Mapping the biosynthetic pathway of a hybrid polyketide-nonribosomal peptide in a metazoan

Likui Feng 1,2, Matthew T. Gordon 1, Ying Liu 1, Kari B. Basso 1 & Rebecca A. Butcher 1

Polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) hybrid systems typically use complex protein-protein interactions to facilitate direct transfer of intermediates between these multimodular megaenzymes. In the canal-associated neurons (CANs) of Caenorhabditis elegans, PKS-1 and NRPS-1 produce the nemamides, the only known hybrid polyketide-nonribosomal peptides biosynthesized by animals, through a poorly understood mechanism. Here, we use genome editing and mass spectrometry to map the roles of individual PKS-1 and NRPS-1 enzymatic domains in nemamide biosynthesis. Furthermore, we show that nemamide biosynthesis requires at least five additional enzymes expressed in the CANs that are encoded by genes distributed across the worm genome. We identify the roles of these enzymes and discover a mechanism for trafficking intermediates between a PKS and an NRPS. Specifically, the enzyme PKAL-1 activates an advanced polyketide intermediate as an adenylate and directly loads it onto a carrier protein in NRPS-1. This trafficking mechanism provides a means by which a PKS-NRPS system can expand its biosynthetic potential and is likely important for the regulation of nemamide biosynthesis.
In the past 15 years, it has become clear that animal genomes encode biosynthetic pathways for many microbial-like secondary metabolites. Although in some cases these pathways were acquired from microorganisms through horizontal gene transfer, in the majority of cases these pathways are thought to have evolved independently in animals. Investigating the rich biochemistry of animals thus has the potential to uncover many important chemical insights and biosynthetic strategies. Furthermore, these studies are poised to reveal how the animal biosynthetic machinery is integrated with the higher-order complexity found in animals, including multiple organelles and tissues, integrated signaling pathways, and complex life-history traits.

Nematodes, in particular, have been shown to have a rich biosynthetic repertoire. In fact, many nematode genomes encode multi-module Type I PKSs and NRPSs for assembly-line-type biosynthesis of polyketides and nonribosomal peptides. Using the megasynthetases PKS-1 and NRPS-1, the model nematode Caenorhabditis elegans has been shown to produce a remarkable class of hybrid polyketide-nonribosomal peptides known as the nemamides in two essential neurons, the canal-associated neurons (CANs) (Fig. 1a). The nemamides promote the survival of the worm during starvation possibly through their effects on insulin gene expression and in a manner that is at least partially independent of the transcription factor DAF-16/FOXO. Since PKS-1 and NRPS-1 homologs are found in most nematode species, including parasitic ones, it is likely that other nematode species produce nemamide-like molecules and that these natural products play a conserved role in nematode biology. However, the biosynthesis of the nemamides, as well as how it is regulated, is poorly understood.

In Type I PKSs and NRPSs, each module is responsible for incorporating a different building block, such as malonyl- or methylmalonyl-CoA in the case of PKSs and proteinogenic or nonproteinogenic amino acids in the case of NRPSs. Acyltransferase (AT) domains are responsible for loading building blocks onto acyl carrier proteins (ACPs) in PKSs, while adenylation (A) domains are responsible for loading building blocks onto peptidyl carrier proteins (PCPs) in NRPSs. The building blocks are then linked together by ketosynthase (KS) domains in PKSs and condensation (C) domains in NRPSs. Additional domains can be present in the modules, such as ketoreductase (KR) and dehydratase (DH) domains in PKSs that control the oxidation state of the β-carbon, as well as methyltransferase and aminotransferase domains in PKSs and NRPSs that can further modify the natural product. Once synthesis of the natural product on a PKS or NRPS is complete, it is typically cleaved from the synthetase by a thioesterase (TE) domain. Further complexity can be introduced through hybrid systems that combine PKS and NRPS modules. These two types of modules can be encoded in the same protein, but they are often encoded in separate proteins. In this latter case, protein–protein interactions facilitate the direct passage of biosynthetic intermediates between the two proteins.

