Article

CellPress

Biosynthesis of Crocacin Involves an Unusual Hydrolytic Release Domain Showing Similarity to Condensation Domains

Stefan Müller,[1] Shwan Rachid,[1,3] Thomas Hoffmann,[1] Frank Surup,[2] Carsten Volz,[1] Nestor Zaburannyi,[1] and Rolf Müller[1,*]

[1]Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Saarland 66123, Germany
[2]Department of Microbial Drugs, Helmholtz Centre for Infection Research, Braunschweig, Niedersachsen 38124, Germany
[3]Present address: Faculty of Science and Health, University of Koya, Koya, Kurdistan KOY45, Iraq
*Correspondence: rolf.mueller@helmholtz-hzi.de
http://dx.doi.org/10.1016/j.chembiol.2014.05.012

SUMMARY

The crocacins are potent antifungal and cytotoxic natural compounds from myxobacteria of the genus *Chondromyces*. Although total synthesis approaches have been reported, the molecular and biochemical basis guiding the formation of the linear crocacin scaffold has remained unknown. Along with the identification and functional analysis of the crocacin biosynthetic gene cluster from *Chondromyces crocatus* Cm c5, we here present the identification and biochemical characterization of an unusual chain termination domain homologous to condensation domains responsible for hydrolytic release of the product from the assembly line. In particular, gene inactivation studies and in vitro experiments using the heterologously produced domain CroK-C2 confirm this surprising role giving rise to the linear carboxylic acid. Additionally, we determined the kinetic parameters of CroK-C2 by monitoring hydrolytic cleavage of the substrate mimic *N*-acetyl-cysteaminyl-crocacin B using an innovative high-performance liquid chromatography mass spectrometry-based assay.

INTRODUCTION

Many pharmaceutically relevant natural products are produced by terrestrial microorganisms (Chin et al., 2006), including numerous compounds isolated from actinobacteria (Lucas et al., 2013) and myxobacteria (Wenzel and Müller, 2009) still holding great promise for future drug candidates (Müller and Wink, 2014). One example is the crocacins, which were isolated from *Chondromyces crocatus* Cm c5 and found to exhibit strong antifungal and cytotoxic activities (Jansen et al., 1999). Although chemical syntheses have been described, it is still difficult to produce crocacins and their analogs via the synthetic route, because of the complexity of the whole molecule. One alternative path forward is biosynthetic engineering of the producer strain, for which we have been able to set the stage in recent

efforts dealing with other bioactive compounds found in the multiproducer *C. crocatus* Cm c5. During this work, genetic tools were developed, and the biosynthesis of chondramides (Rachid et al., 2006), ajudazols (Buntin et al., 2008), chondrochlorens (Rachid et al., 2009), thuggacins (Buntin et al., 2010), and croca-peptins (Viehrig et al., 2013) was studied, revealing a number of exceptional biosynthetic mechanisms and demonstrating possibilities for biosynthetic engineering in this species.

As the basis for biosynthesis, mostly large multifunctional enzyme complexes such as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) account for product formation via successive assembly of these secondary metabolites from small building blocks (Hertweck, 2009; Sieber and Marahiel, 2005). A minimal type I PKS module consists of a β-ketoacyl synthase (KS) an acyltransferase (AT) plus an acyl carrier protein (ACP) collectively expanding the polyketide chain via Claisen condensation with an extender unit, a process similar to fatty acid biosynthesis (Hopwood, 1997). Similarly, an NRPS module is composed of different domains performing distinct functions. In detail, an adenylation (A) domain activates a specific amino acid in an ATP-dependent manner and subsequently transfers it to the 4′-phosphopantetheine (Ppant) arm of a peptidyl carrier protein (PCP). Finally, a condensation (C) domain catalyzes amide bond formation between two adjacent PCP-bound amino acids to yield the dipeptide. Once the growing intermediate reaches its full length, it must be released from the last carrier protein. This reaction is commonly catalyzed by a thioesterase (TE) domain. The TE domain-mediated product release is the most widespread mechanism found in natural product biosynthesis. Usually, a TE is located at the end of a PKS or NRPS assembly line and contains a conserved catalytic triad (Ser-His-Asp) required for product release (Tsai et al., 2002). Particularly, the assembled compound is transferred from the last carrier protein to the active site Ser of the TE, thus forming the respective ester bound to the TE domain. Depending on the characteristics of the TE, this intermediate undergoes either hydrolysis or macrocyclization. In the first case, the ester is hydrolyzed by the nucleophilic attack of water, resulting in a linear carboxylate, as reported for numerous natural products, such as the δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine tripeptide (Byford et al., 1997) and bacillaene (Butcher et al., 2007; Moldenhauer et al., 2007). Contrary to this, macrocyclization occurs when the respective TE-ester undergoes an

intramolecular transesterification with a suitable hydroxyl, amino, or thiol group, as reported for the thuggacins (Buntin et al., 2010), nystatin (Brautaset et al., 2000), thiocoraline (Robbel et al., 2009), and many others.

However, nature has evolved several alternative chain termination strategies involving terminal reductase (Red) domains and terminal C domain variants leading to altered products (Du and Lou, 2010). Chain termination catalyzed by Red domains demonstrates an additional way to release a product from the Ppant moiety of its cognate carrier protein. In general, these domains are capable of catalyzing two- and four-electron reduction of the respective carrier protein-bound thioester, thereby releasing the former ester as aldehyde or alcohol. The first biochemically characterized example thereof was described by the Walsh group in the context of lysine biosynthesis (Ehmann et al., 1999). Primal observation of a similar mechanism in natural product biosynthesis was reported for myxochelin formation (Gaitatzis et al., 2001). Additional genetic and biochemical investigations proved that the PCP-bound intermediate is reduced to the aldehyde state, which is either trapped in a reductive aminotransferase reaction, giving rise to the terminal amine in myxochelin B, or further reduced to the alcohol, as found in myxochelin A (Gaitatzis et al., 2001; Li et al., 2008).

As an additional alternative, variants of C domains from NRPS systems are as well capable of performing versatile chain termination reactions. The most common route is typically found in fungi, in which terminal C domains catalyze macrocyclization of the respective product through nucleophilic attack of the PCP-bound thioester by an internal amine (Keating et al., 2001; Gao et al., 2012). Accordingly, the biosynthesis of the fungal metabolites cyclosporin (Weber et al., 1994), enniatin (Pieper et al., 1995), HC-toxin (Scott-Craig et al., 1992), and PF1022A (Weckwerth et al., 2000) is believed to follow such a mechanism. Noteworthy, this chain termination strategy is not restricted to the formation of an intramolecular amide but can also occur between the carrier protein-bound carboxyl group of the assembled compound and a primary amine of an acceptor substrate, as found in several biosynthetic pathways, such as rapamycin (Schwecke et al., 1995), FK506 (Motamedi and Shafiee, 1998), aeruginosin (Ishida et al., 2007), and streptothricin (Maruyama et al., 2012). Additionally, esterification displays a second and rather uncommon method of C domain-mediated chain release. The first C domain variant that was shown to catalyze esterification in vitro is that of the PCP-C didomain protein Fum14p, involved in fumonisin biosynthesis (Zaleta-Rivera et al., 2006). Particularly, this survey suggests that the C domain of Fump14p mediates chain release of a tricarboxylic PCP-bound intermediate while generating its fumonisin esters. Another instance of a C domain-catalyzed chain release via transesterification was reported for the free-standing condensation enzyme SgcC5 taking part in the biosynthesis of C-1027 (Lin et al., 2009). In that study, the authors experimentally demonstrated that SgcC5 transfers the PCP-bound (S)-3-chloro-5-hydroxy-β-tyrosyl intermediate onto the enediyne core mimic (R)-1-phenyl-1,2-ethanediol while forming the respective ester. Furthermore, they stated that this enzyme still retained reduced canonical amide-forming activity of typical NRPS C domains, thereby providing an evolutionary link between amide- and ester-forming condensation enzymes.

On the basis of current knowledge, it is increasingly evident that the catalytic activity of C domain variants of NRPS machineries expands behind the classic formation of an amide bond: Along with the identification and functional analysis of the crocacin biosynthetic gene cluster from *C. crocatus* Cm c5, we herein demonstrate a C domain variant exhibiting hydrolytic activity, which is responsible for the formation of the free carboxylic acid moiety of crocacin B. Thus, this study adds a third function to this enzyme class and additionally provides interesting insights for genetic engineering approaches to generate novel natural products.

