View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Harvard University - DASH

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

DASH.HARVARD.EDU

HARVARD LIBRARY

Office for Scholarly Communication

The Genetic and Molecular Basis for Sunscreen Biosynthesis in Cyanobacteria

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Balskus, E. P., and C. T. Walsh. 2010. "The Genetic and Molecular Basis for Sunscreen Biosynthesis in Cyanobacteria." Science 329 (5999) (September 2): 1653–1656. doi:10.1126/science.1193637.
Published Version	doi:10.1126/science.1193637
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:33946912
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

NIH Public Access
Author Manuscript
Science. Author manuscript; available in PMC 2011 June 16.

Published in final edited form as:
Science. 2010 September 24; 329(5999): 1653–1656. doi:10.1126/science.1193637.

The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria

Emily P. Balskus and **Christopher T. Walsh**[*]
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

Abstract

UV-A and UV-B radiation are harmful to living systems, causing damage to biological macromolecules. An important strategy for dealing with UV exposure is the biosynthesis of small molecule sunscreens. Among such metabolites, the mycosporine and mycosporine-like amino acids (MAAs) are remarkable for their wide phylogenetic distribution and their unique chemical structures. Here we report the identification of a MAA biosynthetic gene cluster in a cyanobacterium and the discovery of analogous pathways in other sequenced organisms. We have expressed the cluster in a heterologous bacterial host and characterized all four biosynthetic enzymes *in vitro*. In addition to clarifying the origin of the MAAs, these efforts have revealed two unprecedented enzymatic strategies for imine formation.

Mycosporines and MAAs are UV-absorbing small molecules (λ_{max} = 310–360 nm) (1) that are synthesized by cyanobacteria, fungi, and eukaryotic micro- and macroalgae (2–5). These metabolites are also ingested and accumulated by numerous marine animals. Discovered and named for mediating light-induced fungal sporulation (6–8), mycosporines and MAAs are involved in many other biological processes, including UV photoprotection of organisms (2–5,9) and their embryos (10), osmotic regulation (11), and defense against oxidative stress (12). Additionally, their ability to prevent UV-induced skin damage *in vivo* (13) has led to their commercialization as Helioguard 365™, a formulation containing MAAs shinorine and porphyra-334 that is used in skin care and cosmetic products.

Structurally this large natural product family is divided into two groups: the mycosporines, which contain a single proteogenic or modified amino acid residue linked to a cyclohexenone core, and the MAAs, which have two such substituents (Figure 1a). The amino acids are attached to the central ring through imine linkages, an arrangement best represented as a combination of resonance tautomers (Figure 1b). This extensive conjugation facilitates absorption of UV (Figure 1c). Fungi produce only mycosporines, while cyanobacteria, algae, and marine organisms can synthesize both mycosporines and MAAs.

The occurrence, distribution, and biology of mycosporines and MAAs have been subjects of intense investigation, but elucidation of the precise biochemical mechanisms underlying their functions has been hampered by a lack of knowledge regarding the molecular and genetic basis for their biosynthesis. Previous feeding studies with cyanobacteria and fungi revealed that the amino acid substituents were added in the later stages of biosynthesis and suggested, but did not prove, that the mycosporine core might originate from shikimate pathway intermediate 3-dehydroquinate (14,15). Despite these efforts, no genes have been definitively linked to mycosporine or MAA production in any organism.

[*]To whom correspondence should be addressed. christopher_walsh@hms.harvard.edu.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We attempted to locate candidate biosynthetic gene clusters using a genome mining approach prompted by the discovery of genes encoding shikimate pathway enzymes in the genome of sea anenome *Nematostella vectensis* (16), including a dehydroquinate synthase homolog (DHQS) adjacent to an O-methyltransferase (O-MT). Recognizing that a methyltransferase was likely involved in mycosporine assembly, we considered the DHQS and O-MT to be candidate biosynthetic genes. BLAST searches with both protein sequences revealed clustered homologs in genomes of dinoflagellates, cyanobacteria, and fungi (17).

Examining the hits in their genomic contexts revealed that open reading frames (ORFs) encoding DHQSs and O-MTs were accompanied by a third conserved ORF, usually annotated as a hypothetical protein. A conserved domain search using this gene revealed homology to the ATP-grasp superfamily. As these enzymes catalyze peptide bond formation, we reasoned that this gene product might be involved in amino acid installation. We also noted genetic variations between cyanobacterial and fungal clusters consistent with the structural differences between cyanobacterial MAAs and fungal mycosporines; cyanobacteria had a second putative amide bond forming enzyme, either an ATP-grasp or nonribosomal peptide synthetase (NRPS) homolog, while fungi contained additional enzymes or domains annotated as reductases. Recently others identified the same DHQS and O-MT genes, but did not note the conserved ATP-grasp homolog nor demonstrate a genetic or biochemical connection to MAA production (18).