Domain analysis of PKS-1 and NRPS-1 provides a number of tantalizing clues indicating that the biosynthesis of the nemamides deviates from a canonical pathway in several respects. Both PKS-1 and NRPS-1 contain stretches of protein sequence with no homology to known enzymatic domains, and the enzymatic domains that can be identified are out of order and/or have diverged significantly from those found in other systems. For example, the substrate preferences of the A domains cannot be predicted based on the presence of key residues present in their active sites. Furthermore, there are no obvious domains that would enable certain structural features present in the nemamides, such as the O-methyl and amino groups, to be incorporated. Given that biosynthetic genes are not clustered in animals as they are in microorganisms, the identification of these missing domains in the worm represents a challenge. Additionally, while there is a TE domain at the C-terminus of NRPS-1, which presumably cleaves the natural product from the synthetase, there is also, strangely, a TE domain at the C-terminus of PKS-1 of unknown function.

Here, we map the biosynthesis of a complex metabolite in an animal system through genetic manipulation of the biosynthetic genes in vivo followed by comparative metabolomics. We show that the biosynthesis of the nemamides requires at least seven genes distributed across the worm genome that are united by their common expression in the CANs (Fig. 1a, b). Furthermore, we uncover the biosynthetic roles of these genes and show that the polyketide-ACP ligase PKAL-1 acts in trans and is required.

![Diagram](image-url)

Fig. 1 Enzymes required for nemamide biosynthesis. a The domain organization of PKS-1 and NRPS-1 is shown, along with five additional free-standing enzymes (NEMT-1, PKAL-1, C32E8.6, C24A3.4, and Y71H2B.1) that were demonstrated in this study to be required for nemamide biosynthesis. To facilitate annotation of the mutant worm strains generated in this study, the enzyme domains have been numbered according to the order of their appearance in PKS-1 and NRPS-1. The ACP domain was identified and its functional role was confirmed in this study. Domain abbreviations: acyl carrier protein (ACP), acyltransferase (AT), ketosynthase (KS), ketoreductase (KR), dehydratase (DH), peptidyl carrier protein (PCP), adenylation (A), condensation (C), thioesterase (TE). b The approximate chromosomal location in C. elegans of pks-1, nrps-1, and the five additional genes demonstrated to be required for nemamide biosynthesis in this study.
for the trafficking of intermediates between PKS-1 and NRPS-1. PKAL-1 represents a unique enzyme in that it loads a complex polyketide intermediate onto an NRPS for further chemical elaboration.

Results

NRPS-1 is responsible for incorporating all of the amino acid components of the nemamides. Assuming a linear, assembly-line mechanism in nemamide biosynthesis, we predicted that the β-Ala moiety in the nemamides is installed by the C-terminal NRPS module of PKS-1 and that the two D-Asn and final L-Asn moieties are installed by the NRPS modules of NRPS-1. However, the lack of sequence homology of the A domains in PKS-1 and NRPS-1 to bacterial and fungal A domains precluded predictions of the amino acid substrate specificities of these domains that might have supported our model. Furthermore, our inability to express in Escherichia coli any of the PKS-1 or NRPS-1 A domains (either as excised domains or as part of multi-domain constructs) precluded us from analyzing the substrate preferences of the A domains in vitro. Therefore, we decided to inactivate specific domains in PKS-1 and NRPS-1 in the worm in order to map their biosynthetic roles. First, we used CRISPR–Cas9 to inactivate the TE domain in NRPS-1 by replacing the catalytic serine with alanine to generate strain \( \text{nrps-1}[\text{TE}_2 \_\text{S2803A}] \) (Supplementary Fig. 1). This worm strain does not make the nemamides, consistent with the role of the NRPS-1 TE domain in cleaving nemamide from the synthetase through the formation of a macrolactam (Supplementary Fig. 2).

Unexpectedly, the \( \text{nrps-1}[\text{TE}_2 \_\text{S2803A}] \) strain accumulates a number of intermediates in the biosynthesis of the nemamides, one with no amino acids incorporated (3), one with β-Ala incorporated (4), one with β-Ala-D-Asn incorporated (5), and one with β-Ala-D-Asn-D-Asn incorporated (6) (Fig. 2a, b; Supplementary Fig. 3). Given that the nemamides are produced in very low amounts in *C. elegans*, and given that the biosynthetic intermediates are produced at even lower amounts, the intermediates had to be partially purified to enable identification. Totally, 2–3 L of worms grown in high-density axenic culture enabled the production of 3–5 g of worms, which were used to generate extracts that were then purified through two chromatographic steps. Intermediates were followed based on the characteristic ultraviolet (UV) spectrum of the triene or tetaene moiety that is present in them. Fortunately, this UV signature appears to be quite unique to nemamide and nemamide intermediates in *C. elegans* extracts. We verified the accumulation of intermediates 3–6 in the NRPS-1 TE domain mutant strain using high-resolution liquid chromatography–mass spectrometry (LC-MS/MS) (Supplementary Figs. 4–7). The amounts of the intermediates detected in this mutant were less than 10% of the mean amount of nemamide A in wild-type worms (Supplementary Fig. 3). The fact that the NRPS-1 TE domain mutant strain does not accumulate the linear form of nemamide A with the last L-Asn incorporated, suggests that this intermediate likely remains covalently attached to the NRPS-1 terminal PCP (PCP$_4$) in the mutant.