RESULTS AND DISCUSSION

Identification of the Crocacin Biosynthetic Gene Cluster
The crocacins are linear natural products most likely derived from a NRPS/PKS hybrid biosynthetic gene cluster. A retrobiosynthetic analysis of the structure suggested two NRPS modules being specific for the incorporation of glycine. On the basis of this assumption, we used the antiSMASH pipeline (Medema et al., 2011) to screen a draft genome of *C. crocatus* Cm c5, which we achieved previously by performing several sequencing attempts for the presence of such a biosynthetic gene cluster. In fact, this search revealed a distinct DNA segment on the chromosome matching the required criteria, and a detailed in silico analysis strongly suggested this region as the crocacin biosynthetic gene cluster (Figure 1). To unambiguously assign this DNA segment and the encoded functions to crocacin biosynthesis, the suicide vector pSUP_croI_KO was introduced into the genome of *C. crocatus* Cm c5 (Figure S2 available online). Integration of this plasmid into the genome via single crossover should lead to the inactivation of *croI*, a hybrid PKS-NRPS gene that should be essential for crocacin biosynthesis and furthermore might affect the expression of the subsequent genes because of polar effects. As expected, analysis of the resulting mutants Cm c5::pSUP_croI_KO by high-performance liquid chromatography mass spectrometry (HPLC-MS) showed abolishment of crocacin production and thereby confirmed the participation of this genomic region in crocacin biosynthesis (Figure 2). This result, taken together with additional experiments described below, clearly confirmed our assumption that this DNA segment indeed bears the encoded functions responsible for crocacin biosynthesis. The respective DNA fragment was therefore referred to as the *cro* locus.

Sequence Analysis of the Crocacin Biosynthetic Gene Cluster and the Contiguous Regions
The *cro* locus has a size of about 57 kb, with a GC content of 67.3%. It contains four PKS (*croA-C* and *croJ*), one NRPS (*croK*), and one hybrid PKS-NRPS (*croI*) gene (Figure 1A; Table 1). The β-branching unit consists of five genes (*croD-H*) and is flanked by *croC* and *croI*. Interestingly, the assembly line stops with a C domain (CroK-C2), but no TE is found. Moreover, no TE domain seems to be encoded in the whole gene cluster, as observed in an in-depth in silico analysis of the whole region (see below).

At this point, we focused on PKS-NRPS-related domains regarding the conservation of characteristic motifs and key amino acid residues. First, the substrate specificity of all AT
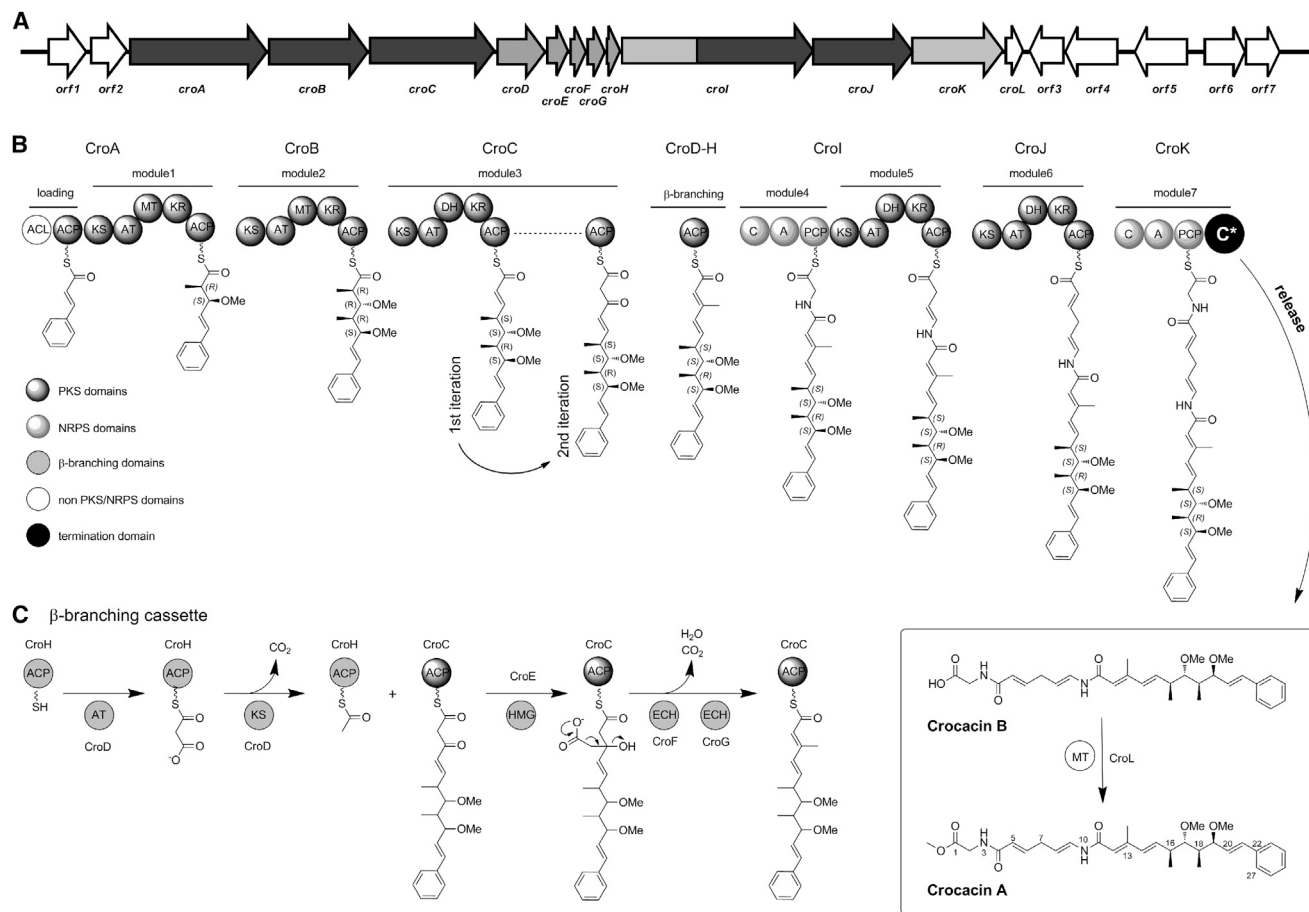
Figure 1. Gene Cluster and Proposed Pathway for Biosynthesis of Crocacin in *C. crocatus* Cm c5
(A) Genetic organization of the crocacin gene cluster.
(B) Proposed biosynthetic pathway for crocacin production in *C. crocatus* Cm c5.
(C) Proposed process of β-branching at C13. ECH, enoyl-CoA hydratase homolog.

domains was evaluated as described previously (Yadav et al., 2003). Each AT domain contains the consensus motif GXSXG, bearing the catalytic serine residue in its center, and can therefore be considered as an active domain. With respect to their substrate specificity, CroA-AT and CroB-AT are predicted to load methylmalonyl-coenzyme A (CoA), while the remaining ones use malonyl-CoA. This is in agreement with the observed structure of the crocacins. Furthermore, the gene cluster contains three dehydratase (DH) domains, which all should be active as judged by the presence of the common consensus sequence of the respective active site (Keatinge-Clay, 2008). The analysis of all five ketoreduction (KR) domains confirmed the presence of the exact Rossman fold motif GXGXXG for CroA-KR, CroI-KR, and CroJ-KR. The other two KR domains deviate from this motif at the second conserved position. In particular, CroB-KR contains a serine and CroC-KR features an aspartate. Because in both cases the deviating amino acid is followed by a glycine, this adjustment should still provide enough space for NADH binding, and therefore these domains are considered functional. The substrate specificity of the two adenylation (A) domains CroI-A and CroK-A was predicted by analyzing their primary sequences with the NRPSpredictor2

tool (Röttig et al., 2011). After this analysis, both A domains should have specificity for the amino acid glycine. Three methyltransferases (MTs) are encoded in the *cro* locus. CroA-MT and CroB-MT are part of the respective modules 1 and 2, while the third one is encoded by the single standing gene *croL*. All domains can be classified as O-MTs when compared with the described conserved motifs (Ansari et al., 2008). A rather unusual feature of the *cro* cluster is the β-branching cassette encoded by *croD-H*. Commonly, such systems consist of a hydroxymethylglutaryl-CoA synthase homolog (HMG), two enoyl-CoA hydratase homologs, and in each case a free-standing KS, AT, and ACP domain (Calderone, 2008). In the present case, the KS and AT are located on the bifunctional protein CroD. Additionally, this KS does not feature the typical Cys-to-Ser mutation in the active site, which hinders it from taking part in chain elongation but still allows decarboxylation of the extender unit. These differences might be because CroD is a relic of a former PKS complex now taking part in β-branching. Apart from that, no further deviations affecting the β-branching cassette seem to be present.

