the gene products and with the proposed biosynthetic route to anatoxin-a and homoanatoxin-a. There are two main stages in this proposed biosynthesis, first, the steps that allow the formation of the actual starter for the PKS and, second, the reactions that take place on the three PKSs. In the first stage L-proline is thought to be activated by AnaC and then loaded on an acyl carrier protein (ACP), AnaD. In a second step we proposed that AnaB would oxidize the pyrrolidine ring to form a dehydropyrroline, probably (S)-1-pyrroline-5-carboxylate (P5C). In the second stage, that is, the steps that are catalyzed by the PKSs on the tethered substrate, the proposed chemical events were deduced from bioinformatic analysis of the PKS amino acid sequences. Using the colinearity principle, we predicted that the starter should be elongated, cyclized, methylated, and then released and decarboxylated to form either anatoxin-a or homoanatoxin-a. We showed in a preliminary report using *in vitro* experiments (15) that AnaD was an ACP that was transformed in its holo form by the *Bacillus subtilis* Sfp phosphopantetheinyl transferase (PPTase) and that AnaC was capable of activating L-proline for acylation on holo-AnaD. These preliminary data confirmed the putative functions of these proteins predicted from the bioinformatic analysis.

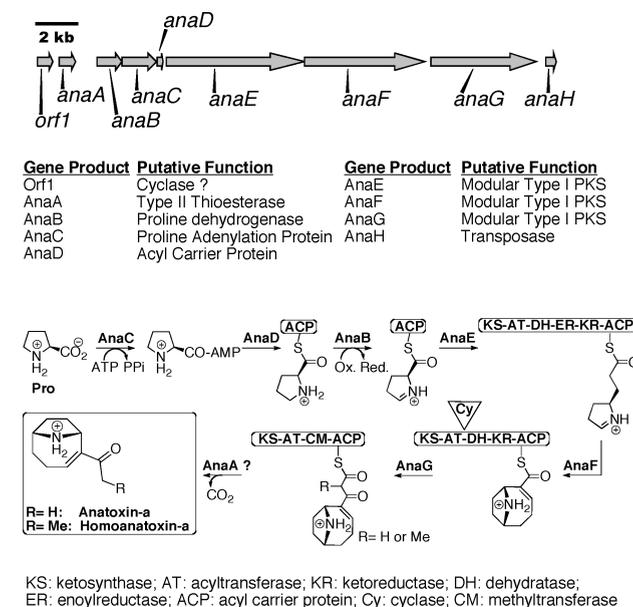
In the present report, we describe the reconstitution, *in vitro*, of the first stage of the biosynthesis of anatoxin-a and homoanatoxin-a in *Oscillatoria* PCC 6506, that is, the transformation of L-proline into AnaD-bound dehydropyrroline. We have identified the gene in *Oscillatoria* PCC 6506 that codes for the PPTase of this organism and showed that the gene product is able to catalyze the transformation of apo-AnaD into holo-AnaD. We have also fully characterized the *in vitro* activity of AnaC, the proline adenylation protein, and we have shown that pure AnaB is indeed a prolyl-AnaD dehydrogenase. Thus, we present experimental evidence that confirms the proposed biosynthetic pathway for the formation of the neurotoxins anatoxin-a and homoanatoxin-a in *Oscillatoria* PCC 6506.

## EXPERIMENTAL PROCEDURES

**General.** Alignments of protein sequences were performed using the basic local alignment search tool (BLAST) supported by the National Center for Biotechnology Information or the Magnifying Genome (MaGe) web servers. Specific alignments

<sup>1</sup>Abbreviations: ACAD, acyl-CoA dehydrogenases; ACP, acyl carrier protein; BLAST, basic local alignment search tool; BsPPT, Sfp phosphopantetheinyl transferase from *Bacillus subtilis*; DCPIP, 2,6-dichlorophenolindophenol; 3,4-Dhp, 3,4-dehydro-L-proline; ESI-MS, electrospray ionization mass spectrometry; HPLC, high-performance liquid chromatography; IPTG, isopropyl  $\beta$ -D-thiogalactoside; LB, Luria-Bertani; LC-MS/MS, liquid chromatography coupled to tandem mass spectrometry; L-ProSNAC, N-acetylcysteamine thioester of L-proline; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; NRP, nonribosomal peptide; NRPS, nonribosomal peptide synthase; oAB, ortho-aminobenzaldehyde; OsPPT, *Oscillatoria* PCC 6506 phosphopantetheinyl transferase; P5C, (S)-1-pyrroline-5-carboxylic acid; PCC, Pasteur culture collection of cyanobacteria; PCP, peptidyl carrier protein; PK, polyketide; PKS, polyketide synthase; PMS, phenazine methosulfate; PPTase, phosphopantetheinyl transferase; ProDh, proline dehydrogenase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis in denaturing conditions; TCEP, tris(2-carboxyethyl)phosphine; TFA, trifluoroacetic acid; TOF-MS, time-of-flight mass spectrometry.

Scheme 1: The *ana* Gene Cluster Responsible for Biosynthesis of Anatoxin-a and Homoanatoxin-a in Cyanobacteria and the Postulated Biosynthetic Route Leading to These Neurotoxins



were performed using the Clustal X software (17). Chemicals and biochemicals were purchased from Sigma-Aldrich. High-performance liquid chromatography (HPLC) was run on a Hewlett-Packard 1050 series apparatus equipped with a Jupiter 5  $\mu$ m C18 300 Å Phenomenex column. Elutions were performed using a linear gradient (from 60% to 80% (v/v) in 20 min) of methanol in aqueous 0.1% trifluoroacetic acid (TFA) at 1 mL/min, and the detector was set at 220 nm. UV-visible spectra and absorbance measurements were recorded on an Uvikon 930 Kontron spectrophotometer. NMR spectra were recorded at the ENSCP NMR laboratory. Protein purifications were achieved using an integrated protein purification system (AKTApurifier; GE Healthcare). PCR and electrophoresis separations (agarose and acrylamide) were run as already described (15). Axenic strain *Oscillatoria* PCC 6506 was obtained from the Pasteur Culture Collection of Cyanobacteria (PCC) and was grown and maintained as already described (18, 19). Protein concentrations were estimated by using the Bradford colorimetric assay (Bio-Rad) (20).

**Cloning, Expression, and Purification of the Recombinant Histidine-Tagged *Oscillatoria* PCC 6506 PPTase (*OsPPT*).** The *Oscillatoria* PCC 6506 PPTase gene (*osppt*) was identified by automatic annotation of the draft genome of PCC 6506 (7 Mbases, 419 contigs) using the MaGe system (21) (A. Méjean, O. Ploux, and C. Médigues, unpublished data). The *osppt* gene was amplified from genomic DNA of PCC 6506 using the *Pfu* DNA polymerase (Promega) and standard conditions, as described by the manufacturer, and the following set of primers (Eurogentec): *osppt*F, 5'-CGGAATTCATGAATGTTTCTGATTGCCTC-3', *Eco*RI site underlined, and *osppt*R, 5'-GCCTCGAGTCAGCTTAAGATATAATTCTCTG-3', *Xho*I site underlined. The amplified DNA, identified as a single band on agarose gel electrophoresis, was purified (PCR Prep; Promega) and digested with the corresponding restriction enzymes (Promega) and ligated (T4 DNA ligase from Promega) into pET28a vector (Novagen) previously cut with the same enzymes. After transformation in *Escherichia coli* JM109 competent cells (CaCl<sub>2</sub> treated) and selection of positive clones,

the plasmid, pET28H6*osppt*, was purified (WizardPlus Miniprep; Promega) and the insert sequenced (GATC company) for verification. The recombinant protein was expressed in *E. coli* BL21(DE3) cells (Novagen) transformed with pET28H6*osppt*. Cells were grown in 1 L of Luria–Bertani (LB) medium at 37 °C under agitation (200 rpm), supplemented with 50 µg/mL kanamycin, and the expression was induced at midlog phase (absorbance at 600 nm = 0.6) with 0.250 µM isopropyl β-D-thiogalactoside (IPTG). The cells were then left overnight at 25 °C with agitation and collected by centrifugation (4000g, at 4 °C for 20 min, Sorval RC5B centrifuge, SLA-3000 rotor). The cell paste was resuspended in 10 mL of 500 mM NaCl and 100 mM sodium phosphate buffer, pH 8.0, and the cells were disrupted by sonication (Misonix sonicator 500 W, 3 mm probe, 4 min at 4 °C with cycles of 15 s sonication followed by cooling periods of 15 s). The resulting suspension was centrifuged (25000g, at 4 °C for 30 min, Sorval RC5B centrifuge, SS-34 rotor), and the supernatant was loaded on a Ni-affinity column (1 mL HisTrap column; GE Healthcare) previously equilibrated with buffer A: 500 mM NaCl and 50 mM sodium phosphate buffer, pH 8.0, containing 20 mM imidazole. After the column was washed (20 mL of buffer A at pH 8.0) the protein was then eluted using a linear gradient of imidazole from 20 mM to 0.5 M in 500 mM NaCl and 100 mM sodium phosphate buffer, pH 8.0, during 20 min at 1 mL/min. The collected fractions were analyzed by polyacrylamide gel electrophoresis under denaturing conditions (SDS–PAGE), and the fractions containing the pure protein were pooled. The protein solution was desalted on a PD10 column (GE Healthcare) using 50 mM Tris-HCl buffer, pH 7.7, and 100 mM NaCl, as eluent. After the protein was concentrated using a concentrator device (Amicon Ultra-15 centrifugal filter unit 10k), sorbitol was added to a final concentration of 7% (w/v), and the pure protein was stored at –20 °C.