We hypothesized that we might be able to map the biosynthesis of the nemamides by NRPS-1 by targeting individual domains in NRPS-1 and analyzing the intermediates that accumulate. We began with domains acting at the end of the nemamide biosynthetic pathway and worked our way backward. The \( \text{nrps-1}[\text{C}_4 \_\text{S1934N}_\text{D1971N}] \) strain, in which an active site Ser of the NRPS-1 C$_4$ domain was mutated, does not make the nemamides and accumulates intermediates 3–6 (Fig. 2a, c; Supplementary Figs. 1–3). This result suggests that the C$_4$ domain in NRPS-1 incorporates the L-Asn into the growing chain before cyclization by the NRPS-1 TE domain to form the nemamides. The \( \text{nrps-1}[\text{C}_3 \_\text{H1486A}] \) strain, in which the catalytic histidine of the NRPS-1 C$_3$ domain was mutated, does not make the nemamides and accumulates intermediates 3 and 4 (Fig. 2a, d; Supplementary Figs. 1–3). The \( \text{nrps-1}[\text{A}_1 \_\text{G2337D}] \) strain, in which a key glycine in the NRPS-1 A$_3$ domain was mutated, also accumulates predominantly intermediates 3 and 4 (Fig. 2a, e; Supplementary Figs. 1–3). These results suggest that the C$_3$ and A$_3$ domains incorporate the two β-Asn moieties into the nemamides. Certain C domains catalyze both the epimerization and condensation of the amino acids that they incorporate, but none of the C domains in NRPS-1 have the characteristic features of these domains (Supplementary Fig. 8). Thus, it is unclear whether incorporation of the β-Asn moieties into the nemamides involves the A$_3$ domain loading D-Asn or the A$_2$ domain loading l-Asn, followed by epimerization to β-Asn by an unidentified epimerase.

The N-terminus of NRPS-1 contains a stretch of an unannotated sequence of 400 amino acids that has no obvious homology to PKS or NRPS domains, based on antiSMASH or BLAST analysis. Using the modeling program SWISS-model, however, we were able to identify an ACP domain between residues 259 and 361 of this sequence (annotated as ACP$_7$). The \( \text{nrps-1}[\text{ACP}_7 \_\text{S307V}] \) strain, in which the predicted site of phosphopantetheinylation in ACP$_7$ is mutated, does not make the nemamides, demonstrating that this ACP domain is functionally relevant (Supplementary Figs. 1 and 2). Surprisingly, this strain specifically accumulates intermediate 3, but not 4, revealing that the β-alanine residue in the nemamides is not incorporated by the C-terminal NRPS module of PKS-1, as we had originally proposed, but by the N-terminal NRPS module of NRPS-1 (Fig. 2a, f; Supplementary Fig. 3). In further support of this model, the \( \text{nrps-1}[\text{A}_2 \_\text{G964D}] \) strain, in which a key glycine in the NRPS-1 A$_2$ domain is mutated, accumulates intermediate 3, but not 4, indicating that A$_2$ loads β-alanine onto NRPS-1 for incorporation into the nemamides (Fig. 2a, g; Supplementary Figs. 1–3). Since disruption of the NRPS-1 ACP$_2$, A$_2$, A$_3$, C$_2$, and TE$_2$ domains leads to the accumulation of intermediate 3, this intermediate is most likely generated by PKS-1 and transferred onto NRPS-1 ACP$_2$ for further elongation. Furthermore, because intermediate 3 contains the amino and O-methyl groups, incorporation of these groups must precede the biosynthetic steps carried out by NRPS-1.

Role of the C-terminal NRPS module and TE domain in PKS-1.

Given that the β-alanine moiety is incorporated by NRPS-1, it was unclear whether the C-terminal NRPS module in PKS-1 would be required for nemamide biosynthesis. The domain organization of this NRPS module in PKS-1, as well as the sequence of the A domain in this module, are highly conserved across nematode evolution, suggesting that the NRPS module of PKS-1 does play an important role in the biosynthesis (Supplementary Table 1). Three mutant worm strains containing mutations in either the A$_1$, C$_1$, or PCP$_2$ domains of the NRPS module in PKS-1 were generated (Fig. 3; Supplementary Fig. 1). While the C$_1$ domain mutant could still produce some nemamides (less than 40%), the A$_1$ domain mutant made only very minor amounts of nemamides and the PCP$_2$ domain mutant did not make any at all (Fig. 3; Supplementary Fig. 9). Thus, although the C-terminal NRPS module of PKS-1 is not involved in incorporating β-alanine, it is required for nemamide biosynthesis. None of these strains accumulated any intermediates with UV signatures characteristic of trienes or tetaenones, suggesting that if any such intermediates do accumulate, they remain covalently linked to the synthetase.

The NRPS module in PKS-1 is followed by a C-terminal TE domain of unknown function. Comparison of this TE domain to
Identification of trans-acting enzymes involved in nemamide biosynthesis. We hypothesized that certain steps in nemamide biosynthesis would require additional enzymes beyond PKS-1 and NRPS-1 and that these enzymes would likely also be expressed in the CANs. Single-cell gene expression profiling has been performed in multiple cell types in *C. elegans*, including the CANs31,32,33,34,35,36,37,38 genes, including *pks-1* and *nrps-1*, show selective expression in the CANs that is at least fivefold higher than in the next tissue in which they are expressed (Supplementary Table 2)32,33. Based on their homology and predicted enzymatic activities, we screened a number of these genes by analyzing nemamide production in loss-of-function mutant worm strains. These data show that at least five additional genes are required for nemamide biosynthesis, F49C12.10, T20F7.7, C32E8.6, C24A3.4, and Y71H2B.1 (Fig. 4a; Supplementary Fig. 10). F49C12.10, which we named NEMT-1 (C32E8.6) and Y71H2B.1 (C32E8.6) genes to inter-rescue suggests that they play distinct roles in nemamide biosynthesis (Supplementary Fig. 11). T20F7.7 was previously named ACS-9 due to its homology to acyl-CoA synthetase, but the failure of the T20F7.7 and C32E8.6 genes to inter-rescue suggests that they play distinct roles in nemamide biosynthesis (Supplementary Fig. 11). T20F7.7 was previously named ACS-9 due to its homology to acyl-CoA synthetase, but
The nemt-1 mutant strain accumulates intermediate 7, which is similar to 3, but lacks the methyl group (Fig. 4c, e; Supplementary Fig. 14). This result indicates that NEMT-1 is the O-methyltransferase that is responsible for the methoxy group that is present in the nemamides. The nemt-1 mutant strain also produces 8, which is similar to nemamide A, but lacks the methyl group (Fig. 4d, e; Supplementary Fig. 15). The fact that the nemt-1 mutant can make desmethyl-nemamide 8 suggests that PKAL-1 can load 7 onto the ACP7 domain of NRPS-1, even though it is lacking the methyl group. However, because the nemt-1 mutant accumulates intermediate 7, PKAL-1 and/or NRPS-1 likely prefers methylated substrates over unmethylated ones (Fig. 4e).

Unlike the nemt-1 single mutant, the nemt-1: pks-1[TE1_S7593A] double mutant does not accumulate intermediate 7 or desmethyl-nemamide 8, suggesting a possible model in which NEMT-1 functions downstream of the PKS-1 TE1 domain. Thus, our genetic data suggest that NEMT-1 methylates intermediate 7 to form 3, and then PKAL-1 is involved in loading 3 onto the ACP7 domain of NRPS-1 for further extension.

No biosynthetic intermediates were found in Y71H2B.1, C24A3.4, or C32E8.6 loss-of-function mutant worm strains. Given that intermediates also do not accumulate in pks-1 mutant worm strains, these three genes may function in earlier stages of nemamide biosynthesis, potentially either in the initiation or in other steps involving PKS-1.

Biochemical activity of PKAL-1. Based on its sequence homology, we hypothesized that PKAL-1 might function similarly to a fatty acyl-CoA ligase (FACL). That is, PKAL-1 might activate 3 as the adenylate and then react that intermediate with CoA to form a CoA-thioester that is subsequently loaded onto the ACP7 domain of NRPS-1. To biochemically characterize the role of PKAL-1 in trafficking intermediates between PKS-1 and NRPS-1, we cloned the enzyme from a cDNA library, expressed it in *E. coli*, and purified it for biochemical characterization. We then incubated PKAL-1 with fatty acids of various lengths, ATP, and CoA and analyzed the products by LC–MS. PKAL-1 can activate a variety of medium and long-chain fatty acids as the corresponding fatty acyl-AMP but cannot further convert them to the corresponding fatty acyl-CoA (Fig. 5a, Supplementary Fig. 16). A negative control, PKAL-1(K488A), in which a lysine predicted to be important for catalysis was mutated, showed no activity. These results suggest that PKAL-1 is analogous to a fatty acyl-AMP ligase (FAAL) rather than a FACL. FAAL enzymes activate fatty acids as the corresponding fatty acyl-AMP but then transfer the fatty acyl group to the phosphopantetheinyl arm of a carrier protein instead of CoA. FAAL enzymes typically have an insertion motif that prevents the movement between the larger N-terminal domain and the smaller C-terminal domain that occurs between the adenylation and CoA-ligase reactions. Certain FAAL enzymes lack the insertion motif but instead have additional interactions between the N-terminal and C-terminal domains that are thought to prevent the CoA-ligase reaction from occurring. Although sequence alignment of PKAL-1 demonstrates that it is missing the insertion motif (Supplementary Fig. 17), structural modeling of PKAL-1 indicates that it might lack CoA ligase activity due to the absence of an effective binding site for CoA (Supplementary Fig. 18).

FAAL enzymes play an important role in the biosynthesis of lipid-modified polyketides and nonribosomal peptides, such as mycolic acids that are essential for mycobacterial growth, isonitrile lipopeptides that promote mycobacterial virulence, and diverse lipopeptides that are widespread in cyanobacteria. In order to
determine whether PKAL-1 can load fatty acids that are similar in length to intermediate 3 onto the ACP7 domain of NRPS-1, the excised ACP7 domain was first co-expressed in E. coli with the promiscuous phosphopantetheinyl transferase Sfp, enabling purification of the holo-ACP7 domain. This holo-ACP7 domain was incubated with PKAL-1, ATP, and fatty acids of various lengths, and the products were analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF). PKAL-1 could efficiently load fatty acids that were 9–18 carbons in length and could load a minor amount of C20 fatty acids. (Fig.5b). This result is consistent with a role for PKAL-1 in activating and loading 3, which is 18 carbons in length. PKAL-1 was unable to load fatty acids onto a different ACP, ACP1, from PKS-1 (Supplementary Fig. 19). Thus, PKAL-1 must engage in specific interactions with the ACP7 domain of NRPS-1 in order to load fatty acids onto it. To characterize further PKAL-1’s substrate preferences, we utilized an enzyme-coupled continuous kinetic assay in which hydroxylamine was used instead of an ACP as the substrate acceptor. Importantly, PKAL-1 is much more active towards long-chain fatty acids than short- and medium-chain fatty acids, with C16 being preferred over C14, which is in turn preferred over C12, which is in turn preferred over C10 and C8 (Supplementary Fig. 20). Fatty acids longer than C16 could not be tested due to solubility issues and/or detergent effects. The kinetic assay thus demonstrates that the substrate preferences of PKAL-1 are consistent with its proposed role in activating and loading 3. While FAAL enzymes have been shown to activate fatty acids to initiate polyketide and nonribosomal peptide biosynthesis, PKAL-1 is unusual in that it activates an advanced polyketide intermediate for loading onto an NRPS for further elaboration.

**Discussion**

Using comparative metabolomics, we have mapped the biosynthetic pathway to the nemamides, a remarkable family of hybrid polyketide-nonribosomal peptides biosynthesized in the CANs of...
**Fig. 5 In vitro activity of PKAL-1 against fatty acid substrates.** a Activity of PKAL-1 towards fatty acids of various lengths in the presence of ATP and CoA. Extracted ion chromatograms are representative of at least three independent experiments. The identities of the peaks were verified based on their masses in both positive and negative modes, as well as their UV spectra. The catalytic mutant PKAL-1(K488A) did not show any product formation. The inability of PKAL-1 to form CoA-thioesters was verified with synthetic standards (Supplementary Fig. 16). b MALDI-TOF analysis of the holo-ACP₂ domain from NRPS-1 incubated with PKAL-1 (light green) or PKAL-1(K488A) (gray), fatty acids of various lengths, and ATP. MALDI-spectra are representative of at least three independent experiments.