An investigation of the surrounding region of the *cro* cluster revealed two open reading frames (*orfs*) upstream of *croA*.
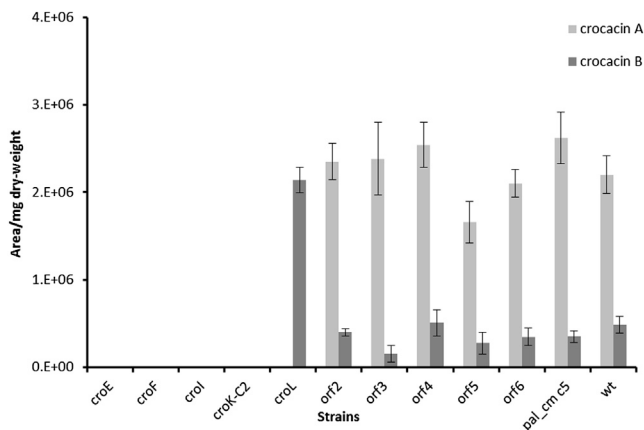
Figure 2. Production of Crocacin A and Crocacin B of Various Mutants and the Wild-Type Strain

Mutants are altered in the indicated gene. All LC-MS-derived area values are normalized to the cell dry weight of each sample. The depicted values are mean values from three independent mutants. Error bars show calculated SDs. wt, wild-type.

A BLAST analysis suggests that *orf1* and *orf2* encode putative proteins comprising an ABC-1 domain. These domains are often found in protein kinases involved in signal transduction and regulation of cellular processes (Leonard et al., 1998), supporting the idea that these proteins might be involved in regulation of crocacin biosynthesis. A similar adjustment of genes was already reported for the rhizopodin gene cluster, but no putative function in the biosynthesis could be assigned (Pistorius and Müller, 2012). To assign or to exclude a function of *orf1* and *orf2* in crocacin biosynthesis, we aimed to disrupt both by integrating derivatives of pSUP_Hyg harboring a homology fragment of the respective loci into the genome of *C. crocatus* Cm c5. Unfortunately, we could only obtain mutants altered in *orf2*. These mutants showed no significant changes in crocacin production when compared with the wild-type strain, indicating that the gene product is not involved in crocacin production under the applied conditions (Figure 2).

The downstream region of *croL* comprises five putative genes (*orf3*–*orf7*). Two of them, *orf3* and *orf4*, lie on the opposite strand and encode a putative cytochrome P450 (Orf3) and a putative long-chain fatty acid-CoA ligase (Orf4). To clarify their function, we disrupted both genes in separate attempts. Resulting mutants showed no significant changes in crocacin production when compared with the wild-type (Figure 2). Therefore, these genes are not considered to be associated with crocacin biosynthesis. Gene *orf5* encodes a putative di-heme-cytochrome-c peroxidase. Targeted inactivation of this gene resulted in a slight decrease of crocacin production. However, this might be regarded as a secondary effect. The available sequence closes with the overlapping genes *orf6* and *orf7*, encoding a putative two-component system consisting of a signal transduction histidine kinase and a LytR family transcriptional regulator. Such systems usually regulate transcription of target genes upon certain stimuli (West and Stock, 2001) and may regulate crocacin formation too. To investigate the role of this system, we integrated plasmid pSUP_orf6_KO into the genome of *C. crocatus*

Cm c5, thereby impeding expression of the histidine kinase. Again, no significant effect on crocacin production was observed. Nevertheless, these first inactivation experiments allowed us to determine the boundaries of the *cro* cluster. The results of the in-depth sequence analysis served as a basis for all further experiments, which led to the presentation of a detailed model for crocacin biosynthesis in *C. crocactus* Cm c5, as shown in Figure 1.

Proposed Model of Crocacin Biosynthesis

On the basis of the structure of crocacin, biosynthesis should most likely start with the activation and tethering of *trans*-cinnamic acid to the ACP domain of the loading module (Figure 1B). To confirm this, we supplemented growing cultures of *C. crocatus* Cm c5 with isotope-labeled d7-trans-cinnamic acid and identified the incorporation of this starter unit by HPLC-MS analysis of the methanolic extract. A mass increase of 7 Da indicates incorporation of the full molecule (Figure S1). The use of *trans*-cinnamoyl-CoA as starter unit in natural product biosynthesis has already been reported for the stilbenes produced by *Photorhabdus luminescens* TTO1 (Joyce et al., 2008). Detailed studies have shown that in *P. luminescens* TTO1 *trans*-cinnamic acid is derived through nonoxidative deamination of L-phenylalanine by the action of StlA (NP_929491, phenylalanine ammonia lyase [PAL], Enzyme Commission [EC] No. 4.3.1.5.) (Williams et al., 2005). In order to shed light on the origin of the starter molecule in *C. crocatus* Cm c5, we fed ring-labeled $^{13}C_6$-L-phenylalanine and in fact observed a mass increase of 6 Da for the crocacins (Figure S1). This suggests a similar origin of *trans*-cinnamic acid in *C. crocatus* Cm c5. To further study this, we analyzed the genome of *C. crocatus* Cm c5 regarding the presence of a PAL-encoding gene by using the amino acid sequence of StlA in a tBLASTn search and found a candidate gene (*pal*^{cmc5}). In this context, it must be mentioned that PALs share very high sequence similarity to histidine ammonia lyases (HALs; EC No. 4.3.1.3.), but they can be distinguished on the basis of some key amino acid residues (Xiang and Moore, 2002; Calabrese et al., 2004). Accordingly, Pal^{cmc5} could be classified as a HAL, and consequently, targeted inactivation of this gene had no effect on crocacin production (Figure 2). Hence, the origin of the *trans*-cinnamic acid starter unit remains elusive, although its incorporation could be verified.

Chain elongation is mediated by PKS modules 1 and 2 using methylmalonyl-CoA as extender unit. Their modular architecture of type KS-AT-MT-KR-ACP perfectly corresponds to the β-methoxy moieties present in crocacin. The origin of the methoxy groups was verified by feeding experiments using isotope-labeled d3-methionine. Mass shifts of +3, +6, and +9 Da are in accordance with the activity of *S*-adenosylmethionine (SAM)-dependent MTs (CroA-MT, CroB-MT and CroL; see below) responsible for three distinct methylations (Figure S1). Assuming that crocacin biosynthesis follows the "colinearity rule," one would expect two distinct PKS modules after CroB introducing two malonyl-CoA-extender units. However, and strikingly, only one PKS module (CroC) with the domain architecture KS-AT-DH-KR-ACP is located behind CroB. To explain this unusual finding, we hypothesize that CroC acts in a programmed iterative manner similarly to PKS modules involved in stigmatellin (Gaitatzis et al., 2002) and aureothin (He and

Table 1. Putative Functions and Domain Organization within the Crocacin Gene Cluster and the Contiguous Region