**PPTase Activity Assay.** Histidine-tagged recombinant apo-AnaD was produced and purified as already described (15). The histidine-tagged Sfp PPTase from *B. subtilis* (BsPPT) was obtained as already described (15), using an overproducing strain kindly provided by Peter F. Leadlay and S. Dickens (Department of Biochemistry, University of Cambridge; strain number PFL-C1043). A typical PPTase assay (100 µL total volume) consisted of 100 mM Tris-HCl, pH 7.7, 10 mM MgCl<sub>2</sub>, 5 mM tris-(2-carboxyethyl)phosphine (TCEP), 50 µM CoASH, 0.5 µM PPTase, either the BsPPT or the OsPPT, and 50 µM apo-AnaD. The reaction was incubated at 28 °C and stopped by the addition of 150 µL of a stop solution: 50% (v/v) methanol in 0.1% (v/v) aqueous TFA. The solution was kept at 0 °C prior to analysis by HPLC. The apo-AnaD (retention time = 14.2 min) and holo-AnaD (retention time = 14.8 min) were separated and quantified by HPLC. The holo-AnaD production, in this assay, was linear with time, from 0 to 15 min, and linear with PPTase concentration, from 0 to 0.15 µM. The effect of pH on the OsPPT activity was examined using the above assay using a complex buffer mixture (0.1 M acetic acid, 0.1 M 2-(*N*-morpholino)ethanesulfonic acid, 0.2 M Tris (22)) set at different pHs from 4.0 to 9.0. Large quantities of holo-AnaD were prepared as follows: the assay (1.5–5 mL final volume) consisted of 100 mM Tris-HCl, pH 7.7, 10 mM MgCl<sub>2</sub>, 5 mM TCEP, 60 µM CoASH, 2 µM PPTase, either the BsPPT or the OsPPT, and 50 µM apo-AnaD. The reaction mixture was incubated for 2 h at 20 °C, and the holo-AnaD thus formed was desalted on a PD-10 column (GE Healthcare) using 100 mM Tris-HCl, pH 7.7, and 1 mM TCEP as eluent. The protein was then concentrated by ultrafiltration

(Amicon Ultra-15 centrifugal filter unit 10k), and sorbitol was added to the solution up to 7% (w/v) before storage at –20 °C.

**Determination of AnaC Kinetic Parameters and Specificity.** Histidine-tagged AnaC was produced and purified as already described (15). The kinetic parameters and specificity of AnaC were determined using the PP<sub>i</sub>–ATP exchange reaction as already described (15). Briefly, the assay (100 µL total volume) consisted of 100 mM Tris-HCl, pH 7.7, 5 mM TCEP, 10 mM MgCl<sub>2</sub>, 5 mM ATP, 5 mM substrate, 1 mM [<sup>32</sup>P]pyrophosphate (specific radioactivity 1.2 Ci/mol; Amersham), and 0.2 µM AnaC. The reactions were initiated by the addition of the enzyme and incubated at 28 °C for 30 min. The reactions were quenched by the addition of 1 mL of stop solution: 1.6% (w/v) activated charcoal, 4.5% (w/v) tetrasodium pyrophosphate, and 3.5% (v/v) perchloric acid in water. The samples were filtered on Whatman paper filters and rinsed three times with 15 mL of water. After drying, the filters were counted using 6 mL of Picofluor scintillation counting liquid on a scintillation counter (Beckman LS6500; Beckman Coulter). Each reaction was run in triplicate, and background counting (reaction without enzyme) was subtracted. The kinetic parameters for L-proline and 3,4-dehydro-L-proline (3,4-Dhp) were determined using the standard assay but using various concentrations of the substrate (from 0 to 6 mM). The data points were fitted to the Michaelis–Menten equation using nonlinear regression analysis (Kaleidagraph software).

**Preparation of Prolyl-AnaD.** The loading of L-proline on the holo-AnaD was achieved as already described (15). The assay (100 µL total volume) consisted of 100 mM Tris-HCl, pH 7.7, 5 mM TCEP, 10 mM MgCl<sub>2</sub>, 5 mM ATP, 5 mM L-proline or [U-<sup>13</sup>C, <sup>15</sup>N]-L-proline (Cambridge Isotope), 50 µM holo-AnaD, and 0.2 µM AnaC, and the reaction was incubated at 28 °C for 30 min. For preparation of large quantities of prolyl-AnaC the assay was adapted to 2 mL total volume, and the concentration of AnaC was 2 µM. The prolyl-AnaD was then desalted on a PD-10 column (GE Healthcare) using 100 mM Tris-HCl, pH 7.0, as eluent. The protein was then concentrated by ultrafiltration (Amicon Ultra-15 centrifugal filter unit 10k) and used immediately, or the solution was supplemented with 7% (w/v) sorbitol before storage at –20 °C. The purified prolyl-AnaD eluted at 13.8 min using the HPLC conditions described above.

**Synthesis of (*S*)-1-Pyrroline-5-carboxylic Acid (P5C).** P5C was prepared as described by Gerratana et al. (23, 24) by the deprotection in acidic conditions of (2*S*)-*tert*-butyl-*N*-Boc-5-hydroxyproline. P5C was then purified by ion-exchange chromatography and assayed using *o*-aminobenzaldehyde (oAB) as already described (25, 26). Pure fractions of P5C were concentrated to dryness under vacuum, solubilized in 0.1 M HCl to a final concentration of 132 mM, and stored at 4 °C.

**Synthesis of the *N*-Acetylcysteamine Thioester of L-Proline (L-ProSNAC).** L-ProSNAC was prepared as described by Ehmann et al. (27).

**Cloning, Expression, and Purification of the Recombinant Histidine-Tagged *Oscillatoria* PCC 6506 AnaB.** The *anaB* gene (15) was PCR-amplified (GoTaq polymerase; Promega) from genomic DNA from PCC 6506 using the following primers: *anaBF*, 5'-CGCGCGAATTCGATTTTG-CCTGGAACAGTC-3', *EcoRI* site underlined, and *anaBR*, 5'-CGGCCTCGAGTTACAATCCCAAGAATTTA-3', *XhoI* site underlined. The PCR product was purified and cloned into a pET28a vector as described above for the *osppt* gene. The recombinant AnaB histidine-tagged protein was produced in *E. coli* BL21-CodonPlus (DE3)-RIPL strain (Stratagene).

Cells were grown in LB medium (1 L) at 37 °C under agitation (250 rpm) in the presence of 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. When the culture reached an absorbance of 0.6 at 600 nm, the medium was cooled to 25 °C, and the induction was initiated with 0.2 mM IPTG. The agitation was maintained overnight at 25 °C. Cells were harvested by centrifugation (4000g at 4 °C for 20 min; SLA-3000 rotor) and resuspended in 5 mL of 20 mM Tris-HCl buffer, pH 7.4, containing 0.5 M NaCl, 20 mM imidazole, 10% (v/v) glycerol, and 0.1 mM FAD. Lysis was performed by sonication at 4 °C for 3 min with 10 s pulses spaced by 20 s cooling periods. After centrifugation (two runs at 26000g, SS-34 rotor, 15 min) the supernatant was loaded on a 1 mL HisTrap column (GE Healthcare). The desired protein was eluted with a linear gradient of imidazole (20–500 mM) in Tris-HCl buffer, pH 8.0, containing 0.5 M NaCl and 10% (v/v) glycerol. Pure fractions of AnaB were pooled, and the protein was desalted and concentrated by ultrafiltration (Amicon Ultra-15 centrifugal filter unit 30k) and stored in 20 mM Tris-HCl buffer, pH 7.4, and 20% (v/v) glycerol. The overall yield of production of AnaB was 5 mg/L of culture. The UV-visible absorption spectrum of pure AnaB was recorded on a pure sample of AnaB at 0.56 mg/mL (12 µM) in 20 mM Tris-HCl, pH 7.4, containing 10% (v/v) glycerol.