The cyanobacterium *Anabaena variabilis* ATCC 29413 is an established producer of the MAA shinorine (Figure 1a) (19) and contains a putative gene cluster consisting of four ORFs (Figure 2a). We hypothesized that DHQS homolog Ava_3858 and O-MT 3857 would assemble 4-deoxygadusol, a cyclohexenone natural product that resembles the mycosporine core (20), and that ATP-grasp homolog Ava_3856 and NRPS-like enzyme Ava_3855 might be responsible for attaching glycine and serine to 4-deoxygadusol via imine linkages (Figure 2b). We demonstrated a direct connection of these genes to MAA biosynthesis by cloning the entire 6.5 kb shinorine biosynthetic gene cluster, as well as two truncated constructs lacking either Ava_3855 or both Ava_3855 and Ava_3856, into the heterologous host *Escherichia coli*. After inducing gene expression, we detected shinorine production only in the strain harboring the complete gene cluster (Figure 2c). Strains containing truncated clusters produced mycosporine glycine (Ava_3855 truncation) and 4-deoxygadusol (Ava_3855/3856 truncation). All three products were isolated from culture extracts and characterized (17). These experiments clearly demonstrate the involvement of this gene cluster in shinorine production and confirm the identities of key biosynthetic intermediates. The synthesis of shinorine by *E. coli* suggests that no additional genes are required beyond those encoded in the *A. variabilis* gene cluster and that the pathway originates from a metabolic intermediate common to both organisms. Finally, these results provide strong evidence that homologous clusters are involved in mycosporine and MAA synthesis.

Seeking an understanding of the chemistry underlying MAA assembly, we characterized each of the shinorine biosynthetic enzymes *in vitro*. Based on past successes expressing enzymes from the shinorine-producing cyanobacterium *Nostoc punctiforme* ATCC 29133 (18) in *E. coli*, we chose to examine the DHQS homolog (NpR5600) and O-MT (NpR5599) from this organism. Unexpectedly, we failed to observe any 4-deoxygadusol production when the putative substrate 3-dehydroquinate was incubated with both enzymes and typical cofactors SAM, NAD^+, and Co^{2+} (21). Closer examination of the homology between NpR5600 and canonical DHQSs revealed that the active site residues of enzymes associated with mycosporine biosynthesis resemble DHQS homolog 2-*epi*-5-*epi*-valionate synthase (EVS) rather than a canonical DHQS (18,22,23), suggesting that these enzymes might share a common substrate, pentose phosphate pathway intermediate sedoheptulose 7-phosphate (SH 7-P). Gratifyingly, treatment of SH 7-P with NpR5600 and NpR5599 in the presence of

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

SAM, NAD^+, and Co^{2+} produced a single product, 4-deoxygadusol. Kinetic analysis and a mechanistic proposal are detailed in the Supporting Information (Figures S5–S7 and Scheme S1). This result is consistent with previous feeding studies, yet contradicts a longstanding assumption that MAA biosynthesis involves a shikimate pathway intermediate.

We next investigated the final steps of shinorine biosynthesis, installation of glycine and serine substituents. Imines, or Schiff bases, are important functional groups in biological systems and are typically formed by condensation of an amine nucleophile and a carbonyl electrophile in an ATP-independent manner. This type of imine formation is utilized by many enzyme families, including class I aldolases and pyridoxal phosphate (PLP)-dependent enzymes. Previous studies suggested a distinct mechanism might operate in mycosporine biosynthesis, as mycosporine glycine will not condense with amino acids to form MAAs (24). Accordingly, the above heterologous expressions confirmed that homologs of ATP-dependent peptide bond-forming enzymes were required for amino acid attachment. As the imine linkages in mycosporines and MAAs differ from peptide bonds, the mechanisms underlying these ligation steps were of great interest.

Attempts to overexpress putative imine-forming genes from the *N. punctiforme* shinorine cluster were unsuccessful; we therefore returned to *A. variabilis*, cloning both ATP-grasp family member Ava_3856 and NRPS-like enzyme Ava_3855 and overexpressing the corresponding proteins in *E. coli*. HPLC assays of Ava_3856 confirmed its ability to convert 4-deoxygadusol and glycine into mycosporine glycine in the presence of ATP and Mg^{2+} cofactors. Two potential mechanisms could account for this transformation, one involving phosphorylation of 4-deoxygadusol followed by 1,4-addition of the amino acid nitrogen to the activated cyclohexenone, and the other invoking phosphorylation of glycine to give an acyl phosphate that could be attacked by 4-deoxygadusol. The acylated enol intermediate formed could then undergo an O to N rearrangement driven by an intramolecular 1,4-addition of the glycine amino group to the cyclohexenone core (Scheme S2).