C. elegans. Furthermore, we have determined the biosynthetic roles of five additional enzymes from C. elegans that function in trans in nemamide biosynthesis. Although the different genes that are required for nemamide biosynthesis are encoded in disparate locations across the worm genome, these genes share a common feature in that they are expressed in the CANs. Our work has shown that at least 7 of the 38 genes that have enriched expression in the CANs are involved in nemamide biosynthesis, suggesting that nemamide biosynthesis may be a primary function of these enigmatic neurons.

Our data show that all of the amino acids in the nemamides, including the β-alanine, are incorporated by NRPS-1 and that the 3-terminal NRPS module of PKS-1 has other functions in nemamide biosynthesis (Fig. 6). Mutation of the PCP₂ domain of PKS-1 completely blocks nemamide production, and mutation of the A₁ domain of PKS-1 enables only trace amounts of nemamides to be produced; thus, the NRPS module of PKS-1 does play a role in nemamide biosynthesis. Furthermore, both the domain organization of this NRPS module in PKS-1 and the sequence of the A₁ domain have been conserved across nematode species. As we have not been able to identify an aminotransferase for incorporation of the amino group into the nemamides, this module may be involved in this process. A possible mechanism for the incorporation of an amino group into the nemamides is suggested by the biosynthesis of the β-amino fatty acid starter unit in the macrolactam family of antibiotics that includes BE-14106 and ML-449 [37–40]. The biosynthesis of the β-amino fatty acid is thought to require a free-standing A domain, a free-standing PCP domain, and a free-standing glycine oxidase to incorporate the amino group from glycine into an α-β unsaturated fatty acyl precursor [37,38]. It should be noted that a C₁ domain is not involved in this proposed mechanism, potentially explaining why the C₁ domain of PKS-1 is not required for nemamide biosynthesis. Importantly, however, we have not yet identified a candidate glycine oxidase in the worm genome, and thus, it is unclear whether the amino group in the nemamides is incorporated in an analogous fashion as the amino group is incorporated in BE-14106 and ML-449.

 Trafficking between PKS and NRPS enzymes typically involves protein–protein interactions between the megasynthetases that enable direct passage of intermediates from the carrier domain on an upstream PKS or NRPS to the KS or C domain on the downstream PKS or NRPS, respectively [16]. Here, we have shown that trafficking of intermediates between PKS-1 and NRPS-1 involves production of a complex polyketide intermediate by PKS-1 and NEMT-1 (i.e., 3) that is activated and loaded by PKAL-1 onto the N-terminal ACP domain of NRPS-1 for subsequent production of the nemamides (Fig. 6). Specifically, we have shown that mutant worms containing mutations in either PKAL-1 or the ACP domain of NRPS-1 accumulate 3. Furthermore, through biochemical experiments, we have demonstrated that PKAL-1 activates and loads fatty acid substrates of similar length to 3 onto the ACP domain of NRPS-1, but not onto other ACP domains. Since an unmethylated version of 3 (i.e., 7) accumulates in the nemt-1 mutant strain, it is likely that PKS-1 produces 7, which is then methylated by NEMT-1 to make 3 (Fig. 6). It remains possible, however, that NEMT-1 acts earlier in the biosynthetic pathway on an intermediate that is attached to the PKS-1 assembly line. However, for this alternative model to be correct, we must assume that in the nemt-1 mutant, the PKS-1 assembly line can continue to process unmethylated intermediates as well as it can the methylated ones that are present in wild-type worms. Furthermore, we must assume that in the nemt-1 mutant the unmethylated polyketide product 7 accumulates simply due to it being a non-preferred substrate for PKAL-1. This model seems less likely given that the nemt-1 strain is able to produce desmethyl-nemamide A (8), indicating that PKAL-1 (and NRPS-1) can process desmethyl intermediates. It is unclear why only small amounts of 7 and 3 accumulate in the nemt-1 and pkal-1 mutants, respectively, relative to the amount of nemamides in wild type worms. This result could indicate that these biosynthetic intermediates are less stable than the nemamides, or it could indicate that there is some type of feedback mechanism that prevents these intermediates from accumulating. The mechanism of trafficking intermediates between PKS-1 and NRPS-1 may be important for the movement of intermediates in nemamide biosynthesis across membranes or between cellular compartments.
The PKAL-1 enzyme is unique in that it activates an advanced polyketide intermediate as the adenylate and loads it onto the ACP, domain of NRPS-1 in order to biosynthesize nemamide A, 1. An analogous pathway biosynthesizes nemamide B, 2. Our data suggest that PKS-1 produces and releases 7, which is then methylated by NEMT-1 to generate 3. However, it is also possible that NEMT-1 acts earlier in the pathway on a polyketide intermediate attached to PKS-1, such that PKS-1 produces and releases 3, not 7 (see “Discussion”). Domain abbreviations: acyl carrier protein (ACP), peptidyl carrier protein (PCP), adenylation (A), condensation (C), and thioesterase (TE).

**Methods**

**Worm strains.** Worms were maintained on OP50 using standard methods. Strains used in this study were obtained from the Caenorhabditis Genetics Center or were generated through genome editing by CRISPR–Cas9 or through transgenesis (Supplementary Table 3). Double mutants were generated from single mutants, using standard genetic crossing methods. The presence of alleles was verified through single worm polymerase chain reaction (PCR) using the primers in Supplementary Table 4, and the PCR products were analyzed by restriction site digestion, as shown in Supplementary Table 5.

**Single worm PCR and CRISPR-Cas9.** Most mutant strains containing deletions and point mutations were generated based on the Fire lab’s CRISPR–Cas9 protocol. The concentration of the Cas9 vector used was 50 ng/µL. The plasmid for expressing the dpy-10 sgRNA was used at 25 ng/µL, and the other plasmids for expressing the sgRNA sequences for target genes were used at 50–100 ng/µL (see Supplementary Table 6 for sequences). The dpy-10 (cn64) donor oligonucleotide was used at a concentration of 500 nM, and other donor oligonucleotides for generating desired mutations were used at 500–750 nM (see Supplementary Table 7 for sequences). To generate deletion mutants, no donor oligonucleotide was used for either dpy-10 or targeted genes. After microinjection, F1 worms with dpy (for deletions) and roller (for point mutations) phenotypes were picked for single worm PCR using the primers listed in Supplementary Table 4 and restriction digestion (only for mutants with point mutations) of the PCR products as shown in Supplementary Table 5. The PCR products of positive candidates were sequenced to verify the deletion (or targeted genes). After microinjection, F1 worms with dpy (for deletions) and roller (for point mutations) phenotypes were picked for single worm PCR using the primers listed in Supplementary Table 4 and subsequent restriction digestion of the PCR products as shown in Supplementary Table 5. The PCR products of positive candidates were sequenced to confirm their alleles.

**Transgenesis.** Transcriptional reporter plasmids were made by amplifying the promoter from genomic DNA using the primers in Supplementary Table 8 and inserting the product into pPD114.108. Translational reporter plasmids were made by amplifying the promoter and gene from genomic DNA using the primers in Supplementary Table 8 and inserting the product into pBS77–SL2-mCherry. The reporter plasmids were injected into the corresponding mutant worm strains at 50 ng/µL. The coel::dsred plasmid was used as a co-injection marker at 25 ng/µL, and the total concentration of DNA injected was 100 ng/µL including pUC18. At least three independent transgenic lines were analyzed.