**Alternate Substrates for AnaB.** Two assays were used to detect any AnaB activity using free L-proline, L-proline methyl ester, or L-ProSNAC as alternate substrates. 2,6-Dichlorophenolindophenol (DCPIP) assay (28): the assay consisted of 50 mM Tris-HCl buffer, pH 7.5, 50 µM DCPIP, 100 µM FAD, 1 mM phenazine methosulfate (PMS), 1–5 mM substrate, and 1 µM AnaB. The solution was incubated at 30 °C, and the absorbance at 610 nm was monitored against time. oAB assay (29): the assay consisted of 50 mM Tris-HCl buffer, pH 7.5, 1–5 mM substrate, and 1 µM AnaB in a total volume of 210 µL. After various incubation times (0, 5, 10, 20, 60 min) at 30 °C, 200 µL of 10% (v/v) aqueous trichloroacetic acid and 40 µL of 0.1 M oAB in 40% (v/v) aqueous ethanol were added to the assay solution, and the absorbance at 440 nm was measured.

**Oxidation Reaction Catalyzed by AnaB.** Prolyl-AnaD (25 µM), prepared as described above, was incubated in the presence of pure AnaB (2 µM) in 100 mM Tris-HCl buffer, pH 7.0, in the presence of PMS (1 mM) or in the absence of any electron acceptor except naturally present oxygen. The oxidation reaction was incubated at 28 °C for 30–60 min. The reaction mixture was analyzed using HPLC as described above and mass spectrometry as described below. A control experiment was run in parallel in the absence of AnaB.

**Mass Spectrometry Analyses.** OsPPT, AnaC, and AnaB samples were desalted using ZipTipC4 pipet tips prior to analysis by mass spectrometry. The MALDI-TOF mass spectra were recorded on an ABI 4800 TOF/TOF mass spectrometer (Applied Biosystems) in positive ion linear mode by using a saturated sinapinic acid matrix (50% (v/v) acetonitrile/water, 0.1% (v/v) TFA). The AnaD samples, apo-, holo-, and proly-AnaD, were first desalted by reversed-phase HPLC under the conditions described above. The samples were concentrated in vacuum (SpeedVac system) and analyzed on a QSTAR Elite quadrupole time-of-flight (Q-TOF) mass spectrometer (Applied Biosystems/MDS Sciex) equipped with an external nano-electrospray ionization (nanoES) ion source. NanoES was performed using a coated medium borosilicate capillary (Proxeon Biosystems A/S, Denmark). The electric field required to spray the solution was established by applying a voltage of about 950 V. The solution flow rate was estimated to be between 10 and 40 nL/min as

described by Wilm and Mann (30). Analyst QS software 2.0 (Applied Biosystems, MDS Sciex) was used for the spectrum acquisition and the data analysis. Prolyl-AnaD and its oxidation product formed after incubation in the presence of AnaB, as described above, were analyzed as follows. Two hundred picomoles of the HPLC-purified prolyl-AnaD and of the oxidized product of prolyl-AnaD, prepared as described above, was dried under vacuum (SpeedVac system) and then resuspended by sonication in 100 µL of a solution of 80% (v/v) acetonitrile/25 mM ammonium bicarbonate. Trypsin (Sigma, T6567) digestion was then performed by the addition of 10 µL of a trypsin solution (200 ng/mL) to the suspension, and the samples were incubated on a rotator for 6 h at 37 °C. The trypsin-generated mixture was dried under vacuum and suspended in 20 µL of solvent A (5% (v/v) acetonitrile/water, 0.1% (v/v) formic acid), and 1 µL of the solution was analyzed using an actively split capillary HPLC system (Ultimate 3000; Dionex, Germering, Germany) connected to the Q-TOF mass spectrometer. The digest was loaded onto a C18 PepMap guard column (0.3 mm i.d. × 5 mm long, 5 µm particle size, 100 Å pore size; Dionex S.A.) in solvent A. The sample was washed on the column for 3 min with solvent A prior to separation using a LC Packings PepMap C18 column (75 µm i.d. × 150 mm long, packed with 3 µm particles with 100 Å pore size; Dionex S.A.) using a 60 min linear gradient from 5% to 50% (v/v) of solvent B (80% (v/v) acetonitrile/water, 0.085% (v/v) formic acid) at 200 nL/min. Data acquisition was performed using the Analyst QS Software, set for the positive ion mode with a spray voltage of 2.3 kV. A TOF-MS survey scan was acquired for 1 s over a mass range of  $m/z$  700–800 and MS/MS range of  $m/z$  65–1700. An information-dependent acquisition method was used to acquire product ion scans on the three most intense ions per cycle over a mass range of  $m/z$  65–1700, excluding previously gated ions for 60 s. A Smart setting of 20.0 was used.

## RESULTS

**Identification and Functional Characterization of the *Oscillatoria* PCC 6506 PPTase.** We have undertaken the sequencing of the genome of the neurotoxic cyanobacteria *Oscillatoria* PPC 6506, and a draft genome has already been obtained consisting of 7 Mbases on 419 contigs. This draft genome has been automatically annotated using the MaGe system (A. Méjean, O. Ploux, and C. Médigue, unpublished data), and among the annotated genes we found only one putative PPTase gene that we called *osppt*. Alignment of the sequence of the gene product OsPPT with PPTase sequences that gave the highest BLAST scores (Supporting Information Figure S1 and Table S1) showed that this enzyme belonged to the Sfp-like PPTase family (31, 32) and more specifically to the W/KEA cyanobacterial PPTase subfamily as defined by Neilan and co-workers (33). Indeed, we found the following consensus sequences that are conserved in this class of enzymes: F<sup>87</sup>, GKP<sup>95</sup>, FNX(S/A)HS<sup>113</sup>, (I/L)G(V/D)<sup>130</sup>, E<sup>152</sup>, and F(F/Y)-X<sub>2</sub>WXKEAX<sub>2</sub>K<sup>180</sup> (the numbering refers to the native OsPPT sequence). The closest PPTase, based on BLAST searches, was the enzyme from *Trichodesmium erythraeum* IMS101 (46% identity, 64% similarity, Supporting Information Table S1), another cyanobacterium whose genome has been completely sequenced.

We have cloned the *osppt* gene into an expression vector, and we have purified the N-terminal histidine-tagged fusion protein

Table 1: Mass Spectrometry Analysis of the Purified Proteins Studied in This Report

| protein  | calcd average mass (Da) <sup>a</sup> |               | obsd mass (Da)             |
|--|--------------------------------------|---------------|----------------------------|
|  | intact protein                       | protein – Met |                            |
| OsPPT <sup>b</sup>   | 32851.38                             | 32720.18      | 32720 (M + H) <sup>+</sup> |
| AnaC <sup>b</sup>  | 63693.03                             | 63561.83      | 63513 (M + H) <sup>+</sup> |
| AnaB <sup>b</sup>  | 46548.44                             | 46417.24      | 46519 (M + H) <sup>+</sup> |
| apo-AnaD <sup>c</sup>  | 13941.03                             | 13809.82      | 13809.6                    |
| holo-AnaD <sup>c</sup>   | 14281.36                             | 14150.16      | 14150.2                    |
| prolyl-AnaD <sup>c</sup>                                       | 14378.48                             | 14247.28      | 14247.2                    |
| [U- <sup>13</sup> C, <sup>15</sup> N]-prolyl-AnaD <sup>c</sup> | 14384.40                             | 14253.20      | 14252.8                    |

<sup>a</sup>The calculated masses correspond to the intact recombinant protein (left) or to the protein lacking the N-terminal Met (right). <sup>b</sup>Mass measured by MALDI-TOF. <sup>c</sup>Mass measured by ESI-MS after deconvolution. Minor species, acetylated M + 42, and gluconylated M + 178 (37) were also observed.

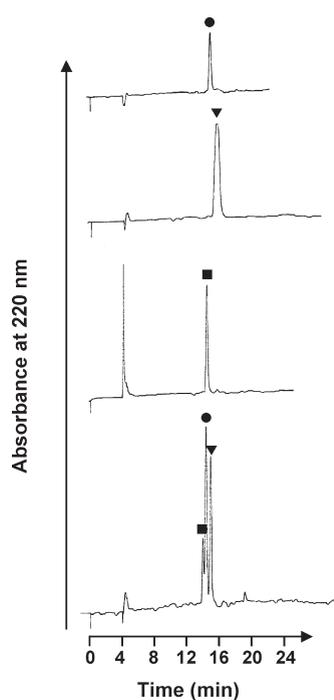


FIGURE 1: HPLC traces of the three different forms of the acyl carrier protein AnaD. From top to bottom: apo-AnaD (●), holo-AnaD (▼), prolyl-AnaD (■), and the mixture of the three samples. The HPLC analyses were run as described in the Experimental Procedures section. Approximately 2.5 nmol (35 μg) of protein samples was injected.

to homogeneity by affinity chromatography. It showed one band by SDS-PAGE and one peak by reversed-phase HPLC and showed an average mass consistent with that predicted (Table 1). The PPTase activity of this new enzyme was probed using an HPLC-based assay and apo-AnaD as the substrate because we could readily separate the apo-AnaD and holo-AnaD forms using this technique (Figure 1). The OsPPT readily transformed apo-AnaD into holo-AnaD at saturating concentrations of substrates, that is, 50 μM apo-AnaD and 50 μM CoASH. The specific activity measured under these conditions was 0.21 μmol min<sup>-1</sup> mg<sup>-1</sup> (apparent  $k_{\text{cat}} = 6.0 \text{ min}^{-1}$ ). The Sfp PPTase from *B. subtilis* was slightly less effective under the same conditions for the transformation of apo-AnaD into holo-AnaD: specific activity of 0.13 μmol min<sup>-1</sup> mg<sup>-1</sup> (apparent  $k_{\text{cat}} = 3.3 \text{ min}^{-1}$ ).

