JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Unprecedented Anthranilate Priming Involving Two Enzymes of the Acyl Adenylating Superfamily in Aurachin Biosynthesis

Dominik Pistorius,[§] Yanyan Li,[§,†] Stéphane Mann,[‡] and Rolf Müller*[,§]

[§]Helmholtz-Institut für Pharmazeutische Forschung Saarland, Helmholtz Zentrum für Infektionsforschung Braunschweig and Pharmazeutische Biotechnologie, Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken, Germany

[‡]Chimie ParisTech, ENSCP, Laboratoire Charles Friedel, UMR 7223-CNRS, 75005 Paris, France

Supporting Information

ABSTRACT: Biosynthesis of many polyketide-derived secondary metabolites is initiated by incorporating starter units other than acetate. Thus, understanding their priming mechanism is of importance for metabolic engineering. Insight into the loading process of anthranilate into the biosynthetic pathway for the quinoline alkaloids aurachins has been provided by the sequencing of a partial biosynthetic gene cluster in the myxobacterium *Stigmatella aurantiaca*. The cluster encodes a predicted aryl:CoA ligase AuaE that was hypothesized to activate and transfer anthranilate to the acyl carrier protein AuaB. However, gene inactivation and *in vitro* experiments described here contradicted this model. Aided by the genome sequence of *S. aurantiaca*, we identified an additional aryl:CoA ligase homologue, AuaEII, encoded in a different gene operon, which is additionally required for anthranilate priming. We report the characterization of both enzymes and the elucidation of a novel non-acetate priming strategy in thio-templated biosynthetic machineries.

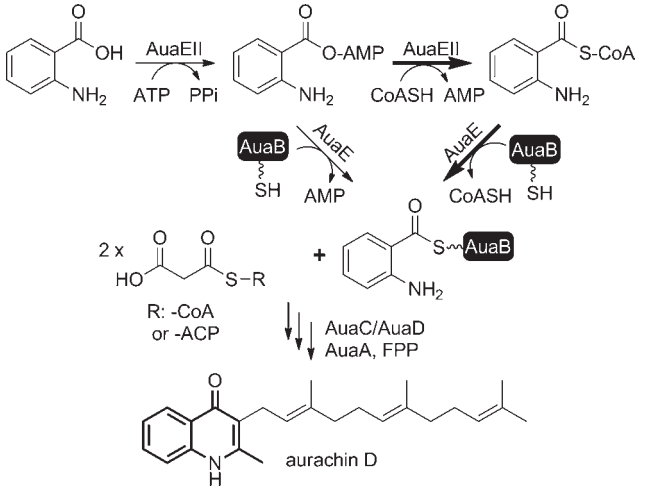
Starter selection and priming is an important process to commence the biosynthesis of polyketide- and non-ribosomal peptide-derived metabolites.[1] In polyketide synthases (PKSs), carboxylic acid starters are typically activated as CoA thioesters and transferred to the loading acyl carrier protein (ACP) by the action of acyltransferases (ATs). For non-acetate starters, dedicated acyl/aryl:CoA ligases are often encoded in the biosynthetic pathway. A different scenario applies for non-ribosomal peptide synthetases (NRPSs), where adenylation (A) domains are responsible for activating and attaching specific amino acids to the phosphopantetheine arm of the peptidyl carrier protein (PCP). CoA ligases and A domains belong to a superfamily of AMP-forming enzymes that also includes firefly luciferases, termed the ANL adenylating superfamily.[2] These enzymes catalyze two independent reactions, sharing the first mechanistic step that yields an acyl-AMP intermediate while differing in the second reaction. We report here the elucidation of an unprecedented anthranilate priming mechanism in aurachin biosynthesis requiring two homologues of aryl:CoA ligases that are responsible for substrate activation and ACP loading, respectively. This novel strategy resembles a hybrid of PKS AT-loading and the NRPS A-domain-mediated priming mechanism.

The aurachins are unique quinoline alkaloids, exhibiting an isoprenyl side chain produced by the myxobacterium *Stigmatella aurantiaca* Sg a15, that inhibit the mitochondrial respiratory chain.[3] We have previously identified a partial gene cluster responsible for aurachin biosynthesis in *S. aurantiaca* Sg a15, which consists of five genes (*auaA*−*auaE*).[4] The functionality of the prenyltransferase AuaA has recently been demonstrated *in vitro*.[5] The remaining encoded proteins show homology to an ACP (AuaB), two β-ketoacyl-ACP synthases II (KAS II, AuaC and AuaD), and a benzoate:CoA ligase (AuaE). Feeding studies[6] and gene inactivation experiments[4] suggested that biosynthesis of the quinoline core structure starts with anthranilate, which is activated and transferred to the ACP by AuaE in an analogous manner, as performed by NRPS A domains. One or both KAS II homologues likely catalyze subsequent condensation of anthraniloyl (AN)-AuaB with two malonyl units (Scheme 1). However, expression of this core *aua* cluster in the heterologous host *Myxococcus xanthus* did not yield any aurachin D, even when the growing culture was supplemented with anthranilate (unpublished data). Experiments employing heterologously expressed AuaB and AuaE proteins, designed to demonstrate the loading of anthranilate to *holo*-AuaB, were similarly not successful (see next paragraph). These findings thus prompted us to search for other candidate genes coding for aryl adenylation activities in the genome of Sg a15. A BLAST search of the draft genome sequences of Sg a15, using AuaE as a probe, allowed identification of a homologous gene (designated *auaEII*) which was found 7.5 kb upstream of the *aua* core cluster. AuaEII shows the highest similarity to a benzoate:CoA ligase from *Geobacter metallireducens* (48% identity/65% similarity). In order to examine its possible involvement in aurachin biosynthesis, targeted gene disruption of *auaEII* was performed. High-pressure liquid chromatography coupled to mass spectrometry (HPLC-MS) analysis of an extract of the mutant revealed that inactivation of gene *auaEII* completely abolished aurachin production (Figure 1), resulting in the same phenotype as the *auaE* knockout mutant described previously.[4] These results confirm that both AuaE and AuaEII are indispensible for the anthranilate loading step and do not represent redundant functionalities.

To elucidate their respective function in detail, both proteins were heterologously expressed in *Escherichia coli* with an N-terminal hexahistidine tag and first assayed *in vitro* for their anthranilate:CoA ligase activity. Reactions were analyzed by HPLC-UV-MS.

Received: April 30, 2011
Published: July 19, 2011

Scheme 1. Proposed Route to Aurachin D Biosynthesis[a]



[a] The quinoline core of aurachin D is highlighted by bold bonds.

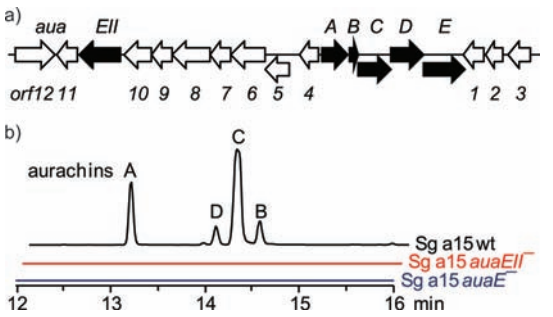


Figure 1. (a) Genetic organization of the *aua* core cluster and surrounding regions. Genes involved in aurachin biosynthesis are depicted in black. Postulated functions encoded by surrounding *orfs* are given in Table S1. (b) HPLC-MS analysis of extracts of *S. aurantiaca* Sg a15 wild type (wt), *auaE*, and *auaEII* knockout mutants. Combined extracted ion chromatograms of 364.3 Da (aurachin D), 380.3 Da (aurachins B and C), and 396.3 Da (aurachin A) are shown.

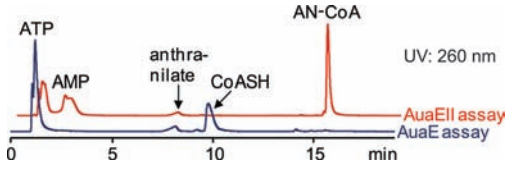


Figure 2. HPLC-UV traces showing *in vitro* CoA ligase activities of AuaE and AuaEII.

Upon incubation with anthranilate, CoASH, ATP, and Mg^{2+}, AuaEII catalyzed the formation of AN-CoA (Figure 2). The identity of AN-CoA was verified by high-resolution (HR)-MS (calcd m/z $[M-H]^-$ = 885.1466, obsd m/z = 885.1462) and NMR (Supporting Information). Further, small amounts of the intermediate AN-AMP could be detected by HPLC-MS, which was confirmed by comparison to a synthetic standard (Supporting Information) and by HR-MS (calcd m/z $[M+H]^+$ = 467.1075, obsd m/z = 467.1080). This intermediate did not accumulate, probably due to rapid turnover. The formation of AN-AMP and AN-CoA was dependent on Mg^{2+}, as its omission and addition of EDTA abolished the production (data not shown). These results confirmed that AuaEII is an authentic anthranilate:CoA ligase. In contrast, AuaE produced neither of these compounds.

Next, we investigated the latter's capacity to load anthranilate to AuaB. Recombinant AuaB was incubated *in vitro* with CoASH and the broad-spectrum phosphopantetheinyl transferase MtaA from *S. aurantiaca* DW4/3-1 to be activated to its *holo* form.[7] After the conversion from *apo* to *holo* was completed, MtaA was removed by anion-exchange chromatography. Purified *holo*-AuaB was incubated with either AuaE, AuaEII, or both under several conditions and was monitored for mass changes of AuaB by HPLC-ESI-MS (Table 1 and Figure 3, *holo*-/AN-AuaB calcd masses 10 194/10 313 Da). An anticipated increase of 119 Da, corresponding to AN tethering on *holo*-AuaB, was observed in the incubation with AuaE and AuaEII in the presence of anthranilic acid, Mg^{2+}, ATP, and CoASH (referred to as standard reaction) at 37 °C for 30 min. Omission of either AuaE or AuaEII led to complete loss of priming. In conjunction with the results obtained for the anthranilate:CoA ligase activity, it is obvious that AuaE is the enzyme that transfers the AN group to *holo*-AuaB from AN-AMP or AN-CoA, both of which are generated by AuaEII. To investigate the substrate nature of AuaE, in one incubation CoASH was omitted from the standard reaction, which should result in AuaEII-mediated formation of the AN-AMP intermediate. For another incubation, AN-CoA was enzymatically prepared and purified from the AuaEII reaction (Supporting Information) to be incubated with *holo*-AuaB and AuaE. In both cases, AN-AuaB was detected. Respective controls without AuaE did not lead to AuaB priming. In contrast to the anthranilate:CoA ligase activity of AuaEII, the transfer reaction catalyzed by AuaE was not dependent on the presence of Mg^{2+}. In order to identify the biologically relevant substrate of AuaE, we chemically synthesized AN-AMP and performed kinetic studies (Supporting Information). The apparent K_m for AN-CoA was 27.8 ± 1.6 μM, whereas AN-AMP exhibited a very high K_m of 20.1 ± 1.8 mM (Figure S2). These results unambiguously prove that AuaE functions as AN-CoA:ACP transacylase. Investigation of AuaE's ability to transfer other CoA substrates to *holo*-AuaB revealed relative activities of 85% for benzoyl-CoA and 24% for buturyl-CoA compared to AN-CoA, and no transfer for malonyl-CoA (Figure S3).[8] These experiments clearly demonstrate that AuaE exhibits a preference for aromatic substrates. Thus, priming of AuaB in aurachin biosynthesis requires concerted actions of two enzymes: AuaEII adenylates anthranilate and produces AN-CoA, and AuaE catalyzes the subsequent transthioesterification reaction to form AN-AuaB.

This priming mechanism is reminiscent of EncN-mediated loading of benzoate in enterocin biosynthesis;[9] however, involvement of two distantly located CoA ligase homologues is unprecedented. The ANL superfamily of adenylating enzymes is defined by a conserved adenylate binding motif. The members are structural homologues, despite sharing weak sequence homology (20−40% identity). In the case of CoA ligases and A domains of NRPSs, synthesis of acyl or aryl thioesters is carried out in a two-step reaction. The first half-reaction involves adenylation of the carboxylic group of the substrate to form an acyl- or aryl-AMP intermediate. In the second half-reaction, the thiol group of CoA or of the phosphopantetheinyl arm of the PCP attacks the acyl carbon, with subsequent elimination of AMP. To catalyze the concerted two- step reaction, a dramatic conformational

Table 1. Loading of Anthranilate to *holo*-AuaB under Different Conditions

reactants	presence in the respective incubations							
holo-AuaB	+	+	+	+	+	+	+	+
AuaE	+	–	+	+	+	–	+	+
AuaEII	+	+	–	+	+	–	+	–
CoASH	+	+	+	–	+	–	–	–
ATP	+	+	+	+	–	–	–	–
anthranilate	+	+	+	+	+	–	–	–
AN-CoA	–	–	–	–	–	+	+	+
observed AuaB species	AN-AuaB	*holo*-AuaB	*holo*-AuaB	AN-AuaB	*holo*-AuaB	*holo*-AuaB	AN-AuaB	AN-AuaB

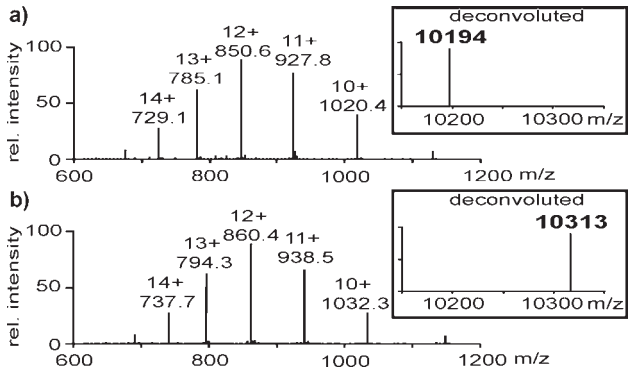


Figure 3. ESI-MS spectra of (a) *holo*-AuaB and (b) AN-AuaB.

