2014

Characterization of an Orphan Diterpenoid Biosynthetic Operon from Salinispora arenicola

Meimei Xu
_Iowa State University, xumm@iastate.edu_

Matthew L. Hillwig
_Iowa State University_

Amy L. Lane
_University of California - San Diego_

Mollie Tiernan
_Iowa State University_

Bradley S. Moore
_University of California - San Diego_

Follow this and additional works at: [http://lib.dr.iastate.edu/bbmb_ag_pubs](http://lib.dr.iastate.edu/bbmb_ag_pubs)

Part of the Biochemistry, Biophysics, and Structural Biology Commons, Marine Biology Commons, and the Natural Products Chemistry and Pharmacognosy Commons

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/bbmb_ag_pubs/5](http://lib.dr.iastate.edu/bbmb_ag_pubs/5). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).

This Article is brought to you for free and open access by the Biochemistry, Biophysics and Molecular Biology at Iowa State University Digital Repository. It has been accepted for inclusion in Biochemistry, Biophysics and Molecular Biology Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Characterization of an Orphan Diterpenoid Biosynthetic Operon from Salinispora arenicola

Abstract
While more commonly associated with plants than microbes, diterpenoid natural products have been reported to have profound effects in marine microbe–microbe interactions. Intriguingly, the genome of the marine bacterium *Salinispora arenicola* CNS-205 contains a putative diterpenoid biosynthetic operon, *terp1*. Here recombinant expression studies are reported, indicating that this three-gene operon leads to the production of isopimara-8,15-dien-19-ol (4). Although 4 is not observed in pure cultures of *S. arenicola*, it is plausible that the *terp1* operon is only expressed under certain physiologically relevant conditions such as in the presence of other marine organisms.

Keywords
Salinispora arenicola, marine, terp1

Disciplines
Biochemistry, Biophysics, and Structural Biology | Marine Biology | Natural Products Chemistry and Pharmacognosy

Comments
Reprinted (adapted) with permission from *Journal of Natural Products* 77 (2014): 2144, doi:10.1021/np500422d. Copyright 2014 American Chemical Society.

Rights
One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

Authors
Meimei Xu, Matthew L. Hillwig, Amy L. Lane, Mollie Tiernan, Bradley S. Moore, and Reuben J. Peters

This article is available at Iowa State University Digital Repository: [http://lib.dr.iastate.edu/bbmb_ag_pubs/5](http://lib.dr.iastate.edu/bbmb_ag_pubs/5)
Characterization of an Orphan Diterpenoid Biosynthetic Operon from *Salinispora arenicola*

Meimei Xu,† Matthew L. Hillwig,†,§ Amy L. Lane,‡⊥ Mollie S. Tiernan,† Bradley S. Moore,‡ and Reuben J. Peters*,†

†Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, Iowa 50011 United States
‡Scripps Institution of Oceanography and Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093 United States
§Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, Iowa 50011 United States

**ABSTRACT:** While more commonly associated with plants than microbes, diterpenoid natural products have been reported to have profound effects in marine microbes—microbe interactions. Intriguingly, the genome of the marine bacterium *Salinispora arenicola* CNS-205 contains a putative diterpenoid biosynthetic operon, *terp1*. Here recombinant expression studies are reported, indicating that this three-gene operon leads to the production of isopimara-8,15-dien-19-ol (4). Although 4 is not observed in pure cultures of *S. arenicola*, it is plausible that the *terp1* operon is only expressed under certain physiologically relevant conditions such as in the presence of other marine organisms.

The production of terpenoids is most commonly associated with plants, which have clearly expanded their ability to produce this class of natural products. For example, terpene synthases form moderate sized gene families, with more than 10 such enzymatic genes found in each of the known vascular plant genome sequences.¹ By contrast, there appear to be just over 100 terpene synthases among the more than 1000 sequenced bacterial genomes, suggesting the relative scarcity of terpene synthases and, hence, terpenoid production, among commonly studied bacterial genera.

This relative paucity of bacterial terpenoid production is also illustrated by the labdane-related diterpenoids, a large superfamilly of ~7000 known natural products whose biosynthesis is characterized by the initiating reaction. Specifically, acid—base catalyzed bicyclization of the general diterpenoid precursor (E,E,E)-geranylgeranyl diphosphate (GGPP, 1), which is mediated by class II diterpene cyclases that generally form the eponymous labdadienyl/copalyl diphosphate (CPP). This is typically followed by an additional cyclization and/or rearrangement reaction initiated by ionization of the allylic diphosphate catalyzed by a class I diterpene synthase.³ All vascular plants have at least one class II diterpene cyclase in order to produce the requisite gibberellin phytohormones, and many plant species have multiple such enzymes.⁴ However, less than 100 class II diterpene cyclases are present in the known bacterial genomes⁵ and less than 10 have been characterized.⁶–¹²