Table 2: AnaC Specificity toward Its Amino Acid Substrate As Probed by the PP<sub>i</sub>-ATP Exchange Assay

| amino acid substrate | relative activity (%) | amino acid substrate | relative activity (%) |
|----------------------|-----------------------|----------------------|-----------------------|
| Ala                  | 0.6 ± 0.1             | Leu                  | 1.8 ± 0.2             |
| Arg                  | 0.2 ± 0.1             | Lys                  | 0.4 ± 0.1             |
| Asn                  | 0.0 ± 0.1             | Met                  | 0.7 ± 0.1             |
| Asp                  | 0.0 ± 0.1             | Phe                  | 0.3 ± 0.1             |
| Cys                  | 0.5 ± 0.1             | Pro                  | 100.0 ± 0.3           |
| Gln                  | 0.0 ± 0.1             | Ser                  | 0.3 ± 0.1             |
| Glu                  | 0.0 ± 0.1             | Thr                  | 0.3 ± 0.1             |
| Gly                  | 0.1 ± 0.1             | Trp                  | 0.7 ± 0.1             |
| His                  | 0.2 ± 0.1             | Tyr                  | 0.9 ± 0.1             |
| Ile                  | 0.6 ± 0.1             | Val                  | 0.4 ± 0.1             |

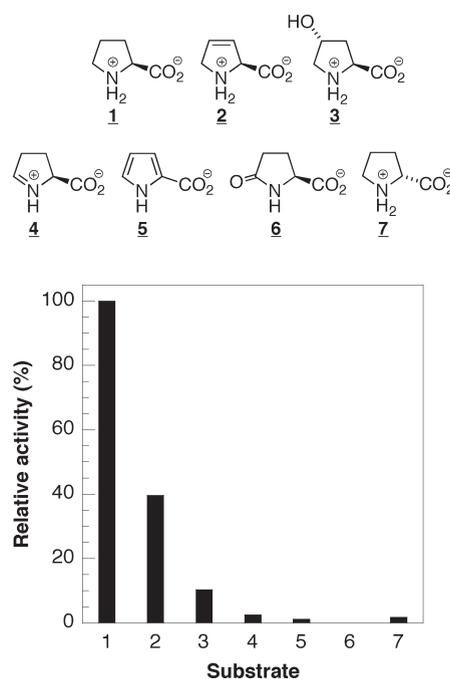


FIGURE 2: Specificity of AnaC-catalyzed PP<sub>i</sub>-ATP exchange toward proline analogues. The assay was run in the presence of 5 mM proline analogue, and the relative initial rates were determined. Each data point represents the average of three independent experiments (standard errors were less than 0.2%). Compounds: 1, L-proline; 2, 3,4-dehydro-L-proline (3,4-Dhp); 3, *trans*-4-hydroxy-L-proline; 4, (*S*)-1-pyrroline-5-carboxylic acid (P5C); 5, pyrrole-2-carboxylic acid; 6, L-pyrroglutamic acid; 7, D-proline.

The pH profile of the specific activity of OsPPT, measured under saturating conditions, was bell-shaped with an optimum pH at 7.0 as observed for similar enzymes (34–36).

**Kinetics and Specificity of the Adenylation Protein AnaC.** We previously showed that the adenylation protein AnaC was capable of activating L-proline for acylation of holo-AnaD to yield prolyl-AnaD (15). Using the classical PP<sub>i</sub>-ATP exchange assay, we tested the ability of this enzyme to discriminate among the 20 proteinogenic amino acids. Table 2 shows that AnaC is very specific for L-proline and does not activate the other 19 amino acids at a reasonable rate. However, slow activation was observed for Leu (1.8% of relative activity). We also tested six close proline analogues for activation by AnaC as shown in Figure 2. We observed a sharp specificity with only 3,4-dehydro-L-proline, 2, being an alternate substrate (relative activity = 39.6 ± 0.2%) and to a lesser extent *trans*-4-hydroxy-L-proline,

**3** (relative activity =  $10.4 \pm 0.2\%$ ). A similar specificity profile among these three compounds, **1**, **2**, and **3**, was observed in the case of the homologous CloN4 and CouN4 enzymes (38), that catalyze the activation of proline in the biosynthesis of clorobiocin and coumermycin A, respectively. The four other compounds tested were very poor to null substrate for AnaC. P5C, **4** (relative activity =  $2.5 \pm 0.1\%$ ), the putative structure formed on AnaD after AnaB oxidation, is a very poor substrate. The aromatic pyrrole-2-carboxylate, **5**, and the pyroglutamate, **6**, were not better substrates for AnaC (relative activities =  $1.1 \pm 0.1\%$  and  $0.0 \pm 0.1\%$ , respectively). Interestingly, D-proline, the enantiomer of the substrate, is not activated by AnaC (relative activity =  $1.7 \pm 0.1\%$ ), and thus this adenylation protein is rather enantioselective, unlike other homologous adenylation domains or proteins (39–41).

Using the  $PP_i$ -ATP exchange reaction, we have also determined the kinetic parameters for activation of L-proline and 3,4-Dhp (Figure 3):  $K_m(\text{Pro}) = 0.97 \pm 0.17$  mM;  $k_{\text{cat}}(\text{Pro}) = 68.5 \pm 3.8$  min<sup>-1</sup>;  $k_{\text{cat}}/K_m(\text{Pro}) = 1\,177$  M<sup>-1</sup> s<sup>-1</sup>;  $K_m(3,4\text{-Dhp}) = 1.54 \pm 0.29$  mM;  $k_{\text{cat}}(3,4\text{-Dhp}) = 29.4 \pm 2.1$  min<sup>-1</sup>;  $k_{\text{cat}}/K_m(3,4\text{-Dhp}) = 318$  M<sup>-1</sup> s<sup>-1</sup>. Thus, the discrimination of Pro versus 3,4-Dhp is largely determined by the catalytic constant. The kinetic parameters determined for L-proline activation by AnaB were in close agreement with those obtained for the homologous CloN4 and CouN4 enzymes (38).

Using the HPLC-based assay (Figure 1), we showed that AnaC readily transformed holo-AnaD into prolyl-AnaD. The conversion of 50 μM holo-AnaD was completed in 15–30 min in

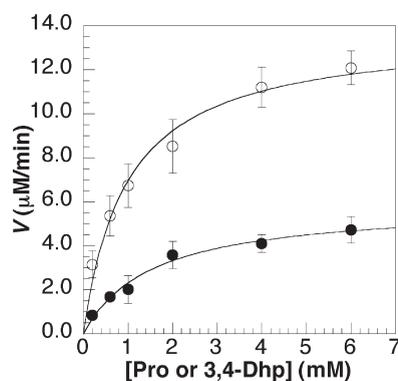


FIGURE 3: Determination of the kinetic parameters of AnaC for L-proline (○) and 3,4-Dhp (●) using the  $PP_i$ -ATP exchange assay. The assay was run in the presence of various concentrations of the substrate, and the initial rates were determined. Each data point represents the average of three independent experiments. The data points were fitted to the Michaelis–Menten equation using nonlinear regression analysis.

the presence of 0.2 μM AnaC. The mass spectrometry data were in perfect agreement with the formation of a prolyl thioester on the phosphopantetheinyl arm of the holo-AnaD using either unlabeled proline or [<sup>13</sup>C, <sup>15</sup>N]-L-proline (Table 1). This prolyl-AnaD was rather unstable at pH 7.0 with decomposition into the holo-AnaD form. It is not clear at the moment what reaction takes place, but preliminary data (S. Mann and O. Ploux, unpublished data) showed that the thioester bond is hydrolyzed to form the thiol and free proline. Thus, the prolyl-AnaD could only be kept for 1 or 2 days before its use (estimated half-life of 2 h in buffered solution at pH 7.0 and 28 °C).

Given the fact that AnaC is very specific for L-proline, we have aligned the amino acid sequence of this adenylation protein with those of nine other adenylation proteins that activate proline and that are involved in the biosynthesis of pyrrole containing secondary metabolites: pyoluteorin (42), coumermycin (43), prodigine (44), clorobiocin (45), prodigiosin (46), pyrrolomycin and dioxapyrrolomycin (49), Dkxanthene (50), and indanomycin (51). As expected, these enzymes are closely related with sequence identity to the AnaC sequence ranging from 31% to 40% (Supporting Information Table S2). Because several authors have proposed a code to predict the specificity of any adenylation protein or domain (40, 52), based on its amino acid sequence and on the three-dimensional structure of GrsA, we have aligned the GrsA and AnaC sequences together with the nine proline adenylation protein sequences quoted above (42–51). Inspecting the alignment of the ten proline-specific adenylation proteins (Supporting Information Figure S2) showed that the residues corresponding to the GrsA residues implicated in the amino acid substrate binding pocket were almost strictly conserved within the proline-specific enzymes (Table 3). However, the codes previously proposed for the prediction of adenylation of proline by adenylation proteins were not fully respected in AnaC-related proteins. Besides Asp<sup>235</sup> and Lys<sup>517</sup> (numbering of GrsA) which cannot be used for the prediction because they are almost strictly conserved in all adenylation proteins, we only found three positions out of eight that conformed to the previous codes.

*Production and Characterization of the Prolyl-AnaD Dehydrogenase, AnaB.* In the cluster of genes responsible for the biosynthesis of anatoxin-a and homoanatoxin-a we identified a putative prolyl-AnaD dehydrogenase coded by the gene *anaB* (Scheme 1, Supporting Information Figure S3 and Table S3). We speculated that this enzyme would oxidize the prolyl-AnaD to the (S)-pyrroline-5-carboxyl-AnaD (P5C-AnaD) based on similar biosynthetic schemes (42–51) and on reasonable chemistry that could lead to anatoxin-a and homoanatoxin-a (15). We have cloned the *anaB* gene in a suitable expression vector to produce a

Table 3: A Plausible Code for Predicting Proline-Specific Adenylation Proteins

| sequence/consensus                            | amino acid found on the selected positions |                  |                  |                  |                  |                         |                  |                         |                  |                         |
|---|--|------------------|------------------|------------------|------------------|-------------------------|------------------|-------------------------|------------------|-------------------------|
| GrsA <sup>a</sup>                             | D <sup>235</sup>                           | A <sup>236</sup> | W <sup>239</sup> | T <sup>278</sup> | I <sup>299</sup> | A <sup>301</sup>        | A <sup>322</sup> | I <sup>330</sup>        | C <sup>331</sup> | K <sup>517</sup>        |
| code for Pro <sup>b</sup>                     | <b>D</b>                                   | V                | Q                | L                | I                | <b>A</b>                | H                | V                       | V                | <b>K</b>                |
| code for Pro <sup>c</sup>                     | <b>D</b>                                   | V                | Q                | S/Y/V/F          | I/A              | <b>A</b>                | H                | V                       | –                | <b>K</b>                |
| AnaC <sup>d</sup>                             | <b>D</b> <sup>219</sup>                    | L <sup>220</sup> | F <sup>223</sup> | Y <sup>262</sup> | L <sup>289</sup> | <b>A</b> <sup>291</sup> | L <sup>315</sup> | <b>V</b> <sup>323</sup> | C <sup>324</sup> | <b>K</b> <sup>527</sup> |
| consensus found in present study <sup>e</sup> | <b>D</b>                                   | L                | F/L              | Y                | L/I              | <b>A</b>                | L/W              | V                       | C                | <b>K</b>                |

<sup>a</sup>These ten residues constitute the Phe binding site in GrsA, as shown by Marahiel et al. (39, 40). GrsA numbering. <sup>b</sup>As defined by Marahiel et al. (40). <sup>c</sup>As defined by Townsend et al. (52). <sup>d</sup>Residues of AnaC corresponding to the ten residues of GrsA implicated in the amino acid binding site, as determined by sequence alignment. AnaC numbering. <sup>e</sup>Residues found in the sequence of the nine proline-specific adenylation proteins under study, at the selected positions thought to be involved in proline binding, as determined by multiple alignment. Strictly conserved residues in the different codes are in bold type.

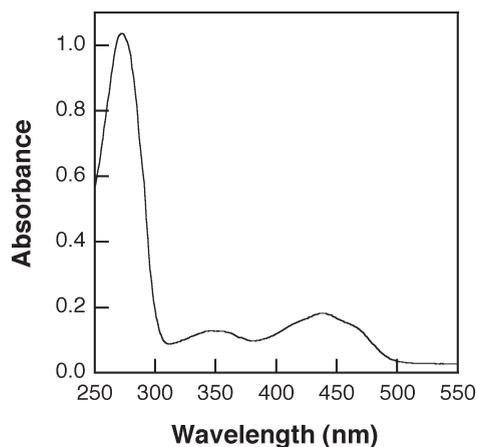


FIGURE 4: UV-visible absorption spectrum of pure AnaB at 0.56 mg/mL (12  $\mu$ M) in 20 mM Tris-HCl, pH 7.4, containing 10% (v/v) glycerol.

recombinant N-terminal histidine-tagged protein. The enzyme was readily produced in the *E. coli* BL21-CodonPlus (DE3)-RIPL strain and purified to homogeneity by affinity chromatography. The pure enzyme, as judged by SDS-PAGE and HPLC, had a mass in agreement with that predicted from the sequence (Table 1). The enzyme exhibited an electronic absorption spectrum in the UV-visible range with bands at 272, 345, and 438 nm and shoulders at 407 and 466 nm (Figure 4). The ratio of absorbance at 438 nm and at 272 nm was 0.175. This spectrum is characteristic of FAD-dependent enzymes and is consistent with the fact that a BLAST search with the protein sequence of AnaB showed that this enzyme belonged to the acyl-CoA dehydrogenase (ACAD) superfamily (Supporting Information Figure S3 and Table S3).

To detect the AnaB-catalyzed oxidation product, we incubated prolyl-AnaD (25  $\mu$ M) in the presence of pure AnaB (2.5  $\mu$ M) at pH 7.0 using either PMS or oxygen as the electron acceptor. Unfortunately, the HPLC traces did not reveal the presence of a new product, either because the substrate and the product coeluted or because the reaction yield was too low. We thus decided to use mass spectrometry to detect this oxidation product. Because a loss of 2 Da over 14247 Da, the expected decrease in mass after oxidation of prolyl-AnaD by AnaB would be impossible to detect as already noted by others (38), we decided to detect the reaction product by digesting the protein and by analyzing the fragments using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). When the substrate of the reaction, prolyl-AnaD, was analyzed using this strategy (Figure 5A), we detected the tryptic fragment **F1** containing the conserved serine at position 41 of the apo-AnaD (native sequence; see Supporting Information Figure S4 and Table S4), with the phosphopantetheinyl arm attached to it. Fragmentation of this ion, **F1**, was observed, and we could attribute the structure of the different major fragments formed, **F2** to **F6** (Figure 5A and Table 4). Fragment **F1** is the intact tryptic fragment which gives fragments **F2** and **F3** with the modified serine at position 41 as a dehydroalanine or a phosphoserine, respectively. Fragment **F4** corresponds to the entire phosphopantetheinyl arm with the prolyl thioester attached to it. This fragment gives rise to fragment **F5** by the loss of  $\text{H}_3\text{PO}_4$ , which in turn loses the prolyl moiety to yield fragment **F6**. The sequence of the tryptic fragment together with the localization of the pantothetheinyl arm was confirmed by

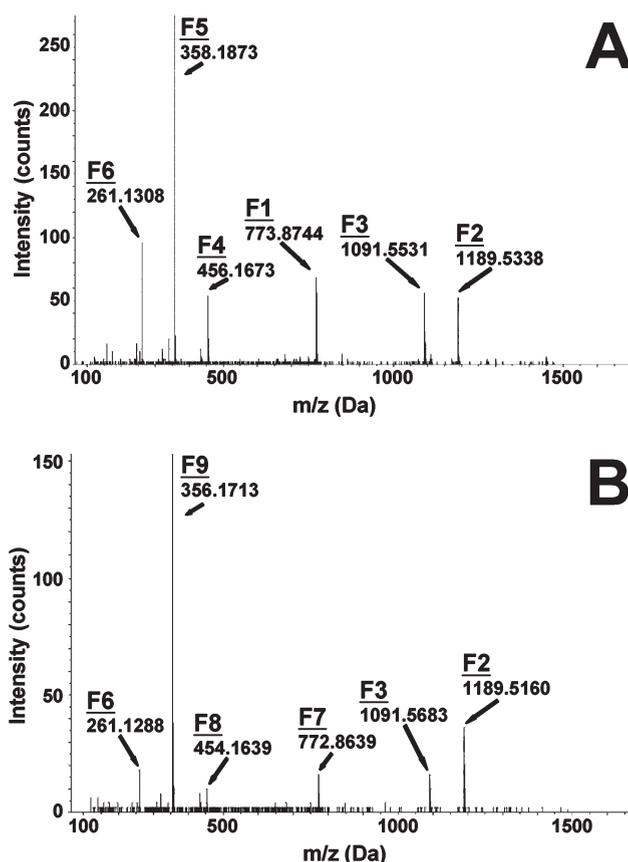
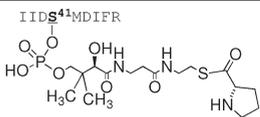
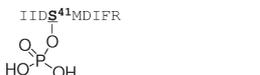
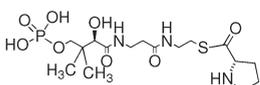
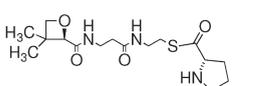
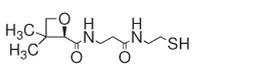
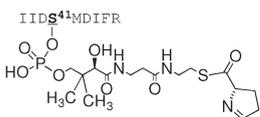
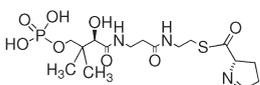
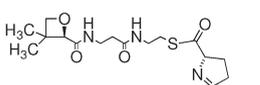


FIGURE 5: Mass spectrometry analysis by LC-MS/MS of the reaction catalyzed by AnaB. The spectra were recorded after tryptic digestion and liquid chromatography separation of either prolyl-AnaD before oxidation (A) or after AnaB-catalyzed oxidation (B). (A) Mass spectrum (the spectrum was zoomed twice along the y axis for clarity) of the prolyl-AnaD tryptic peptide, from positions 38 to 46 of AnaD that contains the phosphopantetheinyl arm. The molecular ion is symbolized as **F1**  $m/z$  773.8744 for  $(M + 2\text{H})^{2+}$ . (B) Mass spectrum (the spectrum was zoomed twice along the y axis for clarity) of the tryptic peptide, from positions 38 to 46 of AnaD that contains the phosphopantetheinyl arm, of the AnaB-catalyzed oxidation product. The molecular ion is symbolized as **F7**  $m/z$  772.8639 for  $(M + 2\text{H})^{2+}$ . See Table 4 and text for full interpretation of the fragmentations.

interpretation of the y and b fragments formed (Supporting Information Figure S5).

After oxidation of prolyl-AnaD catalyzed by AnaB in the presence of oxygen, the reaction product was analyzed using the same strategy (Figure 5B). We clearly observed fragments **F2**, **F3**, **F7**, **F8**, and **F9**. The tryptic fragment **F7** had lost 2 Da when compared to the corresponding fragment **F1**. The same reduction of 2 Da was observed on fragments **F8** and **F9**, corresponding to fragments **F4** and **F5**, respectively, of the mass spectrum shown in Figure 5A but not on fragments **F2**, **F3**, and **F6** that were present in the two mass spectra. No oxidation product was observed in the control experiment without AnaB. This unambiguously shows that the two-electron oxidation took place on the proline moiety, as expected. Thus AnaB is a prolyl-AnaD dehydrogenase (Scheme 1). The oxidation reaction was not complete because in the oxidation mixture we observed the presence of the tryptic fragment that came from prolyl-AnaD. The yield was low based on the ratio of the tryptic fragments that came from the substrate and from the product in the LC-MS chromatogram (total ion current). Interestingly, we did not detect, in the reaction mixture,

Table 4: Fragments Observed after Tryptic Digestion and LC-MS/MS Analysis of Prolyl-AnaD and Its Oxidation Product after AnaB Catalysis

| Fragment number | Proposed structure  | Calculated monoisotopic $m/z$ value for (M+H) <sup>+</sup><br>(Da) | Observed $m/z$ value<br>(Da) |
|-----------------|---|--|------------------------------|
| <b>F1</b>       |    | 773.79 (M+2H) <sup>2+</sup>  | 773.87                       |
| <b>F2</b>       |    | 1189.57  | 1189.53                      |
| <b>F3</b>       |    | 1091.57  | 1091.55                      |
| <b>F4</b>       |    | 456.16   | 456.17                       |
| <b>F5</b>       |    | 358.18   | 358.19                       |
| <b>F6</b>       |    | 261.13   | 261.13                       |
| <b>F7</b>       |   | 772.78 (M+2H) <sup>2+</sup>  | 772.86                       |
| <b>F8</b>       |  | 454.14   | 454.16                       |
| <b>F9</b>       |  | 356.17   | 356.17                       |

the presence of pyrrolyl-AnaD, the putative four-electron oxidation product.

In an attempt to perform kinetic analysis of the AnaB-catalyzed reaction we searched for alternate substrates for this enzyme because of the inherent difficulties in performing the reaction using prolyl-AnaD. Unfortunately, free L-proline, L-proline methyl ester, or the *N*-acetylcysteamine thioester of L-proline (ProSNAC) were not substrates for AnaB. We used two different assays to detect any AnaB activity: the reduction of DCPIP or the formation of an adduct with the putative pyrroline product and oAB. As noted by Stachelhaus and co-workers (27) the ProSNAC is rather unstable at neutral or slightly basic pH. We found that this molecule decomposed into proline and *N*-acetylcysteamine probably via a hydrolysis catalyzed by the free amine of proline (S. Mann and O. Ploux, unpublished observation). Nevertheless, using optimized conditions, we could not observe the oxidation of this prolyl-AnaD surrogate by AnaB.

## DISCUSSION

Anatoxin-a and its higher homologue homoanatoxin-a are two potent cyanobacterial neurotoxins. Cases of animal death after ingestion of water contaminated by cyanobacteria producing

these toxins have been reported, dramatically highlighting the risk that represents the presence of these neurotoxic cyanobacteria in the environment and in water reservoirs. The biosynthesis of these harmful neurotoxins remained unclear until we published recently the identification of the gene cluster responsible for the biosynthesis of anatoxin-a and homoanatoxin-a in *Oscillatoria* PCC 6506 (15). Using different approaches we concluded that the biosynthesis started from free L-proline and involved three PKSs to yield anatoxin-a or homoanatoxin-a.

In this study, we have reconstituted the first steps of the anatoxin-a/homoanatoxin-a biosynthesis in *Oscillatoria* PCC 6506. We have purified the four enzymes implicated in the transformation of free L-proline to acyl carrier protein bound dehydroproline that will serve as the actual starter for the PKS-mediated following steps (Scheme 1).

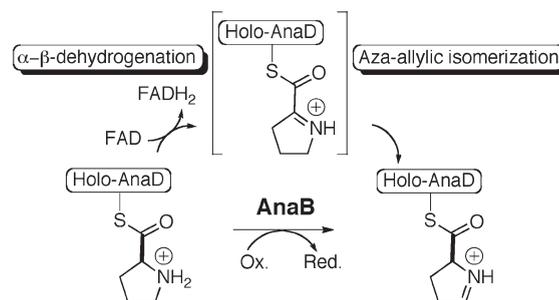
The PPTase responsible for the attachment of the phosphopantetheinyl arm on the ACP apo-AnaD was identified by annotation of the genome of PCC 6506. The *osppt* gene was cloned in an expression vector, the enzyme was purified, and its activity was clearly demonstrated using an HPLC-based assay. The catalytic properties of this newly described PPTase, the OsPPT, were similar to those reported for the PPTase from *Synechocystis* PCC 6803 (35) and from *B. subtilis*, the standard PPTase (34). Alignment of the sequence of this OsPPT with other

PPTases showed that it belonged to the Sfp superfamily, as expected, and the closest enzyme found was the PPTase from *T. erythraeum* IMS101. As for this cyanobacterium, whose genome sequence determination has been completed, we only found one PPTase in the draft genome of *Oscillatoria* PCC 6506. It cannot be excluded that another PPTase enzyme exists in *Oscillatoria* PCC 6506 since the genome sequencing is not finished, but it seems reasonable to postulate that this OsPPT serves as the transferase for all ACPs in this cyanobacterium, including ACPs involved in fatty acid metabolism.

The second step of anatoxin-a biosynthesis that we studied in this report was the loading of L-proline to the holo-AnaD, catalyzed by the adenylation protein AnaC. We showed that this enzyme was very specific for its substrate L-proline and that even close structural analogues, such as P5C, were not substrates for AnaD. The 3,4-Dhp analogue was the only alternate substrate with a  $k_{\text{cat}}/K_m$  4 times lower than that for L-proline, essentially due to a reduction in  $k_{\text{cat}}$ . Thus, AnaD discriminates among the amino acid substrates on the catalytic constant, that is, on transition state stabilization rather than on the substrate recognition. Because there are numerous adenylation proteins or domains involved in NRPS, PKS, or NRPS/PKS hybrids, several authors have tried to find a simple code that could predict the specificity of a particular adenylation protein or domain from its amino acid sequence (40, 52). These codes were based on the three-dimensional structure of GrsA (39), a Phe-specific adenylation domain, and on the amino acid sequences of several adenylation proteins or domains for which the specificity was known. The two different proposed codes for proline-specific adenylation proteins, shown in Table 3, were similar with only one position (position 278) occupied by different residues. We thus aligned the amino acid sequences of AnaC homologues together with GrsA (Supporting Information Figure S2 and Table S2) and found that none of the proposed codes were respected in these proline-specific adenylation proteins. Another putative code (Table 3) actually emerged from these alignments. Of course, we compared quite similar proteins (Supporting Information Table S2), but the fact is that the previous codes are not predictive. Determination of the three-dimensional structure of AnaC will certainly help in understanding the structural determinants for the amino acid discrimination in this class of enzymes. However, as we showed for the discrimination of proline versus 3,4-Dhp by AnaC, the driving force is the stabilization of the transition state rather than the stabilization of the substrate. Thus, it is perhaps impossible to predict the specificity just by modeling alternate substrate in the active site and by looking at favorable or unfavorable interactions. It is not surprising that simple codes based on primary sequences are thus not very helpful in predicting specificities.