Gene/Protein	Length (bp/aa)	Modules	Domains (Coordinates in Protein Sequence [aa])
croA/CroA	7,713/2,570	loading	CoA-Lig$_L$ (32–477), ACP$_L$ (606–672)
		module 1	KS (691–1,064), AT (1,222–1,504), *O*-MT (1,620–1,876), KR (2,160–2,354), ACP (2,456–2,520)
croB/CroB	5,700/1,899	module 2	KS (36–412), AT (568–885), *O*-MT (969–1,224) KR (1,509–1,688), ACP (1,795–1,857)
croC/CroC	5,637/1,878	module 3	KS (59–401), AT (555–871), DH (925–1,096), KR (1,251–1,358), ACP (1,771–1,828)
croD/CroD	2,793/930	β-branching	KS (39–408), AT (566–886)
croE/CroE	1,254/417	β-branching	hydroxymethylglutaryl-CoA synthase homolog
croF/CroF	798/265	β-branching	enoyl-CoA hydratase/isomerase homolog
croG/CroG	789/262	β-branching	enoyl-CoA hydratase/isomerase homolog
croH/CroH	201/66	β-branching	ACP (6–57)
croI/CroI	8,892/2,963	module 4	C (28–321), A (521–918), PCP (1,007–1,070)
		module 5	KS (1,108–1,483), AT (1,638–1,955), DH (2,004–2,182), KR (2,557–2,735), ACP (2,984–2,913)
croJ/CroJ	5,685/1,894	module 6	KS (39–414), AT (568–885), DH (935–1,102), KR (1,475–1,664), ACP (1,768–1,831)
croK/CroK	4,611/1,536	module 7	C1 (33–328), A (526–923), PCP (1,012–1,071), C2 (1,109–1,404)
croL/CroL	723/240	MT	MT (41–144)

Orf	Length (bp/aa)	Accession No. of Closest Homolog	Similarity/Identity [%]	Putative Function/Homolog
1	1,827/608	ZP_06890144	54/38	ABC-1 domain protein
2	1,698/565	ZP_06890144	51/28	ABC-1 domain protein
3	1,248/415	ZP_21238480	69/50	cytochrome P450
4	2,037/679	ZP_21238481	66/51	long-chain-fatty-acid-CoA ligase
5	1,722/573	ZP_11029676	85/76	di-heme cytochrome-c peroxidase
6	1,095/364	ZP_11029675	89/81	two-component histidine kinase
7	783/260	ZP_11029674	93/85	regulatory protein

Hertweck, 2003) biosynthesis. The way this iteration is likely to occur is extraordinary, as the first iteration cycle would use the full set of catalytic domains of module 3 to form an α,β-unsaturated product, while the second iteration would be nonreductive. Given the structure of crocacin, it seems likely that the nascent polyketide chain directly undergoes β-branching at C13 after the second iteration step, notably without any β carbon processing. A plausible explanation for this is that CroE catalyzes the aldol addition of the acetyl group from CroH to the β-keto functionality of C13 much faster than the motif would undergo reduction through CroC-KR and CroC-DH. This hypothesis is supported by the fact that we did not encounter any crocacin derivatives, which would result from a less strict process, such as iteration without β-branching or noniterative use without β-branching. Moreover, the methyl group at C13 is not SAM derived, as shown by the aforementioned feeding experiment. In particular, the methyl branch is likely to be incorporated as follows (Figure 1C): CroD-AT loads malonyl-CoA onto the free-standing ACP (CroH). Subsequently, CroD-KS decarboxylates the malonyl-thioester and CroE catalyzes the C-C bond formation between the acetyl moiety and the β-carbon C13 of the enzyme-bound template. Finally, the enoyl-CoA hydratase homologs CroF and CroG process the intermediate by dehydration and decarboxylation to yield the β-methyl moiety. To support this model, we disrupted *croE* and *croF*. The respective mutant strains were analyzed by HPLC-MS and showed loss of crocacin production, which suggests the protein's participation in crocacin

biosynthesis (Figure 2). Such β-methyl moiety introducing biochemistry on growing PKS templates has been reported for the biosynthesis of other natural products, such as mupirocin (El-Sayed et al., 2003), curacin (Chang et al., 2004), myxovirescin (Simunovic et al., 2006), bryostatin (Sudek et al., 2007), and most recently myxopyronin (Sucipto et al., 2013). Besides that, in vitro investigations of the β-branching machineries from *pksX* and myxovirescin pathways proved the mechanism proposed earlier (Calderone et al., 2006; Calderone et al., 2007).

After β-branching, the biosynthetic intermediate is passed through module 4 (C-A-PCP) and thereby extended with glycine. Afterward, two PKS elongation steps with malonyl-CoA are performed by module 5 and module 6, both resulting in an unsaturated polyketide chain. In the case of module 5, an unusual position of the double bond is observed, which could be attributed to the conjugational system, which is extended for such a constitution. Although the respective DH domain does not obviously differ from common DH domains, it may either catalyze an isomerization (Leesong et al., 1996) or a γ-β elimination, as suggested for rhizoxin (Kusebauch et al., 2010).

The assembly of crocacin ends with module 7, comprising the domain order C-A-PCP-C. The first C domain CroK-C1 catalyzes the condensation of a second glycine, and the resulting crocacin intermediate remains bound to PCP7. Given the structure of crocacin B, one would expect the presence and subsequent action of a TE domain, leading to the formation of the free carboxylic acid upon hydrolytic release. Surprisingly, the very last domain CroK-C2 of this megasynthetase seems to be another C domain,

as could be classified by a homology analysis. In silico analysis of the *cro* locus showed no indications of an additional TE-encoding gene, pointing toward an unusual release function of CroK-C2. Hereupon, we introduced plasmid pSUP_croK-C2_KO into the genome of *C. crocatus* Cm c5, resulting in an inactivation of *croK-C2* and consequently in mutants that might be able to produce PCP7-bound crocacin. Assuming that a distinct TE domain encoded somewhere else in the genome would be capable of releasing PCP-bound crocacin, this particular gene disruption would result in an unaltered production of crocacins. Interestingly, all clones analyzed by HPLC-MS did not produce crocacins or shunt products thereof (Figure 2). This indicates the involvement of CroK-C2 in the release of the final product, although we cannot exclude that the mutagenesis itself results in inactive CroK. To gain further insights into the actual release mechanism, we attempted the heterologous production of CroK-C2 alongside with in vitro experiments, as described below.

Crocacin A, the main derivative in *C. crocatus* Cm c5, is most likely derived from crocacin B by a tailoring O-MT acting on the terminal carboxyl group. As described above, the *cro* locus encodes the putative O-MT CroL, which could be responsible for the formation of crocacin A. The targeted inactivation of this gene resulted in the abolishment of crocacin A production, while elevated levels of crocacin B were detectable in extracts when compared with the wild-type (Figure 2). Together with our d3-methionine feeding experiments (Figure S1), these findings confirmed the assumption that CroL is the SAM-dependent tailoring enzyme responsible for O-methylation of crocacin B. This is yet another example of a post-assembly-line modification of a natural compound whereby the product range of the core biosynthetic machinery is broadened.

Heterologous Production and In Vitro Assay of CroK-C2

Gene inactivation experiments revealed that mutants being altered in the last segment of *croK* failed to produce crocacin (see above). This finding, together with the lack of an encoded TE domain in the cluster, suggested an extraordinary mode of action of the terminal C domain CroK-C2 in the release mechanism of crocacin B from the assembly line. Because of these interesting facts, we set out to investigate the protein's function in crocacin biosynthesis. First, we applied an in silico approach using the NaPDoS tool (Ziemert et al., 2012) to classify this protein (Figure S3A). According to this evaluation, CroK-C2 clusters in a small clade named "modified AA" together with three C domains from the bleomycin and microcystin pathways, whereas the remaining domains CroI-C and CroK-C1 group with the classic LCL domains catalyzing condensation of L-amino acids without epimerization. Even though the group modified AA has not been experimentally characterized, its proposed function is to modify incorporated amino acids, such as in the dehydration of serine to dehydroalanine (Du et al., 2000; Tillett et al., 2000). A similar function for CroK-C2 is not expected; nevertheless, this phylogenetic analysis emphasizes its distinctiveness among common C domains, which suggests a catalytic activity that differs from the standard ones. In order to determine how and to what extent CroK-C2 varies from classic C domains, a primary sequence alignment of all C domains of crocacin biosynthesis was performed (Figure S3B) and analyzed for the presence of conserved motifs (Marahiel et al., 1997). Although CroK-C2 shows 36.8% pairwise similarity to CroI-C and 37.5% to CroK-C1, respectively, the classic C domains CroI-C and CroK-C1 are 65.8% similar between each other. Furthermore, CroK-C2 perfectly features the HHXXXDG consensus motif and only moderately deviates from the consensus motifs C2 and C4. Contrary to this, motifs C1 and C5 seem to be not present at all in CroK-C2.