To distinguish between these possibilities we performed an assay with ^{18}O-labeled glycine and analyzed the isotope content of mycosporine glycine using high-resolution liquid chromatography-mass spectrometry (LC-MS). The distribution of ^{18}O in the product was identical to that of the glycine starting material (Figure 3a), indicating that Ava_3856 phosphorylates 4-deoxygadusol rather than glycine, as the latter would be accompanied by loss of one ^{18}O atom via elimination of phosphate. Phosphorylation of 4-deoxygadusol was unexpected; while members of the ATP-grasp family employ various nucleophilic partners, including amines and thiols, the electrophiles activated are typically carboxylic acids (25).

The final enzyme characterized was Ava_3855, an NRPS-like enzyme containing adenylation (A), thiolation (T), and thioesterase (TE) domains (26). HPLC assays confirmed its role in converting mycosporine glycine and serine to shinorine; activity was dependent on the presence of ATP and the alternate amino acid threonine was not utilized by the enzyme. As with Ava_3856, a key mechanistic question regarding the action of Ava_3855 was the identity of the electrophilic reaction partner. In this case, activation of serine as an acyl adenylate was predicted based on conserved A domain active site residues (27) and confirmed using an ATP-$^{32}PP_i$ exchange assay. Covalent attachment of radiolabeled serine onto the T domain of Ava_3855 was also demonstrated.

To establish that the serine thioester was on pathway to shinorine we assayed Ava_3855 using [carboxy-$^{18}O_2$]serine. In this case, the product LC-MS revealed a shift of two mass units, indicating that Ava_3855 indeed operates through activation of the serine carboxylate; adenylation and loading by the A and T domains is accompanied by loss of ^{18}O as adenosine monophosphate (AMP). The mechanism of imine formation may involve TE domain-

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

catalyzed formation of an enol ester intermediate, which could undergo subsequent O to N rearrangement via 1,4-addition of the serine nitrogen to the activated cyclohexenimine core (Figure 3b). We favor this pathway as it is consistent with the poor electrophilicity of mycosporine glycine (24); an alternative mechanism for Ava_3855 involving a direct condensation is detailed in Scheme S3.

The *in vitro* characterization of Ava_3856 and Ava_3855 reveals two distinct, yet complementary, mechanisms of ATP-dependent enzymatic imine formation that differ from conventional chemical methods and biochemical mechanisms. These enzymes have evolved from peptide bond-forming catalysts in distinct ways: ATP-grasp homolog Ava_3856 generates a new type of electrophile using vinylogous acid activation, while NRPS-like enzyme Ava_3855 likely employs an unusual release mechanism. The recruitment of ATP-dependent peptide bond-forming enzymes in this manner is so far unprecedented in natural product biosynthesis and defines a new biosynthetic logic for imine construction. A short four-enzyme pathway thus converts a primary metabolite from the pentose phosphate pathway into a widely distributed class of small molecule biological sunscreens using mechanistically elegant chemistry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the NIH (GM-20011). E.P.B. is the recipient of an NIH NRSA postdoctoral fellowship (GM-084625). We acknowledge Tri-K Industries (Northvale, NJ) for providing a sample of Helioguard 365 containing authentic standards of shinorine and porphyra-334.