BLAST searches have shown that the *anaB* gene product belonged to the ACAD superfamily and was related to prolyl-ACP dehydrogenases implicated in the biosynthesis of pyrrole-containing secondary metabolites: pyoluteorin (42), coumermycin (43), prodiginine (44), clorobiocin (45), prodigiosin (46), pyrrolomycin and dioxapyrrolomycin (49), Dkxanthen (50), and indanomycin (51). These genes, coding for prolyl dehydrogenases, are all found in clusters of genes responsible for the complete biosynthesis of the corresponding secondary metabolites. It is quite interesting to note that the genetic context is conserved in some of these clusters of genes (*ana* (15), *cou* (43), *clo* (45), and *idm* (51)), with the following gene order, prolyl-ACP dehydrogenase gene, adenylation protein gene, and ACP gene,

Scheme 2: Proposed Mechanism for the AnaB Oxidation Yielding to P5C-AnaD<sup>a</sup>



<sup>a</sup>The electron acceptor (Ox/Red) is O<sub>2</sub> in the *in vitro* experiment.

responsible for the loading of proline on an ACP followed by the oxidation of the tethered proline ring. Thus, these clusters, or at least the genes coding for the enzymes catalyzing the first steps of the biosynthesis of the secondary metabolites, are very likely evolutionarily related. Alignments of the amino acid sequence of AnaB with related prolyl-ACP dehydrogenases and representatives of the ACAD superfamily (Supporting Information Figure S3 and Table S3) showed that these FAD-containing prolyl dehydrogenases were homologous to the isovaleryl-CoA dehydrogenase (53–55) and that the active site base (E254 in human isovaleryl-CoA dehydrogenase) is conserved in AnaB and related prolyl-ACP dehydrogenases while the active site base of the pig medium chain acyl-CoA dehydrogenase (E376) was not conserved in this class of enzymes. Thus, the prolyl-AnaD dehydrogenase are likely evolutionarily related to isovaleryl-CoA dehydrogenases.

Walsh and co-workers clearly showed that CloN3, the flavo-protein involved in cloromycin biosynthesis, oxidized the prolyl-ACP to the four-electron oxidation state, that is, pyrrolyl-ACP (38). Based on early feeding experiments (56), reasonable chemistry (15), and the putative functions of the PKSs involved in anatoxin-a biosynthesis, we speculated that AnaB would only oxidize the prolyl-AnaD to dehydropyrryl-AnaD and possibly would form the (S)-pyrroline-5-carboxyl-AnaD (P5C-AnaD). As noted by others (38), it is not easy to detect a two-electron oxidation product bound to an ACP. Furthermore, prolyl-AnaD, the AnaB substrate, is unstable and decomposes into holo-AnaD and presumably proline. We thus used mass spectrometry to analyze the product formed after oxidation by AnaB, using oxygen as the electron acceptor because alternate oxidant such as PMS were not superior in the conditions tested. Digestion by trypsin followed by LC-MS/MS analysis unambiguously showed that the two-electron oxidation reaction took place on the proline moiety. Therefore, while AnaB is homologous to CloN3 and the other prolyl-dehydrogenases that oxidize the prolyl-ACP to pyrrolyl-ACP, it is the first prolyl-ACP dehydrogenase that only oxidizes prolyl-ACP to the pyrroline oxidation state. However, we could not determine, in this study, the pyrroline regioisomer that was formed.

Our mechanistic proposal is depicted in Scheme 2. This proposition is based on the homology of AnaB with ACAD and on the proposed biosynthesis of anatoxin-a in which the (S)-pyrroline-5-carboxylate thioester is believed to be the electrophile in the cyclization reaction catalyzed by the PKS AnaF (Scheme 1). The isomerization proposed here is reminiscent of the allylic isomerization catalyzed by the ACAD enzymes (57, 58). It is interesting to note that oxidation of proline to P5C is

a common reaction in the catabolism of proline and that the enzymes catalyzing this oxidation reaction are flavoenzymes (29). However, this class of enzymes, the so-called proline dehydrogenase (ProDh) family, shares no sequence similarity with AnaB and related prolyl-ACP dehydrogenases. Thus, AnaB and related enzymes would be evolutionarily related to ACAD rather than to ProDhs. The difference between AnaB and the related prolyl-ACP dehydrogenases would be that the former stops at the dehydropyrryl-ACP oxidation state and that the latter fully oxidize the substrate to pyrrolyl-ACP. We are currently trying to identify the isomer formed after AnaB oxidation of prolyl-AnaD and also to find the best electron acceptor in this reaction. Determination of the three-dimensional structure of this interesting enzyme is also underway.

## ACKNOWLEDGMENT

We thank Y. Raoul des Essarts for technical help during the purification and characterization of the OsPPT, L. Giorgi for help in performing the PP<sub>1</sub>-ATP exchange reaction, C. Médigue and her team (Génoscope, Evry) for help during bioinformatic analysis of the PCC 6506 genome data, and P. F. Leadlay (Cambridge, U.K.) for providing the Sfp overproducing strain.