To investigate the actual function of CroK-C2 in crocacin biosynthesis, we heterologously produced the protein in *Escherichia coli* BL21 (DE3) pLysS and purified it to homogeneity as N-terminal His$_6$-tagged form (Figure 3). To test if CroK-C2 is able to release PCP7-bound crocacin, an *N*-acetylcysteaminyl-crocacin B (crocacin-SNAC) mimicking PCP7-bound crocacin was synthesized and used as a substrate-mimic to set up an enzymatic in vitro assay using native and denatured CroK-C2. Indeed, hydrolysis of the crocacin-SNAC ester was detectable only upon incubation with native CroK-C2 (Figure S4A). This initial test confirmed the proposed role of this unusual terminal C domain in the release of crocacin B. As a next step, we aimed to evaluate the kinetic parameters of the CroK-C2-mediated hydrolysis. As there is no cofactor consumption that could be photometrically monitored, we set up an MS-based method to record the reaction. The high salt concentration of the assay obviously hampered a mass spectrometric approach based on direct infusion. Hence, a fast HPLC-MS method was established, which is capable of separating crocacin-SNAC and its hydrolysis product in less than 1 min. In particular, consecutive injections of the assay mixture enabled us to track the time course in steps of 0.9 min (Figure 3C). We were not able to reach a plateau phase in the Michaelis-Menten graph, because of the poor water solubility of the substrate, upon which we could not use substrate concentrations higher than 200 μM. We found that on one hand, DMSO has a strong influence on the enzyme activity (Figure S4C), and on the other hand, a certain concentration of DMSO is mandatory to dissolve the substrate. For these two reasons, we maintained DMSO concentration at 5% for all assays as the most reasonable assay condition. After all, the formation of crocacin B with variable concentrations of crocacin-SNAC exhibited Michaelis-Menten kinetics with parameters as follows: $K_m = 140.3 \pm 26.1$ μM, $k_{cat} = 749$ s^{-1}, and $k_{cat}/K_m = 5.34$ s^{-1} μM^{-1} (Figure 3D). On the basis of these parameters, we conclude that CroK-C2 efficiently releases crocacin B from the assembly line.

Having confirmed the exceptional activity of CroK-C2, we intended to assess whether the hydrolytic cleavage of crocacin-SNAC is unique to this particular enzyme or if any C domain could catalyze this reaction. To evaluate this question, we chose the myxochelin megasynthetase MxcG (C-A-PCP-Red) to be assayed with crocacin-SNAC for two reasons. First, its functionality has already been shown in vitro (Gaitatzis et al., 2001; Li et al., 2008), and second, this NRPS ends with an Red domain instead of a TE, so that an undesired hydrolysis of the substrate can most likely be neglected, especially as even NADPH can be omitted in the assay. Following this idea, MxcG was heterologously produced and purified to homogeneity (Figure 3A). MxcG was then used in an in vitro assay under the same conditions as applied in previous experiments. However, we did not monitor the reaction online but rather stopped the reaction at
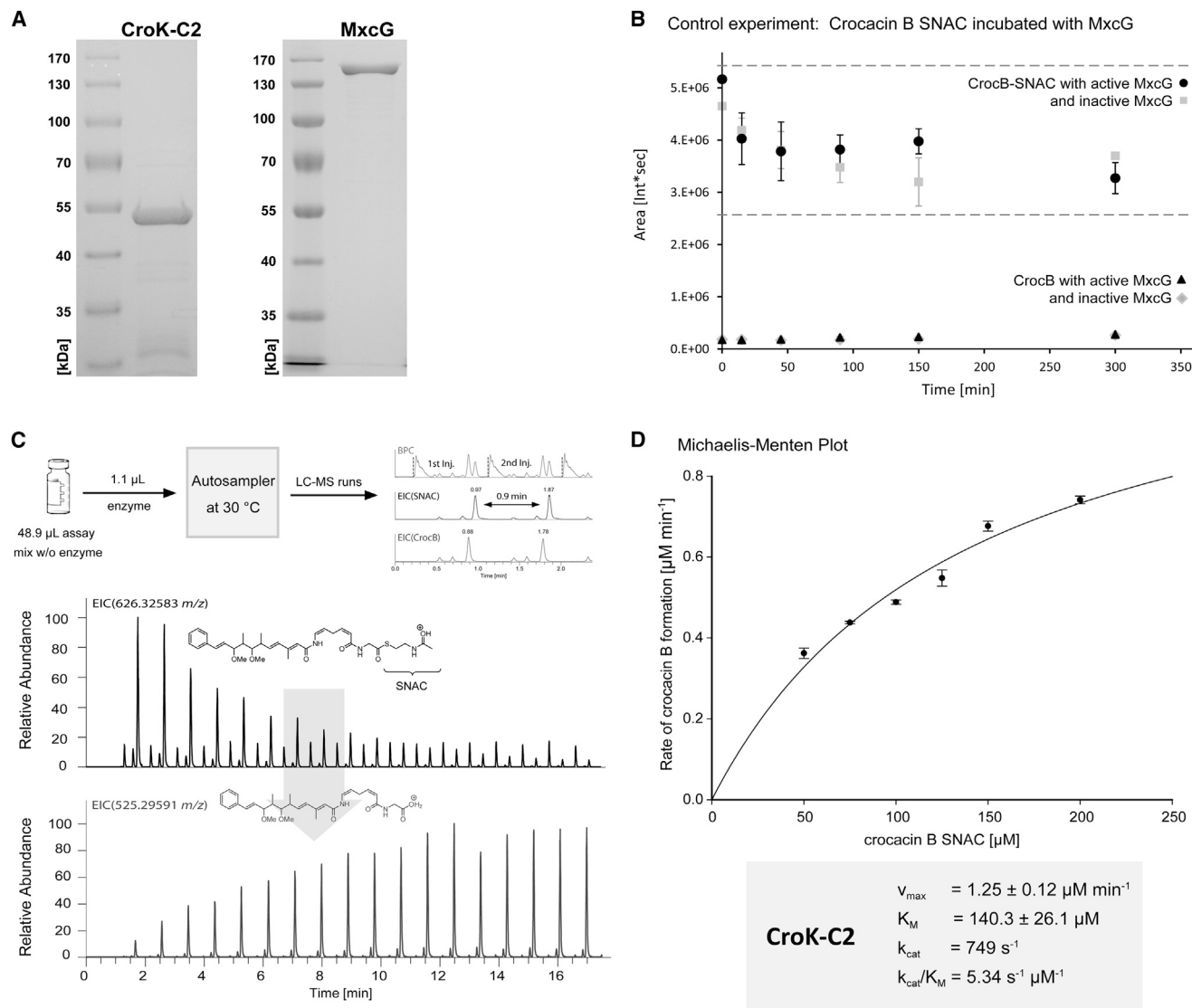
Figure 3. In Vitro Activity Assays of CroK-C2 and MxcG with Respect to Crocacin B SNAC Hydrolysis

(A) SDS-PAGE analysis of heterologously produced and enriched CroK-C2 (52.4 kDa) and MxcG (156.8 kDa).

(B) Crocacin-SNAC hydrolysis experiment with active and inactive MxcG shows no effect of MxcG on this substrate. Error bars show calculated SDs.

(C) Fast HPLC-MS assay used to monitor the formation of crocacin B as a measure for CroK-C2 activity. The reaction mixture was consecutively probed every 0.9 min. Different substrate concentrations were measured with this technique to determine initial reaction rates.

(D) Michaelis-Menten plot and respective parameters of CroK-C2 as calculated on the basis of crocacin B formation in the assay. Error bars show calculated SDs.

different time points. We observed neither an increase of the product (crocacin B) nor a significant decrease of the substrate (crocacin-SNAC) and thereby concluded that MxcG, contrary to CroK-C2, is not able to act hydrolytically on crocacin-SNAC (Figure 3B).