Nevertheless, there are a handful of bacteria that produce labdane-related diterpenoids of significant interest. The relevant class II diterpene cyclases have been identified for a number of these bacterial natural products, including such enzymes involved in the production of gibberellin phytohormones by plant symbiotic rhizobia and phytopathogens,⁹,¹¹,¹² the potential antibacterial and antidiabetic compounds platencin and platensimycin by *Streptomyces platensis*,¹⁰ and an immunomodulatory factor by *Mycobacterium tuberculosis.*⁸ Intriguingly, it has been reported that an epiphytic marine bacterium produces a labdane-related diterpenoid that acts at subpicomolar levels to promote the aggregation of marine green macroalgae (e.g., sea lettuce).¹³

With the advent of next generation sequencing, there has been a tremendous increase in the availability of microbial genome sequences,¹⁴ which has generally revealed that there are many more putative natural product biosynthetic gene clusters than known metabolites.¹⁵ A variety of approaches have been taken toward identifying the compounds resulting from such orphan operons.¹⁶ Here is reported the use of recombinant expression to characterize a bioinformatics-predicted labdane-related diterpenoid biosynthetic operon from the marine bacterium *Salinispora arenicola* CNS-205.

The genome sequence of the CNS-205 strain of the widely distributed marine actinobacterium *S. arenicola* was previously reported, with bioinformatic analysis indicating the presence of 30 putative natural product biosynthetic gene clusters.¹⁷ Included among these was *terp1*, annotated as producing an unidentified diterpene. This small operon contains genes encoding a putative class II diterpene cyclase (Sare_1288) and class I (di)terpene synthase (Sare_1287) as well as a cytochrome P450 (CYP) mono-oxygenase (Sare_1286) that has been assigned as CYP1051A1 (Figure 1). Thus, the *terp1*
The putative class II diterpene cyclase is most closely related to two previously characterized ent-CPP synthases (CPSs) from Streptomyces species of terrestrial actinobacteria (42–45% amino acid sequence identity) and is referred to henceforth as SaDTS. While class II diterpene cyclases almost invariably contain a DxDD motif that cooperatively acts as the catalytic acid, SaCPS contains a Thr in place of the last Asp amino acid, suggesting that SaCPS might be functional.

Similarly, the putative class I diterpene synthase is most closely related to a previously characterized ent-pimaradiene synthase from Streptomyces sp. KO-3988 (∼30% amino acid sequence identity) and is referred to henceforth as SaDTS. While class II diterpene cyclases almost invariably contain a DxDD motif involved in binding the requisite divalent magnesium ion cofactors, SaDTS does not contain a corresponding sequence. Amino acid sequence alignments indicate that the corresponding sequence in SaDTS is EDWQVD₈₃₋₈₈ instead (SI, Figure S2). While the ent-kaurene synthase from S. platensis also does not have the canonical DxDD motif at this position, sequence conforming to the DxDD motif can be found nearby, which is not true in SaDTS. On the other hand, SaDTS does contain the NDxxSxxxE motif also involved in binding the requisite divalent magnesium ion co-factors, leaving open the possibility that SaDTS may be active as well.

Characterization of the bioinformatics-predicted labdane-related diterpene product of this operon was undertaken with a previously developed modular metabolic engineering system. This enabled recombinant expression of SaCPS with a GGPP synthase (GGPS) in Escherichia coli, which led to the production of CPP, observed as the dephosphorylated copalol by GC-MS analysis of hexane extracts of the induced culture (Figure 2A). To determine the absolute configuration of this CPS, SaCPS was co-expressed with GGPS and selective class I diterpene synthases, much as previously described for investigation of other class II diterpene cyclases. Briefly, SaCPS and the GGPS were co-expressed with either the ent-kaurene synthase from Arabidopsis thaliana (AtKS) that selectively reacts with ent-CPP, or a mutant of the abietadiene synthase from Abies grandis (AgAS) that no longer exhibits class II activity and only reacts with normal CPP (AgAS:D404A). No ent-kaurene was observed upon co-expression of SaCPS (and GGPS) with AtKS, while the same products made by wild-type AgAS were readily observed upon co-expression with AgAS:D440A (SI, Figure S3). Thus, SaCPS makes normal (5S,9S,10S) CPP (2).