## SUPPORTING INFORMATION AVAILABLE

Sequence alignments and BLAST scores for OsPPTase, AnaC, AnaB, and AnaD (Figures S1–S4 and Tables S1–S4) and determination of the sequence of the tryptic fragment of AnaD containing serine 41 (Figure S5). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- van Apeldoorn, M. E., van Egmond, H. P., Speijers, G. J., and Bakker, G. J. (2007) Toxins of cyanobacteria. *Mol. Nutr. Food Res.* 51, 7–60.
- Edwards, C., Beattie, K., Scrimgeour, C., and Codd, G. (1992) Identification of anatoxin-a in benthic cyanobacteria (blue-green algae) and in associated dog poisoning at Loch Insh, Scotland. *Toxicon* 30, 1165–1175.
- Gunn, G. J., Rafferty, A. G., Rafferty, G. C., Cockburn, N., Edwards, C., Beattie, K. A., and Codd, G. A. (1992) Fatal canine neurotoxicosis attributed to blue-green algae (cyanobacteria). *Vet. Rec.* 130, 301–302.
- Gugger, M., Lenoir, S., Berger, C., Ledreux, A., Druart, J. C., Humbert, J. F., Guette, C., and Bernard, C. (2005) First report in a river in France of the benthic cyanobacterium *Phormidium favosum* producing anatoxin-a associated with dog neurotoxicosis. *Toxicon* 45, 919–928.
- Cadel-Six, S., Peyraud-Thomas, C., Briant, L., Tandeau de Marsac, N., Rippka, R., and Méjean, A. (2007) Different genotypes of anatoxin-producing cyanobacteria coexist in the Tarn River, France. *Appl. Environ. Microbiol.* 73, 7605–7614.
- Wood, S. A., Selwood, A. I., Rueckert, A., Holland, P. T., Milne, J. R., Smith, K. F., Smits, B., Watts, L. F., and Cary, C. S. (2007) First report of homoanatoxin-a and associated dog neurotoxicosis in New Zealand. *Toxicon* 50, 292–301.
- Puschner, B., Hoff, B., and Tor, E. R. (2008) Diagnosis of anatoxin-a poisoning in dogs from North America. *J. Vet. Diagn. Invest.* 20, 89–92.
- Wonnacott, S., and Gallagher, T. (2006) The chemistry and pharmacology of anatoxin-a and related homotropans with respect to nicotinic acetylcholine receptors. *Mar. Drugs* 4, 228–254.
- Chorus, I., and Bartram, J. (1999) Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring and Management, E&FN Spon, London.
- Kaneko, T., Nakajima, N., Okamoto, S., Suzuki, I., Tanabe, Y., Tamaoki, M., Nakamura, Y., Kasai, F., Watanabe, A., Kawashima, K., Kishida, Y., Ono, A., Shimizu, Y., Takahashi, C., Minami, C., Fujishiro, T., Kohara, M., Katoh, M., Nakazaki, N., Nakayama, S., Yamada, M., Tabata, S., and Watanabe, M. M. (2007) Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res.* 14, 247–256.
- Franguel, L., Quillardet, P., Castets, A. M., Humbert, J. F., Matthijs, H. C., Cortez, D., Tolonen, A., Zhang, C. C., Gribaldo, S., Kehr, J. C., Zilliges, Y., Ziemert, N., Becker, S., Talla, E., Latifi, A., Billault, A., Lepelletier, A., Dittmann, E., Bouchier, C., and de Marsac, N. T. (2008) Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. *BMC Genomics* 9, 274.
- Tillett, D., Dittmann, E., Erhard, M., von Döhren, H., Börner, T., and Neilan, B. A. (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. *Chem. Biol.* 7, 753–764.
- Noguchi, T., Shinohara, A., Nishizawa, A., Asayama, M., Nakano, T., Hasegawa, M., Harada, K., Nishizawa, and T., Shirai, M. (2009) Genetic analysis of the microcystin biosynthesis gene cluster in *Microcystis* strains from four bodies of eutrophic water in Japan. *J. Gen. Appl. Microbiol.* 55, 111–123.
- Kellmann, R., Mihali, T. K., Jeon, Y. J., Pickford, R., Pomati, F., and Neilan, B. A. (2008) Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Appl. Environ. Microbiol.* 74, 4044–4053.
- Méjean, A., Mann, S., Maldiney, T., Vassiliadis, G., Lequin, O., and Ploux, O. (2009) Evidence that biosynthesis of the neurotoxic alkaloids anatoxin-a and homoanatoxin-a in the cyanobacterium *Oscillatoria* PCC 6506 occurs on a modular polyketide synthase initiated by L-proline. *J. Am. Chem. Soc.* 131, 7512–7513.
- Cadel-Six, S., Iteman, I., Peyraud-Thomas, C., Mann, S., Ploux, O., and Méjean, A. (2009) Identification of a polyketide synthase coding sequence specific for anatoxin-a-producing *Oscillatoria* cyanobacteria. *Appl. Environ. Microbiol.* 75, 4909–4912.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- Herdman, M., Iteman, I., and Rippka, R. (2005) Catalogue of cyanobacterial strains, 2nd ed., Institut Pasteur, Paris.
- Rippka, R. (1988) Isolation and purification of cyanobacteria. *Methods Enzymol.* 167, 3–27.
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.
- Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., Lajus, A., Pascal, G., Scarpelli, C., and Médigue, C. (2006) MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* 34, 53–65.
- Ellis, K. E., and Morrison, J. F. (1982) Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol.* 87, 405–426.
- Gerratana, B., Arnett, S. O., Stapon, A., and Townsend, C. A. (2004) Carboxymethylproline synthase from *Pectobacterium carotovora*: a multifaceted member of the crotonase superfamily. *Biochemistry* 43, 15936–15945.
- Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. *Biochemistry* 42, 7836–7847.
- Williams, I., and Frank, L. (1975) Improved chemical synthesis and enzymatic assay of delta-1-pyrroline-5-carboxylic acid. *Anal. Biochem.* 64, 85–97.
- Mezl, V. A., and Knox, W. E. (1976) Properties and analysis of a stable derivative of pyrroline-5-carboxylic acid for use in metabolic studies. *Anal. Biochem.* 74, 430–440.
- Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. *Chem. Biol.* 7, 765–772.
- Becker, D. F., and Thomas, E. A. (2001) Redox properties of the PutA protein from *Escherichia coli* and the influence of the flavin redox state on PutA-DNA interactions. *Biochemistry* 40, 4714–21.
- White, T. A., Krishnan, N., Becker, D. F., and Tanner, J. J. (2007) Structure and kinetics of monofunctional proline dehydrogenase from *Thermus thermophilus*. *J. Biol. Chem.* 282, 14316–14327.
- Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* 68, 1–8.
- Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) A new enzyme superfamily—the phosphopantetheinyl transferases. *Chem. Biol.* 3, 923–936.
- Weissman, K. J., Hong, H., Oliynyk, M., Siskos, A. P., and Leadlay, P. F. (2004) Identification of a phosphopantetheinyl transferase for

- erythromycin biosynthesis in *Saccharopolyspora erythraea*. *ChemBioChem* 5, 116–125.
33. Copp, J. N., and Neilan, B. A. (2006) The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Appl. Environ. Microbiol.* 72, 2298–2305.
  34. Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* 37, 1585–1595.
  35. Copp, J. N., Roberts, A. A., Marahiel, M. A., and Neilan, B. A. (2007) Characterization of PPTNs, a cyanobacterial phosphopantetheinyl transferase does not possess characteristic broad-range activity. *ChemBioChem* 10, 3133–3139.
  36. Roberts, A. A., Copp, J. N., Marahiel, M. A., and Neilan, B. A. (2009) The *Synechocystis* sp. PCC6803 Sfp-type phosphopantetheinyl transferase does not possess characteristic broad-range activity. *ChemBioChem* 10, 1869–1877.
  37. Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V., and Stroh, J. G. (1999) Spontaneous alpha-N-6-phosphogluconoylation of a “His tag” in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* 267, 169–184.
  38. Garneau, S., Dorrestein, P. C., Kelleher, N. L., and Walsh, C. T. (2005) Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A1 biosynthesis. *Biochemistry* 44, 2770–2780.
  39. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* 16, 4174–4183.
  40. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in non-ribosomal peptide synthetases. *Chem. Biol.* 6, 493–505.
  41. Villiers, B. R., and Hollfelder, F. (2009) Mapping the limits of substrate specificity of the adenylation domain of TycA. *ChemBioChem* 10, 671–682.
  42. Nowak-Thompson, B., Chaney, N., Wing, J. S., Gould, S. J., and Loper, J. E. (1999) Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas fluorescens* Pf-5. *J. Bacteriol.* 181, 2166–2174.
  43. Wang, Z. X., Li, S. M., and Heide, L. (2000) Identification of the coumermycin A(1) biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob. Agents Chemother.* 44, 3040–3048.
  44. Cerdeño, A. M., Bibb, M. J., and Challis, G. L. (2001) Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3(2): new mechanisms for chain initiation and termination in modular multienzymes. *Chem. Biol.* 8, 817–829.
  45. Pojer, F., Li, S. M., and Heide, L. (2002) Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics. *Microbiol.* 148, 3901–3911.
  46. Harris, A. K., Williamson, N. R., Slater, H., Cox, A., Abbasi, S., Foulds, I., Simonsen, H. T., Leeper, F. J., and Salmond, G. P. (2004) The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species- and strain-dependent genome context variation. *Microbiology* 150, 3547–3560.
  47. Garneau-Tsodikova, S., Dorrestein, P. C., Kelleher, N. L., and Walsh, C. T. (2006) Protein assembly line components in prodigiosin biosynthesis: characterization of PigA,G,H,I,J. *J. Am. Chem. Soc.* 128, 12600–12601.
  48. Walsh, C. T., Garneau-Tsodikova, S., and Howard-Jones, A. R. (2006) Biological formation of pyrroles: nature’s logic and enzymatic machinery. *Nat. Prod. Rep.* 23, 517–531.
  49. Zhang, X., and Parry, R. J. (2007) Cloning and characterization of the pyrrolomycin biosynthetic gene clusters from *Actinosporangium vitaminophilum* ATCC 31673 and *Streptomyces* sp. strain UC 11065. *Antimicrob. Agents Chemother.* 51, 946–957.
  50. Meiser, P., Weissman, K. J., Bode, H. B., Krug, D., Dickschat, J. S., Sandmann, A., and Müller, R. (2008) DKxanthene biosynthesis: understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem. Biol.* 15, 771–781.
  51. Li, C., Roege, K. E., and Kelly, W. L. (2009) Analysis of the indanomycin biosynthetic gene cluster from *Streptomyces antibioticus* NRRL 8167. *ChemBioChem* 10, 1064–1072.
  52. Challis, G. L., Ravel, R., and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* 7, 211–224.
  53. Kim, J. J., Wang, M., and Paschke, R. (1993) Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc Natl. Acad. Sci. U.S.A.* 90, 7523–7527.
  54. Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J., and Kim, J. J. (1997) Structure of human isovaleryl-CoA dehydrogenase at 2.6 Å resolution: structural basis for substrate specificity. *Biochemistry* 36, 8455–8464.
  55. Battaile, K. P., Nguyen, T. V., Vockley, J., and Kim, J. J. (2004) Structures of isobutyryl-CoA dehydrogenase and enzyme-product complex: comparison with isovaleryl- and short-chain acyl-CoA dehydrogenases. *J. Biol. Chem.* 279, 16526–16534.
  56. Hemscheidt, T., Rapala, J., Sivonen, K., and Skulberg, O. M. (1995) Biosynthesis of anatoxin-a in *Anabaena flos-aquae* and homoanatoxin-a in *Oscillatoria formosa*. *J. Chem. Soc., Chem. Commun.*, 1361–1362.
  57. Dakoji, S., Shin, I., Battaile, K. P., Vockley, J., and Liu, H. W. (1997) Redesigning the active-site of an acyl-CoA dehydrogenase: new evidence supporting a one-base mechanism. *Bioorg. Med. Chem.* 5, 2157–2164.
  58. Dakoji, S., Shin, I., Becker, D. F., Stankovich, M. T., and Liu, H.-w. (1996) Studies of acyl-CoA dehydrogenase catalyzed allylic isomerization: a one-base or two-base mechanism? *J. Am. Chem. Soc.* 118, 10971–10979.