Taken together, these results depict the specialty of CroK-C2 with respect to other C domains and prove its remarkable role in the release of a PCP-bound intermediate by hydrolysis. Obviously, it cannot be ruled out that other hydrolytic C domains present in currently uncharacterized assembly lines exist. For that reason, we aimed to screen the GenBank nucleotide database for the presence of NRPS and NRPS/PKS hybrid clusters of bacterial origin that do not encode an apparent release function

like a TE or terminal Red domain (TD) (Figure S5). Initially, starting from 2,390 sequences, 1,632 were eliminated because no biosynthetic function has been assigned to them. Second, the remaining 758 sequences were analyzed for NRPS signatures, with 241 sequences matching these criteria. Of these sequences, 175 contained TEs and 16 contained TDs. Accordingly, the remaining 50 sequences featured NRPS-related genes, while no TE or TD was detectable. A detailed in silico analysis (Table S1) of these sequences using Pfam (Punta et al., 2012) revealed that 15 additional sequences contained TE domains that either were not found by the antiSMASH software (because of their poor homology to other TEs) or were part of a split cluster in which the other part contained the TE.

Hence, 35 sequences did not harbor any classic termination domains. Six sequences thereof have been defined as a partial sequence in the database, which is why they were excluded from the analysis. From the remaining 29 sequences, 16 encode at least partially characterized secondary metabolite pathways in which chain termination should involve a C domain catalyzing amide or ester bond formation. Finally, the 13 residual ones either have not been linked to secondary metabolites or do not appear to encode terminal C domains, except one polyketide-2 cluster (AB070942) from *Streptomyces avermitilis* (Omura et al., 2001) and the crocacin gene cluster (FN547928) from *C. crocatus* Cm c5, of which an early version was previously submitted to GenBank. The core cluster from *S. avermitilis* most likely consists of two distinct loci, BAB69218.1 and BAB69217.1. The first encodes the loading module, and the second bears the information of the loading modules ACP followed by two PKS modules (KS-AT-DH-KR-ACP) and a terminal C domain as the sole NRPS feature. Of note, a Pfam search revealed the presence of a putative haloacid dehalogenase-like hydrolase (E value = 0.32) after the C domain on BAB69217.1 and a putative abhydrolase (E value = 9.7×10^{-13}) encoded by the following locus BAB69216.1. This suggests that compound release might rather be performed by one of the putative hydrolases than by the C domain, but the actual scenario remains elusive.

According to this in silico evaluation and to the best of our knowledge, CroK-C2 and its hydrolytic chain-terminating activity is currently unique to bacterial NRPS machineries.

SIGNIFICANCE

Myxobacteria are well-known producers of structurally diverse bioactive natural products. Their assembly is performed by large multienzyme PKS and NRPS complexes and often comprises unprecedented biosynthetic features. Therefore, the analysis and characterization of yet undescribed secondary metabolite gene clusters hold great biotechnological potential, especially when new catalytic functions within these machineries are detected. Knowledge gained could help in engineering related biosynthetic routes toward the formation of altered products. Our analysis of the crocacin biosynthetic gene cluster in *C. crocatus* Cm c5 indeed revealed such a catalytic activity of a condensation-type domain being part of a modular NRPS. Along with a programmed iteration module and an uncommon β-branching cassette, the cluster features a highly remarkable domain order in the last module. With a total absence of typical release domains, the observed hydrolytic release of crocacins was initially not understood. This question was addressed by characterizing the terminal C domain CroK-C2 by reconstitution of its activity in vitro, which proved that CroK-C2 is responsible for release of the PCP-bound crocacin. This particular finding can now help in engineering NRPS natural product pathways toward a programmed template release, thus opening the possibility of creating new derivatives with terminal carboxylic acid moieties, but further studies are required to elucidate the actual molecular basis that triggers a former C domain to catalyze hydrolysis.

EXPERIMENTAL PROCEDURES

Strains and Culture Conditions
See Supplemental Experimental Procedures.

Gene Disruption Experiments in *C. crocatus* Cm c5
See Supplemental Experimental Procedures.

Construction of pSUP_Hyg Derivatives Used in This Study
See Supplemental Experimental Procedures.

Isolation and Purification of Crocacin B
Utilized instruments, the strain maintenance and 70 l scale cultivation of *C. crocatus* Cm c5, as well as the obtainment of the crude extract were described previously (Viehrig et al., 2013). As described for the isolation of crocapeptins A_1 to A_3, the crude extract was subjected to solvent partition using 85% aqueous methanol, which was extracted twice with heptane. Subsequently, the aqueous methanol was adjusted to 70% methanol and extracted twice with dichloromethane. The dichloromethane fraction contained the main amount of crocacin B, which was fractionated by reverse-phase (RP) medium-pressure liquid chromatography (column 480 × 30 mm, ODS/AQ C18 [Kronlab], gradient 30%−100% methanol in 60 min, flow 30 ml/min, UV peak detection at 210 nm). Fractions containing crocacin B were combined and further purified by preparative RP HPLC (column 250 × 21 mm, Nucleodur C18 5 μm [Macherey-Nagel], gradient 45%–100% methanol with 0.2% acetic acid in 25 min) provided 10.7 mg of crocacin B.

Synthesis of Crocacin-SNAC
The synthesis strategy used in the generation of crocacin-SNAC is based on the method described Carroll et al. (2002). A solution of crocacin B (6.3 mg, 12 μmol) in 1 ml dichloromethane was cooled down to 0°C. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide (4.6 mg, 24 μmol) and 4-dimethylaminopyridine (∼50 μg, cat.) were added. After 2 hr stirring at 25°C, the solution was cooled down to 0°C, and *N*-acetylcysteamine (7.2 mg, 60 μmol, 5 eq.) was added. Subsequently, the mixture was stirred for 16 hr at 25°C and quenched with 2 ml water. After separation of the layers, dichloromethane was removed in vacuo, and the resulting material was analyzed by HPLC-UV-MS. Subsequently, crocacin-SNAC (1.8 mg, 2.8 μmol, 24%) was isolated from the crude product by preparative HPLC (column 250 × 21 mm, Nucleodur C18 5 μm [Macherey-Nagel], gradient 65%–100% methanol in 25 min).

Heterologous Production and Purification of CroK-C2 and MxcG
CroK-C2 was produced as N-terminal His-tagged fusion protein using the pET28b(+) plasmid (Novagen). For construction of the expression plasmid pET28b_croK-C2, the respective part of *croK* was amplified by PCR using the oligonucleotides CroK-C2-NdeI (CCG CGC CCA TAT GCC TCT CGA CAA GGT GGC CCT G) and CroK-C2-XhoI (ACG CGC TCG AGT CAG GCT GCT GGA GAG GTC G) and integrated into the prepared vector. Production of the His₆ fusion protein was achieved by using *E. coli* BL21 DE3 carrying the pLysS plasmid. Initial cultivation was carried out in a 2 l shaking flask containing 300 ml Luria broth medium supplemented with 25 μg/ml chloramphenicol and 50 μg/ml kanamycin sulfate at 37°C and 200 rpm. When the culture reached an optical density at 600 nm of 0.5 to 0.7, gene expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.25 mM, and the temperature was shifted to 16°C. After further cultivation for approximately 20 hr, the cells were harvested by centrifugation, washed twice with PBS buffer, and resuspended in ice-cold 25 ml HisTrap A buffer (25 mM HEPES, 500 mM NaCl, 25 mM imidazole, and 10% v/v glycerol [pH 8.0]) containing Roche complete protease inhibitors. Cells were disrupted by using a French press three times at 900 psi, and the debris was removed by centrifugation at 20,000 × *g* for 30 min at 4°C. Afterward, the soluble fraction was filtered through a 0.45 μm syringe filter before incubation with Ni²⁺-NTA (Qiagen) for 1 hr at 4°C. The protein was eluted with a stepwise gradient of imidazole using different mixtures of HisTrap A and HisTrap B (25 mM HEPES, 500 mM NaCl, 500 mM imidazole, and 10% v/v glycerol [pH 8.0]). All fractions were subjected to SDS-PAGE, and those that contained the target protein (52.4 kDa) were pooled and concentrated. As a second purification step, size-exclusion chromatography was applied. Therefore, an

Äkta pure M2 system equipped with a Superdex 200 increase 10/300 GL column was used in HisTrap buffer without imidazole at a flow rate of 0.75 ml/min. Fractions of interest were pooled and analyzed by SDS-PAGE and finally subjected to MALDI-TOF analysis to prove their identity.

MxcG was produced and purified according to a previously described protocol (Gaitatzis et al., 2001; Li et al., 2008). The final fraction was concentrated and subjected to size-exclusion chromatography to remove residual dithiothreitol present in the sample. Therefore, an Äkta pure M2 system equipped with a Superdex 200 increase 10/300 GL column at a flow rate of 0.75 ml/min in elution buffer (25 mM Tris HCl [(pH 8.0], 50 mM NaCl, 0.1 mM EDTA) was used. The amount of protein in the respective samples was determined using Bio-Rad Protein Assay dye reagent, and obtained intensities were referred to a calibration line.

In Vitro Assay for Monitoring the Hydrolysis of *N*-Acetylcysteamine Ester of Crocacin B by CroK-C2 and MxcG

Test samples were premixed without protein and preheated to 30°C. The enzyme reaction was started by the addition of active protein or, in the case of the negative control experiments, by the addition of heat-inactivated protein (5 min at 98°C). Activity testing of CroK-C2 was carried out by a fast liquid chromatography mass spectrometry (LC-MS) method in order to measure the time course of the reaction. Reaction mixtures of 50 μl were prepared in conical HPLC vials containing 100 mM Tris HCl (pH 7.6), 50 mM NaCl, 0.1 μM protein, and crocacin-SNAC in concentrations ranging from 50 to 200 μM. The DMSO content was fixed to 5% (v/v) for all assay mixtures. To measure a time course, the autosampler temperature was set to 30°C to allow reasonable enzyme activity. Upon starting the reaction by adding enzyme, the same sample was probed consecutively by drawing 1 μl of the mixture every 0.9 min. This sample was injected to the HPLC by using an overlapping injection protocol. Separation is achieved by an isocratic elution at 58% B on a Waters BEH C18, 50 × 2.1 mm, 1.7 μm dp column using (A) H_2O + 0.1% formic acid (FA) to (B) acetonitrile + 0.1% FA at a flow rate of 800 μl/min and 45°C. Peaks for crocacin B ($[M+H]^+$ = 525.2959 *m/z*) and crocacin-SNAC ($[M+H]^+$ = 626.3258 *m/z*) are monitored in SIM mode (450-700 *m/z*) at a resolution of 7,500 and 200 ms fill time using a Thermo Scientific Orbitrap. The concentration of crocacin B was determined by an area-based calibration curve, which was measured using the same method. Rates of product formation were calculated from the initial linear portion of the obtained time course and fitted to the Michaelis-Menten equation by using the enzyme kinetics tool of Sigma Plot 11.0 (Systat Software).

Activity of MxcG was monitored by quenching enzyme assays at different time points. Reaction mixtures of 70 μl were incubated in HPLC vials at 30°C. Each reaction mixture contained 25 mM Tris HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 5 μM protein, 20 μM crocacin-SNAC, and 2% DMSO. All reactions were started simultaneously by the addition of enzyme and quenched at different time points by the addition of 70 μl methanol. After centrifugation, 2 μl of the reaction mixture was used for LC-MS analysis using the standard method.

Crocacin-SNAC
See Supplemental Experimental Procedures.

LC-MS Standard Method
See Supplemental Experimental Procedures.

In Silico Analysis of NRPS-Encoding Gene Clusters with No Apparent Termination Domain
See Supplemental Experimental Procedures.

ACCESSION NUMBERS

The sequence of crocacin biosynthetic gene cluster from *C. crocatus* Cm c5 has been deposited in GenBank under accession number KJ868728.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2014.05.012.

REFERENCES

Ansari, M.Z., Sharma, J., Gokhale, R.S., and Mohanty, D. (2008). In silico analysis of methyltransferase domains involved in biosynthesis of secondary metabolites. BMC Bioinformatics *9*, 454.

Brautaset, T., Sekurova, O.N., Sletta, H., Ellingsen, T.E., StrŁm, A.R., Valla, S., and Zotchev, S.B. (2000). Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. Chem. Biol. *7*, 395–403.

Buntin, K., Rachid, S., Scharfe, M., Blöcker, H., Weissman, K.J., and Müller, R. (2008). Production of the antifungal isochromanone ajudazols A and B in *Chondromyces crocatus* Cm c5: biosynthetic machinery and cytochrome P450 modifications. Angew. Chem. Int. Ed. Engl. *47*, 4595–4599.

Buntin, K., Irschik, H., Weissman, K.J., Luxenburger, E., Blöcker, H., and Müller, R. (2010). Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity. Chem. Biol. *17*, 342–356.

Butcher, R.A., Schroeder, F.C., Fischbach, M.A., Straight, P.D., Kolter, R., Walsh, C.T., and Clardy, J. (2007). The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *104*, 1506–1509.

Byford, M.F., Baldwin, J.E., Shiau, C.Y., and Schofield, C.J. (1997). The mechanism of ACV synthetase. Chem. Rev. *97*, 2631–2650.

Calabrese, J.C., Jordan, D.B., Boodhoo, A., Sariaslani, S., and Vannelli, T. (2004). Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. Biochemistry *43*, 11403–11416.

Calderone, C.T. (2008). Isoprenoid-like alkylations in polyketide biosynthesis. Nat. Prod. Rep. *25*, 845–853.

Calderone, C.T., Kowtoniuk, W.E., Kelleher, N.L., Walsh, C.T., and Dorrestein, P.C. (2006). Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the *pksX* pathway of *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *103*, 8977–8982.

Calderone, C.T., Iwig, D.F., Dorrestein, P.C., Kelleher, N.L., and Walsh, C.T. (2007). Incorporation of nonmethyl branches by isoprenoid-like logic: multiple beta-alkylation events in the biosynthesis of myxovirescin A1. Chem. Biol. *14*, 835–846.

Carroll, B.J., Moss, S.J., Bai, L., Kato, Y., Toelzer, S., Yu, T.W., and Floss, H.G. (2002). Identification of a set of genes involved in the formation of the substrate for the incorporation of the unusual "glycolate" chain extension unit in ansamitocin biosynthesis. J. Am. Chem. Soc. *124*, 4176–4177.

Chang, Z., Sitachitta, N., Rossi, J.V., Roberts, M.A., Flatt, P.M., Jia, J., Sherman, D.H., and Gerwick, W.H. (2004). Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. J. Nat. Prod. *67*, 1356–1367.

Chin, Y.W., Balunas, M.J., Chai, H.B., and Kinghorn, A.D. (2006). Drug discovery from natural sources. AAPS J. *8*, E239–E253.

Du, L., and Lou, L. (2010). PKS and NRPS release mechanisms. Nat. Prod. Rep. *27*, 255–278.

Du, L., Sánchez, C., Chen, M., Edwards, D.J., and Shen, B. (2000). The biosynthetic gene cluster for the antitumor drug bleomycin from Streptomyces verticillus ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. Chem. Biol. *7*, 623–642.

Ehmann, D.E., Gehring, A.M., and Walsh, C.T. (1999). Lysine biosynthesis in Saccharomyces cerevisiae: mechanism of alpha-aminoadipate reductase

(Lys2) involves posttranslational phosphopantetheinylation by Lys5. Biochemistry *38*, 6171–6177.

El-Sayed, A.K., Hothersall, J., Cooper, S.M., Stephens, E., Simpson, T.J., and Thomas, C.M. (2003). Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. Chem. Biol. *10*, 419–430.

Gaitatzis, N., Kunze, B., and Müller, R. (2001). In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. USA *98*, 11136–11141.

Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. (2002). The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. J. Biol. Chem. *277*, 13082–13090.

Gao, X., Haynes, S.W., Ames, B.D., Wang, P., Vien, L.P., Walsh, C.T., and Tang, Y. (2012). Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. Nat. Chem. Biol. *8*, 823–830.

He, J., and Hertweck, C. (2003). Iteration as programmed event during polyketide assembly; molecular analysis of the aureothin biosynthesis gene cluster. Chem. Biol. *10*, 1225–1232.

Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. Angew. Chem. Int. Ed. Engl. *48*, 4688–4716.

Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. Chem. Rev. *97*, 2465–2498.

Ishida, K., Christiansen, G., Yoshida, W.Y., Kurmayer, R., Welker, M., Valls, N., Bonjoch, J., Hertweck, C., Börner, T., Hemscheidt, T., and Dittmann, E. (2007). Biosynthesis and structure of aeruginoside 126A and 126B, cyanobacterial peptide glycosides bearing a 2-carboxy-6-hydroxyoctahydroindole moiety. Chem. Biol. *14*, 565–576.

Jansen, R., Washausen, P., Kunze, B., Reichenbach, H., and Höfle, G. (1999). Antibiotics from gliding bacteria, LXXXIII - The crocacins, novel antifungal and cytotoxic antibiotics from Chondromyces crocatus and Chondromyces pediculatus (Myxobacteria): Isolation and structure elucidation. Eur. J. Org. Chem. *5*, 1085–1089.

Joyce, S.A., Brachmann, A.O., Glazer, I., Lango, L., Schwär, G., Clarke, D.J., and Bode, H.B. (2008). Bacterial biosynthesis of a multipotent stilbene. Angew. Chem. Int. Ed. Engl. *47*, 1942–1945.

Keating, T.A., Ehmann, D.E., Kohli, R.M., Marshall, C.G., Trauger, J.W., and Walsh, C.T. (2001). Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. ChemBioChem *2*, 99–107.

Keatinge-Clay, A. (2008). Crystal structure of the erythromycin polyketide synthase dehydratase. J. Mol. Biol. *384*, 941–953.

Kusebauch, B., Busch, B., Scherlach, K., Roth, M., and Hertweck, C. (2010). Functionally distinct modules operate two consecutive alpha,beta—>beta,gamma double-bond shifts in the rhizoxin polyketide assembly line. Angew. Chem. Int. Ed. Engl. *49*, 1460–1464.

Leesong, M., Henderson, B.S., Gillig, J.R., Schwab, J.M., and Smith, J.L. (1996). Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. Structure *4*, 253–264.

Leonard, C.J., Aravind, L., and Koonin, E.V. (1998). Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. Genome Res. *8*, 1038–1047.

Li, Y., Weissman, K.J., and Müller, R. (2008). Myxochelin biosynthesis: direct evidence for two- and four-electron reduction of a carrier protein-bound thioester. J. Am. Chem. Soc. *130*, 7554–7555.

Lin, S., Van Lanen, S.G., and Shen, B. (2009). A free-standing condensation enzyme catalyzing ester bond formation in C-1027 biosynthesis. Proc. Natl. Acad. Sci. USA *106*, 4183–4188.

Lucas, X., Senger, C., Erxleben, A., Grüning, B.A., Döring, K., Mosch, J., Flemming, S., and Günther, S. (2013). StreptomeDB: a resource for natural compounds isolated from Streptomyces species. Nucleic Acids Res. *41*, D1130–D1136.

Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. *97*, 2651–2674.

Maruyama, C., Toyoda, J., Kato, Y., Izumikawa, M., Takagi, M., Shin-ya, K., Katano, H., Utagawa, T., and Hamano, Y. (2012). A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. Nat. Chem. Biol. *8*, 791–797.

Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. *39*, W339–W346.

Moldenhauer, J., Chen, X.H., Borriss, R., and Piel, J. (2007). Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. Angew. Chem. Int. Ed. Engl. *46*, 8195–8197.

Motamedi, H., and Shafiee, A. (1998). The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. Eur. J. Biochem. *256*, 528–534.

Müller, R., and Wink, J. (2014). Future potential for anti-infectives from bacteria - how to exploit biodiversity and genomic potential. Int. J. Med. Microbiol. *304*, 3–13.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., et al. (2001). Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites. Proc. Natl. Acad. Sci. USA *98*, 12215–12220.

Pieper, R., Haese, A., Schröder, W., and Zocher, R. (1995). Arrangement of catalytic sites in the multifunctional enzyme enniatin synthetase. Eur. J. Biochem. *230*, 119–126.

Pistorius, D., and Müller, R. (2012). Discovery of the rhizopodin biosynthetic gene cluster in *Stigmatella aurantiaca* Sg a15 by genome mining. Chembiochem *13*, 416–426.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., et al. (2012). The Pfam protein families database. Nucleic Acids Res. *40*, D290–D301.

Rachid, S., Krug, D., Kunze, B., Kochems, I., Scharfe, M., Zabriskie, T.M., Blöcker, H., and Müller, R. (2006). Molecular and biochemical studies of chondramide formation-highly cytotoxic natural products from *Chondromyces crocatus* Cm c5. Chem. Biol. *13*, 667–681.

Rachid, S., Scharfe, M., Blöcker, H., Weissman, K.J., and Müller, R. (2009). Unusual chemistry in the biosynthesis of the antibiotic chondrochlorens. Chem. Biol. *16*, 70–81.

Robbel, L., Hoyer, K.M., and Marahiel, M.A. (2009). TioS T-TE—a prototypical thioesterase responsible for cyclodimerization of the quinoline- and quinoxaline-type class of chromodepsipeptides. FEBS J. *276*, 1641–1653.

Röttig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011). NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res. *39*, W362–W367.

Schwecke, T., Aparicio, J.F., Molnár, I., König, A., Khaw, L.E., Haydock, S.F., Oliynyk, M., Caffrey, P., Cortés, J., Lester, J.B., et al. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. Proc. Natl. Acad. Sci. USA *92*, 7839–7843.

Scott-Craig, J.S., Panaccione, D.G., Pocard, J.A., and Walton, J.D. (1992). The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus Cochliobolus carbonum is encoded by a 15.7-kilobase open reading frame. J. Biol. Chem. *267*, 26044–26049.

Sieber, S.A., and Marahiel, M.A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem. Rev. *105*, 715–738.

Simunovic, V., Zapp, J., Rachid, S., Krug, D., Meiser, P., and Müller, R. (2006). Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and trans-acting acyltransferases. Chembiochem *7*, 1206–1220.

Sucipto, H., Wenzel, S.C., and Müller, R. (2013). Exploring chemical diversity of α-pyrone antibiotics: molecular basis of myxopyronin biosynthesis. Chembiochem 14, 1581–1589.

Sudek, S., Lopanik, N.B., Waggoner, L.E., Hildebrand, M., Anderson, C., Liu, H., Patel, A., Sherman, D.H., and Haygood, M.G. (2007). Identification of the putative bryostatin polyketide synthase gene cluster from "Candidatus Endobugula sertula", the uncultivated microbial symbiont of the marine bryozoan Bugula neritina. J. Nat. Prod. 70, 67–74.

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.

Tsai, S.C., Lu, H., Cane, D.E., Khosla, C., and Stroud, R.M. (2002). Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases. Biochemistry 41, 12598–12606.

Viehrig, K., Surup, F., Harmrolfs, K., Jansen, R., Kunze, B., and Müller, R. (2013). Concerted action of P450 plus helper protein to form the amino-hydroxy-piperidone moiety of the potent protease inhibitor crocapeptin. J. Am. Chem. Soc. 135, 16885–16894.

Weber, G., Schörgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994). The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame. Curr. Genet. 26, 120–125.

Weckwerth, W., Miyamoto, K., Iinuma, K., Krause, M., Glinski, M., Storm, T., Bonse, G., Kleinkauf, H., and Zocher, R. (2000). Biosynthesis of PF1022A and related cyclooctadepsipeptides. J. Biol. Chem. 275, 17909–17915.

Wenzel, S.C., and Müller, R. (2009). The biosynthetic potential of myxobacteria and their impact in drug discovery. Curr. Opin. Drug Discov. Devel. 12, 220–230.

West, A.H., and Stock, A.M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem. Sci. 26, 369–376.

Williams, J.S., Thomas, M., and Clarke, D.J. (2005). The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in Photorhabdus luminescens TT01. Microbiology 151, 2543–2550.

Xiang, L., and Moore, B.S. (2002). Inactivation, complementation, and heterologous expression of encP, a novel bacterial phenylalanine ammonia-lyase gene. J. Biol. Chem. 277, 32505–32509.

Yadav, G., Gokhale, R.S., and Mohanty, D. (2003). Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. J. Mol. Biol. 328, 335–363.

Zaleta-Rivera, K., Xu, C., Yu, F., Butchko, R.A., Proctor, R.H., Hidalgo-Lara, M.E., Raza, A., Dussault, P.H., and Du, L. (2006). A bidomain nonribosomal peptide synthetase encoded by FUM14 catalyzes the formation of tricarballylic esters in the biosynthesis of fumonisins. Biochemistry 45, 2561–2569.

Ziemert, N., Podell, S., Penn, K., Badger, J.H., Allen, E., and Jensen, P.R. (2012). The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. PLoS ONE 7, e34064.