Co-expression of SaDTS with GGPS and SaCPS led to production of an unknown diterpene, observed by GC-MS analysis of hexane extracts of the induced culture. To determine the structure of this compound, SaDTS was co-expressed with GGPS and a plant CPS in E. coli (Figure 2B), which produces larger quantities of 2 (i.e., than SaCPS), and the culture volume was increased to 2 L. This enabled isolation of ∼2 mg for NMR analysis (SI, Figures S4–S6 and Table S1), which indicated that this was a (iso)pimarane-8,15-diene, with resolution of the configuration at C-13 derived from comparison to previously reported NMR chemical shift data, which led to assignment of the SaDTS product as isopimara-8,15-diene (3).

To functionally characterize CYP1051A1, it was necessary to account for the fact that CYPs require the input of electrons, generally obtained from NADPH, and in the case of bacterial CYPs typically provided by a ferredoxin (FdR) that has been reduced by a ferredoxin reductase (Fdr). Accordingly, CYP1051A1 was co-expressed with an Fdx and Fdr from S. arenicola (Sare_4141 and Sare_0646, respectively) in a strain of

Figure 1. Schematic of the terpI diterpene biosynthetic operon from S. arenicola CNS-205.

Figure 2. GC-MS extracted ion (m/z = 257) chromatograms and associated mass spectra for (A) production of CPP (2), detected as dephosphorylated copalol (2'; retention time, RT = 17.16 min), from expression of SaCPS in a strain of E. coli engineered to make GGPP (1) by SaCPS. (B) Isopimara-8,15-diene (3; RT = 15.27 min) from expression of SaDTS in a strain engineered to make 2. (C) Isopimara-8,15-dien-19-ol (4; RT = 17.04 min) from co-expression of CYP1051A1 with an Fdx and Fdr in a strain engineered to make 3.
E. coli also engineered to produce 3. The resulting recombinant strain produced a hydroxylated derivative of 3, observed by GC-MS analysis of an organic solvent extract (Figure 2C). Attempts to scale up production of the CYP1051A1 product were not fruitful, with a net yield of only ~100 μg from 5 L of culture. Fortuitously, it was discovered that CYP99A3 from rice catalyzed the same reaction (SI, Figure S7) and was more amendable to scale-up. Thus, CYP99A3 was used to produce enough compound (~1 mg), which was mixed with the ~0.1 mg isolated from CYP1051A1 for NMR analysis (SI, Figures S4, S8, and S9 and Table S2). Only a single peak was found in the alcohol region of the 13C spectra (SI, Figure S9), consistent with equivalence of the CYP99A3 and CYP1051A1 products, which the collected data indicated was isopimara-8,15-dien-19-ol (4). The configuration at C-4 was suggested by the observation of an NOE signal between the secondary alcohol hydrogens on C-19 and the methyl hydrogens of C-20, whose configuration was already known (vide supra) and verified by comparison to previously reported NMR chemical shift data.28 Thus, the S. arenicola terp1 operon presumably can lead to the production of isopimara-8,15-dien-19-ol (Scheme 1).

It has been previously reported that at least two species of Streptomyces grown in certain media produce the diterpenoid vigulepinol (3α-hydroxy-ent-pimara-9(11),15-diene) and derived oxalo-terpins, which are very similar to viguiepinol (3). Previously described as a colorless solid,29,30 1H and 13C NMR data largely match literature values,28 again with the few significant differences supported here by HMBC correlations (SI, Table S1). Isopimara-8,15-diene (3). Previously described as a colorless solid,29,30 1H and 13C NMR, as well as MS, data largely match literature values,24,25,39 with the few significant differences supported here by HMBC correlations (SI, Table S1).

In conclusion, functional characterization of the orphan terp1 biosynthetic operon from S. arenicola is reported here. While the terp1 operon is only found in strain CNS-205, these results further illuminate the natural products capacity of these widely distributed marine actinobacteria. Although the resulting diterpenoid 4 cannot be found in pure cultures of S. arenicola CNS-205 grown in a variety of media, it is possible that 4 serves an ecological function and is only produced by S. arenicola CNS-205 under specific environmental conditions (e.g., in the presence of other organisms). Alternatively, it has been previously noted that all other characterized bacterial diterpenoid biosynthetic operons contain a GGPS as bacteria do not generally produce 1 otherwise.7,10,12,31–36 and the lack of a GGPS in the S. arenicola CNS-205 terp1 operon may underlie the observed lack of production of 4. In any case, these results demonstrate the applicability of the utilized modular metabolic engineering system to elucidation of (di)terpenoid biosynthetic operons uncovered by genome mining of actinobacteria and potentially other bacteria as well.
Author Contributions
The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS
This work was supported by funds from the NIH, an IRACDA postdoctoral fellowship to A.L.L. (GM068524) and research grants to B.S.M. (GM085770) and R.J.P. (GM076324).

REFERENCES