References and Notes

1. Cockell CS, Knowland J. Biol Rev. 1999; 74:311–345. [PubMed: 10466253]
2. Bandaranayake WM. Nat Prod Rep. 1998; 15:159–172. [PubMed: 9586224]
3. Shick JM, Dunlap WC. Annu Rev Physiol. 2002; 64:223–262. [PubMed: 11826269]
4. Oren A, Gunde-Cimerman N. FEMS Microbiol Lett. 2007; 269:1–10. [PubMed: 17286572]
5. Dunlap WC, Shick JM. J Phycol. 1998; 34:418–430.
6. Trione EJ, Leach CM, Mutch JT. Nature. 1966; 212:163–164. [PubMed: 5972212]
7. Favre-Bonvin J, Arpin N, Brevard C. Can J Chem. 1976; 54:1105–1113.
8. Dehorter B, et al. Phytochemistry. 1980; 19:2311–2315.
9. Neale PJ, Banaszak AT, Jarriel CR. J Phycol. 1998; 34:928–938.
10. Adams NL, Shick JM. Marine Biology. 2001; 138:267–280.
11. Oren A. Geomicrobiol J. 1997; 14:231–240.
12. Yakovleva I, Bhagooli R, Takemura A, Hidaka M. Comp Biochem Physiol B. 2004; 139:720–730.
13. de la Coba F, et al. J Dermatol Sci. 2009; 55:161–169. [PubMed: 19586754]
14. Favre-Bonvin J, Bernillon J, Salin N, Arpin N. Phytochemistry. 1987; 26:2509–2514.
15. Portwich A, Garcia-Pichel F. Phycologia. 2003; 42:384–392.
16. Starcevic A, et al. Proc Natl Acad Sci USA. 2008; 105:2533–2537. [PubMed: 18268342]
17. Supporting information containing materials and methods, details of bioinformatics analyses, and assay data is available on Science Online.
18. Singh SP, Klisch M, Sinha RP, Häder DP. Genomics. 2010; 95:120–128. [PubMed: 19879348]
19. Singh SP, Klisch M, Sinha RP, Häder DP. Photochem Photobiol. 2008; 84:1500–1505. [PubMed: 18557824]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

20. Lemoyne F, Bernillon J, Favre-Bonvin J, Bouillant ML, Arpin N. Z Naturforsch, C: Biosci. 1985; 40:612–616.
21. Bender SL, Medhi S, Knowles JR. Biochemistry. 1989; 28:7555–7560. [PubMed: 2514789]
22. Wu X, et al. ChemBioChem. 2007; 8:239–248. [PubMed: 17195255]
23. Mahmud T, Flatt PM, Wu X. J Nat Prod. 2007; 70:1384–1391. [PubMed: 17661520]
24. White JD, Cammack JH, Sakuma K. J Am Chem Soc. 1989; 111:8970–8972.
25. Galperin MY, Koonin EV. Protein Science. 1997; 6:2639–2643. [PubMed: 9416615]
26. Fischbach MA, Walsh CT. Chem Rev. 2006; 106:3468–3496. [PubMed: 16895337]
27. Stachelhaus T, Mootz HD, Marahiel MA. Chem Biol. 1999; 6:493–505. [PubMed: 10421756]

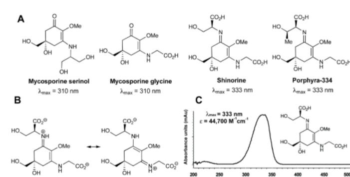

Figure 1. Mycosporines and MAAs

(**A**) Chemical structures of representative mycosporines and MAAs from fungi (mycosporine serinol) and cyanobacteria (mycosporine glycine, shinorine and porphyra-334). (**B**) Shinorine resonance tautomers. (**C**) UV-vis absorbance spectrum of shinorine.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

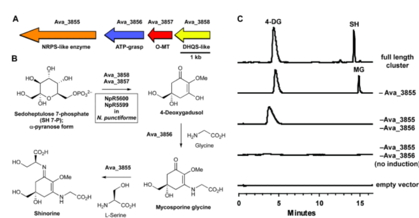

Figure 2. Biosynthesis of shinorine

(**A**) The putative shinorine gene cluster from *A. variabilis*. (**B**) Biosynthetic pathway for the assembly of shinorine from sedoheptulose-7-phosphate. (**C**) HPLC traces of methanol extracts from *E. coli* induced with 500 μM IPTG at 15 °C (310 nm). Abbreviations: SH = shinorine, MG = mycosporine glycine, 4-DG = 4-deoxygadusol.

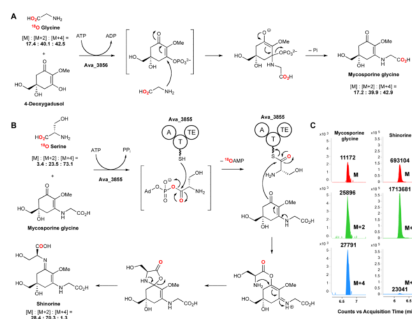

Figure 3. ¹⁸O labeling experiments

(**A**) Reaction of 4-deoxygadusol and ^{18}O glycine with ATP-grasp homolog Ava_3856. (**B**) Reaction of mycosporine glycine and ^{18}O serine with NRPS-like enzyme Ava_3855 (A = adenylation domain, T = thiolation domain, TE = thioesterase domain). (**C**) Extracted ion chromatograms from LC-MS analysis of Ava_3856 and Ava_3855 assays.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript