Thioester reduction and aldehyde transamination are universal steps in actinobacterial polyketide alkaloid biosynthesis†

U. R. Awodi,‡ J. L. Ronan,‡ J. Masschelein,‡ E. L. C. de los Santos and G. L. Challis*

Actinobacteria produce a variety of polyketide alkaloids with unusual structures. Recently, it was shown that a type I modular polyketide synthase (PKS) is involved in the assembly of coelimycin P1, a polyketide alkaloid produced by Streptomyces coelicolor M145. However, the mechanisms for converting the product of the PKS to coelimycin P1 remain to be elucidated. Here we show that the C-terminal thioester reductase (TR) domain of the PKS and an ω-transaminase are responsible for release of the polyketide chain as an aldehyde and its subsequent reductive amination. Bioinformatics analyses identified numerous gene clusters in actinobacterial genomes that encode modular PKSs with a C-terminal TR domain and a homolog of the ω-transaminase. These are predicted to direct the biosynthesis of both known and novel polyketide alkaloids, suggesting that reductive chain release and transamination constitutes a conserved mechanism for the biosynthesis of such metabolites.

Introduction

An astonishing array of structurally-diverse polyketide metabolites with numerous applications in medicine, animal health and agriculture are produced by actinobacteria. Remarkable molecular machines known as type I modular polyketide synthases (PKSs) are responsible for the biosynthesis of the overwhelming majority of these natural products. Modular PKSs employ several distinct strategies for the creation of structural diversity, including the utilization of a variety of starter and extender units for assembly of the polyketide chain, diverse mechanisms for release of the fully-assembled chain from the PKS, and a wide range of on- and post-PKS tailoring reactions. Among these, the mechanism of chain release is arguably of prime importance in defining which structural class a metabolite belongs to. Thus, in macrolide biosynthesis chain release proceeds via macrolactonization, whereas in ansamycin biosynthesis the chain is released via macro lactamation; polyether biosynthesis employs hydrolytic chain release; (spiro)tetronates are assembled by condensation of the polyketide chain with a glyceryl thioester; and chain release in prodigine biosynthesis proceeds via nucleophilic attack of an amino acid ω-carbanion equivalent (see ESI† for further details).

Incorporation experiments with isotope-labeled precursors have shown that several mono-, di-, tri- and tetracyclic alkaloids produced by actinobacteria, including latumycin 1 (also known as abikoviromycin), nigrifacin 2, pyrindicin 3, cyclizidine 4, and streptazolin 5 (ref. 11) (Fig. 1) are derived from polyketide precursors. Several other actinobacterial alkaloids, such as streptazine E 6, 4-hydroxy[2-penta-1,3-dienyl]-piperidine 7, and indolizomycin 8 (ref. 14) can similarly be hypothesized to have a polyketide origin. In 2012, we reported the identification of the unusual alkaloid coelimycin P1 9 (Fig. 1) as the metabolic product of a cryptic polyketide (gpk) biosynthetic gene cluster in Streptomyces coelicolor M145. This

Fig. 1 Structures of polyketide alkaloids from actinobacteria.

Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK. E-mail: g.l.challis@warwick.ac.uk
† Electronic supplementary information (ESI) available: Experimental procedures and additional tables, figures, chromatograms and results from biological assays. See DOI: 10.1039/c6sc02803a
‡ These authors contributed equally.
provided the first biochemical insights into the molecular mechanisms underlying actinobacterial polyketide alkaloid biosynthesis.

The \textit{cpk}A, \textit{cpk}B and \textit{cpk}C genes within the coelimycin biosynthetic gene cluster encode a hexamodular PKS, which is proposed on the basis of the predicted specificities of its catalytic domains to assemble thioester (Fig. 2). This intermediate is thought to be released from the PKS by NAD(P)H-mediated reduction to the corresponding aldehyde, catalyzed by the thioester reductase (TR) domain at the C-terminus of CpkC. Reductive amination of aldehyde, catalyzed by the putative \(\omega\)-transaminase encoded by \textit{cpk}G, is proposed to give amine (or its (Z)-3,4 isomer, depending on the timing of \(\Delta\)-2,3 isomerization). Elaboration of 12 to 9 is believed to proceed via a series of oxidations, a cyclodehydration and bis-epoxide trapping by N-acetyl-L-cysteine, as illuminated by incorporation of labelled precursors.

**Results and discussion**

To experimentally investigate the role played by the CpkC TR domain in coelimycin biosynthesis, it was overproduced in \textit{E. coli} as an N-terminal His\(_6\) fusion protein and purified to homogeneity using nickel affinity chromatography. UHPLC-ESI-Q-TOF-MS analysis confirmed that the purified protein had the expected molecular weight (see ESIf). The catalytic activity of the CpkC TR domain was investigated using octanoyl-CoA as a mimic of the proposed natural substrate. NADH and NADPH consumption was examined by monitoring the decrease in absorbance at 340 nm. A substantial decrease in absorbance with time was observed when NADH was used, whereas no decrease in absorbance was observed for NADPH (Fig. 3), showing that octanoyl-CoA is reduced by the former but not the latter. No NADH consumption was observed in control reactions from which the enzyme was omitted (Fig. 3).

In the absence of CpkG, which is proposed to intercept the octanal formed by reduction of octanoyl-CoA via transamination, CpkC-TR could catalyze further reduction of 14 to 1-octanol (Scheme 1). Thus, the reaction mixture containing CpkC-TR, octanoyl-CoA and NADH was extracted with chloroform and derivatized with \(N_2O\)-bis-(trimethylsilyl)acetamide (BSA) to convert all alcohols present to the corresponding tri-methylsilyl (TMS) ether derivatives. GC-MS analysis of the derivatized extract identified an analyte, absent from a control reaction containing the boiled enzyme, with \(m/z = 187.2\). This corresponds to the fragment ion of trimethylsilyl-1-octanol resulting from neutral loss of a methyl group and was found to

**Fig. 2** The proposed roles played by the CpkABC modular PKS and the CpkG \(\omega\)-transaminase in coelimycin P1 biosynthesis.

**Fig. 3** Analysis of NAD(P)H consumption in the CpkC-TR-catalyzed reduction of octanoyl-CoA using a continuous spectrophotometric assay.

**Scheme 1** Reactions catalyzed by CpkC-TR using octanoyl-CoA and octanal as substrate analogues and conversion of 1-octanol to its TMS ether to facilitate GC-MS analysis.
have the same retention time as the m/z = 187.2 ion resulting from GC-MS analysis of BSA-derivatised 1-octanol (Fig. 4). An analogous experiment in which octanal 14 was used in place of octanoyl-CoA 13 gave a similar result (see ESI†). Taken together, the data show that CpkC-TR catalyzes the NADH-mediated reduction of octanoyl-CoA 13 to octanal 14, which is further reduced to 1-octanol 15 in the absence of a suitable transaminase to trap the transient aldehyde intermediate.

We next turned our attention to CpkG which shows sequence similarity to pyridoxal-5′-phosphate (PLP)-dependent ω-transaminases. CpkG was overproduced in E. coli and purified as described for CpkC-TR. UHPLC-ESI-Q-TOF-MS analysis confirmed that the purified protein had the molecular weight expected for apo-CpkG and gel filtration chromatography showed that it is a homodimer (see ESI†). To investigate the ability of PLP to bind to CpkG, the UV-vis absorbance spectrum of the enzyme purified in the presence of the cofactor was compared to that of PLP alone. At pH 8, a characteristic λmax shift from 410 nm to 418 nm was observed, indicating that PLP forms an aldimine with CpkG (see ESI†). We initially investigated whether holo-CpkG is able to catalyze the conversion of amines 17–21 to the corresponding aldehydes/ketones, along with concomitant conversion of either pyruvate 22 or ω-keto glutarate 23 to alanine 24 or glutamic acid 25, respectively (Scheme 2). A spectrophotometric assay that detects the production of amino acids via formation of a copper(II) complex was used. All of the amines tested were found to stimulate conversion of the two ω-keto acids to the corresponding amino acids, demonstrating that CpkG has broad substrate tolerance (see ESI†). The highest level of activity was observed for octylamine 21 using pyruvate 22 as the co-substrate, indicating that CpkG prefers medium chain alkylamines, consistent with its proposed role in coelimycin biosynthesis. The production of octanal 15 in this reaction was confirmed by derivatization with d-cysteine and LC-MS comparison with an authentic standard of the resulting thiazolidine (see ESI†).

The stereochemical outcome of the CpkG-catalyzed reaction of octylamine 21 with pyruvate 22 was investigated by derivatizing the alanine 24 produced with Marfey’s reagent. LC-MS comparisons with Marfey’s derivatives of l- and d-alanine revealed that only l-alanine is formed (see ESI†), suggesting that CpkG has preference for l-amino acid substrates.

Having established that octylamine 21 and pyruvate 22 are good substrates for CpkG in the reverse of the proposed biochemical reaction, we investigated reductive amination of octanal 14 using alanine 24 as an amino donor. The pyruvate 22 produced was detected using a coupled assay with lactate dehydrogenase and NADH (see ESI†) and octalamine 21 formation was confirmed by LC-MS comparisons with an authentic standard (Fig. 5). Despite the exclusive production of l-alanine in the reverse reaction (see above), CpkG was found to accept both l- and d-alanine as amino donors in the forward reaction.

Taken together, our data strongly support the key proposed roles played by CpkC-TR and CpkG in diverting the product of a modular PKS into alkaloid metabolism. We thus reasoned that homologous enzymes are likely to be ubiquitous in bacterial polyketide alkaloid biosynthesis. To investigate this hypothesis, we searched the GenBank database for gene clusters containing genes encoding a type I modular PKS with a TR domain and a CpkG homolog within 35 kb of each other (see ESI† for details). Functional annotation of the sequences obtained using antiSMASH resulted in the identification of 22

Scheme 2 Amines tested as substrates for CpkG, alongside pyruvate and ω-keto glutarate as co-substrates.

Fig. 4 Extracted ion chromatograms at m/z = 187.2 (corresponding to the fragment ion of 16 resulting from loss of a methyl group) from GC-MS analyses of an authentic standard of 16 (top), the BSA-derivatized product of the CpkC-TR-catalyzed reduction of 13 with NADH (middle) and the BSA-derivatized product of a control reaction containing boiled enzyme (bottom).

Fig. 5 Extracted ion chromatograms at m/z = 130.1590 ± 0.005 (corresponding to the [M + H]+ ion for 21) from UHPLC-ESI-Q-TOF-MS analyses of an authentic standard of 21 (top), the CpkG-catalyzed reaction of 14 with l-alanine (middle) and a control reaction from which the enzyme was omitted (bottom).
In conclusion, we have shown that reductive chain release and transamination are the key steps in the conversion of a modular PKS product to the unusual alkaloid coelimycin P1. Moreover, a further 22 gene clusters in actinobacteria that are known or predicted to direct the biosynthesis of diverse polyketide alkaloids have been identified. All of these encode a TR-terminated modular PKS and a CpkG homolog, suggesting a common enzymatic logic for actinobacterial polyketide alkaloid biosynthesis.

Acknowledgements

This research was supported by a Chancellor’s International Scholarship (to U. R. A.), a Chancellor’s Scholarship (to J. L. R.), a Marie Skłodowska-Curie Fellowship (to J. M., contract no.
Notes and references