Figure 4. Alignment of the core A10 region of the ANL superfamily of adenylate-forming enzymes. The diamond indicates the lysine residue that is essential for formation of the adenylate intermediate. For details see Figure S4.

change is necessary after the adenylation half-reaction has been completed. This domain alteration has been observed in several structures of ANL adenylating enzymes.[10] We performed in-depth sequence analyses of AuaE and AuaEII in order to explain the discrepancy in their activities. AuaEII conserves all nine core motifs typical for ANL adenylating enzymes (Figure S4).[11] In contrast, AuaE shows strong deviations from the consensus in core A3 (Figure S4) and, most remarkably, lacks the A10 motif (Figure 4). This truncation might potentially have been caused by a recombination event, as there is a transcriptional overlap of the endings of genes *auaE* and *orf1*, which are transcribed in opposite directions.[4] The A3 core is responsible for binding and orientation of the β- and γ-phosphates of ATP.[12] The A10 core is implicated in the conformation change, where it is rotated away from the active site for the second half-step.[10] Further, mutagenesis studies confirmed that the invariable conserved lysine residue in the A10 motif plays an indispensible role in the adenylation half-reaction.[13] Taken together, these sequence modifications explain the deficiency of AuaE in adenylating anthranilate.

To the best of our knowledge, AuaE is the first naturally occurring enzyme in the ANL superfamily that catalyzes only the second half-reaction (thioester formation, in this case also transthioesterification). AuaE, which has obviously evolved from AN-CoA ligases, is functionally equivalent to PKS ATs, which belong to the unrelated α/β hydrolase superfamily.[14] Such convergent evolution may be advantageous, as few evolution events would be required to confer substrate specificity. This view is supported by analysis of the 10-amino acid code that predicts specificity of NRPS A domains.[15] The specificity-conferring amino acid residues of AuaE are in very good agreement with those extracted from known bacterial anthranilate-activating enzymes (Table S3). Notably, one of the hallmark features of binding to anthranilate proposed for fungal A domains (Gly235) is also conserved for AuaE.[16]

The two proteins AuaE and AuaEII show only modest sequence homology with each other (30% identity/47% similarity), which argues against a gene duplication event. The gene *auaEII* is located in an operon directly upstream of the *auaA*–*auaE* cluster of genes, which is transcribed in the reverse direction. The proteins encoded by *orf4*–*orf9* might be involved in the regulation of aurachin biosynthesis and the transport of the final product out of the cell. Being the only copy of aryl-CoA ligase encoded in the Sg a15 genome, AuaEII may also be involved in the degradation of aromatic compounds.[17] A previous study has identified a dedicated plant-type 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase for production of the precursor anthranilate located at yet another locus in the Sg a15 genome.[18] Owing to the unusual split gene organization, these findings suggest that aurachin biosynthesis is highly coordinated and involves cross-talks between primary and secondary metabolism. Deciphering its regulation network is thus of considerable interest. Interestingly, the enzymes required for the formation of aurachins C, B, and A are not found within the aurachin core gene cluster.

Characterization of AuaE and AuaEII led to revision of the initially proposed biosynthetic route to aurachin D (Scheme 1).[4,19] Complete *in vitro* validation of quinoline biosynthesis has been hampered so far by difficulties in obtaining soluble AuaC protein, despite numerous efforts. Nevertheless, we performed loading experiments of AuaB with malonyl-CoA in order to clarify the identity of the extender unit. As AuaB is the sole ACP encoded in the aurachin core biosynthetic gene cluster, it is conceivable that malonyl-AuaB would be the substrate for two rounds of PKS extension to form the quinoline core. Inspired by the type II PKS systems, the *fabD* gene encoding a malonyl-CoA:ACP transacylase

(MCAT) involved in fatty acid biosynthesis in the closely related and completely sequenced strain *S. aurantiaca* DW4/3-1 was identified, cloned, and heterologously expressed. Upon incubation of FabD and malonyl-CoA at 37 °C for 30 min, 80% of *holo*-AuaB was converted to malonyl-AuaB, as revealed by HPLC-MS (malonyl-AuaB, calcd and obsd mass 10 280 Da). To investigate the priming preference of AuaB, the protein was incubated with identical concentrations of malonyl-CoA and AN-CoA in presence of equal molarities of FabD and AuaE. Protein concentrations were precisely determined on a protein chip by a 2100 Bioanalyzer (Agilent). This experiment revealed a clear preference for anthraniloyl over malonyl loading onto *holo*-AuaB (86% versus 14%) under the conditions used. Thus, similar to the enterocin type II PKS, malonyl-AuaB may be the extender unit, and the partition of primer and extender CPs could possibly be determined by the ratio of respective ATs.[20] However, due to the lack of complete *in vitro* reconstitution of quinoline biosynthesis, we cannot exclude malonyl-CoA or malonyl tethered on a second ACP as extender unit.

In conclusion, we have elucidated a novel non-acetate priming mechanism divided between a CoA ligase and an atypical acyltransferase from the ANL superfamily, which should inspire strategies of metabolic engineering to incorporate novel starters into thio-templated biosynthetic pathways.

■ ASSOCIATED CONTENT

Supporting Information. Aurachin structures, protein accession numbers, detailed experimental procedures, and supporting figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
rom@mx.uni-saarland.de

Present Addresses
[†]Unité "Molécules de Communication et Adaptation des Micro-organismes", UMR 7245 CNRS/Muséum National d'Histoire Naturelle, 57 rue Cuvier, 75005 Paris, France

■ ACKNOWLEDGMENT

This work was supported by DFG Priority Program 1152 "Evolution of Metabolic Diversity".

■ REFERENCES

(1) Moore, B.; Hertweck, C. *Nat. Prod. Rep.* **2002**, *19*, 70–99.
(2) Gulick, A. M. *ACS Chem. Biol.* **2009**, *4*, 811–827.
(3) Kunze, B.; Höfle, G.; Reichenbach, H. *J. Antibiot.* **1987**, *40*, 258–265.
(4) Sandmann, A.; Dickschat, J.; Jenke-Kodama, H.; Kunze, B.; Dittmann, E.; Müller, R. *Angew. Chem., Int. Ed.* **2007**, *46*, 2712–2716.
(5) Stec, E.; Pistorius, D.; Müller, R.; Li, S. M. *ChemBioChem* **2011**, *12*, 1724–1730.
(6) Höfle, G.; Kunze, B. *J. Nat. Prod.* **2008**, *71*, 1843–1849.
(7) Gaitatzis, N.; Hans, A.; Müller, R.; Beyer, S. *J. Biochem.* **2001**, *129*, 119–124.
(8) Kalaitzis, J. A.; Izumikawa, M.; Xiang, L.; Hertweck, C.; Moore, B. S. *J. Am. Chem. Soc.* **2003**, *125*, 9290–9291.
(9) Izumikawa, M.; Cheng, Q.; Moore, B. S. *J. Am. Chem. Soc.* **2006**, *128*, 1428–1429.
(10) (a) Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C. *Biochemistry* **2003**, *42*, 2866–2873. (b) Reger, A. S.; Wu, R.; Dunaway-Mariano, D.; Gulick, A. M. *Biochemistry* **2008**, *47*, 8016–8025. (c) Kochan, G.; Pilka, E. S.; von Delft, F.; Oppermann, U.; Zue, W. W. *J. Mol. Biol.* **2009**, *388*, 997–1008.
(11) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. *Chem. Rev.* **1997**, *97*, 2651–2674.
(12) Chang, K. H.; Xiang, H.; Dunaway,-Mariano, D. *Biochemistry* **1997**, *36*, 15650–15659.
(13) (a) Horswill, A. R.; Escalante-Semerena, J. C. *Biochemistry* **2002**, *41*, 2379–2387. (b) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Anderson, S. M. *Biochemistry* **2000**, *39*, 5433–5440.
(14) Keatinge-Clay, A. T.; Shelat, A. A.; Savage, D. F.; Tsai, S. C.; Miercke, L. J.; O'Connell, J. D.; Khosla, C.; Stroud, R. M. *Structure* **2003**, *11*, 147–154.
(15) (a) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493–505. (b) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211–224.
(16) Ames, B. D.; Walsh, C. T. *Biochemistry* **2010**, *49*, 3351–3365.
(17) Schühle, K.; Jahn, M.; Ghisla, S.; Fuchs, G. *J. Bacteriol.* **2001**, *183*, 5268–5278.
(18) Silakowski, B.; Kunze, B.; Müller, R. *Arch. Microbiol.* **2000**, *173*, 403–411.
(19) Li, Y.; Müller, R. *Phytochemistry* **2009**, *70*, 1850–1857.
(20) Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. *Nat. Chem. Biol.* **2007**, *3*, 557–558.