**Small-scale worm extraction for nemamide production.** Wild-type and mutant worm strains were grown at room temperature on two NGM agar plates (10 cm)
spread with 0.75 mL 25X OP50 until the food on the plates was almost gone. Then, the worms were transferred to a 1 L Erlenmeyer flask containing 5 medium (300 mL). The worm cultures were grown at 23 °C for 5 days until no food was left and were fed with 3.5 mL of 25X OP50 every day. For sample collection, the culture flasks were placed in an ice bath for 30 min to 1 h to settle the worms. Then, the worms were transferred from the bottom of the flask to a 30 mL centrifuge tube and were centrifuged (220 × g for 5 min) to separate the worms from the medium. The supernatant was repeated until no worms were removed from the flasks. The collected worms were washed with water three times and centrifuged (220 × g for 5 min), and then they were soaked in 10 mL of water for 1 h in a shaking incubator (22.5 °C, 225 rpm) to remove bacteria and ciliate. The worms were collected by centrifugation and freeze-dried. The dried worm pellets were ground with sea sand (2 g sand per 200 mg dried worms) using a mortar and pestle. The ground worms were extracted with 15 mL of 190 proof ethanol for 3.5 h, and the extract was centrifuged (2700 × g for 20 min). The supernatant was collected and dried using a speed vac. The dried worm samples were each resuspended in 100 μL of methanolic sonication (if needed), and centrifuged (18,400 × g for 1 min) before analysis by LC-MS. The samples were analyzed using a Luna 5 μm C18 (2) column (100 × 4.6 mm; Phenomenex) coupled with an Agilent 6130 single mass spectrometer operating in single ion monitoring (SIM) mode for nemamide A ([(M + Na)+ m/z 737]) and nemamide B ([(M + Na)+ m/z 755]). The following solvent gradient was used with a flow rate of 0.7 mL/min: 95% buffer A, 5% buffer B, 0 min; 0% buffer A, 100% buffer B, 20 min; 0% buffer A, 100% buffer B, 22 min; 95% buffer A, 5% buffer B, 23 min; 95% buffer A, 5% buffer B, 26 min (buffer A, water with 0.1% formic acid; buffer B, acetonitrile with 0.1% formic acid). Two pkal-1 mutants, RBAS9 (reb21 allele) and RBAS9 (reb28 allele) were analyzed, and neither produced any nemamides. Two C32E8.6 mutants, RB06 (reb23 allele) and RA661 (reb24 allele) were also analyzed, and neither produced any nemamides.

Large-scale worm extraction for biosynthetic intermediates. Egg prep of wild-type or a mutant strain was performed in order to inoculate 5 mL of C. elegans medium, which was grown for 7–10 d. This 5 mL culture was then used to inoculate 500 mL of C. elegans medium, which was grown for 7–10 d. The 50 mL culture was then used to inoculate 500 mL of C. elegans medium in a 2.8 L baffled flask. The culture was grown for 7–10 d. C. elegans medium containing cow’s milk was initiated by adding 2 μL of 2 mg/mL purified recombinant Escherichia coli (E. coli) alkaline phosphatase (pKal-1) to an Erlenmeyer flask (350 mL each) to give four fractions (A, B, C, and D). The collected worms were washed with water for 20 min. The supernatant was collected and dried using a speed vac. The dried worm samples were each resuspended in 100 μL of methanolic sonication (if needed), and centrifuged (18,400 × g for 1 min) before analysis by LC-MS. The samples were analyzed using a Luna 5 μm C18 (2) column (100 × 4.6 mm; Phenomenex) coupled with an Agilent 6130 single mass spectrometer operating in single ion monitoring (SIM) mode for nemamide A ([(M + Na)+ m/z 737]). The following solvent gradient was used with a flow rate of 0.7 mL/min: 95% buffer A, 5% buffer B, 0 min; 0% buffer A, 100% buffer B, 20 min; 0% buffer A, 100% buffer B, 22 min; 95% buffer A, 5% buffer B, 23 min; 95% buffer A, 5% buffer B, 26 min (buffer A, water with 0.1% formic acid; buffer B, acetonitrile with 0.1% formic acid). Two pkal-1 mutants, RBAS9 (reb21 allele) and RBAS9 (reb28 allele) were analyzed, and neither produced any nemamides. Two C32E8.6 mutants, RB06 (reb23 allele) and RA661 (reb24 allele) were also analyzed, and neither produced any nemamides.

**Plasmid construction, protein overexpression, and purification.** All genes and excised domains were amplified by PCR using Phusion polymerase (New England Biolabs) from a C. elegans DNA library. Specifically, pkal-1, flp-1, ACP, and rep-1, ACPs, were amplified with specific primers (Supplementary Table 9), and the PCR products were inserted using NcoI and Nof into the pET16b-KH01 vector (a modified version of pET-16b) such that the expressed proteins were expressed with C-terminal His tag. All of the sequences were verified by sequencing. The pET16b-KH01 (mutation) was generated via Q5 site-directed mutagenesis kit (New England Biolabs) using a specific primer pair (Supplementary Table 9) and verified through sequencing. Additional mutations were introduced into ACP and ACP2 to allow visualization of the carrier proteins by UV/vis, enabling purification by FPLC and concentration estimation by NanoDrop. The second and third residues (after the start codon) in ACP were modified to Tyr and Trp, respectively, and the second residue (after the start codon) in ACP2, was modified to Trp and Tyr, via the Q5 site-directed mutagenesis kit (Supplementary Table 9). The PKAL-1 construct was transferred into BL21 (DE3) cells, and the cells were grown in LB broth with 150 μg/mL ampicillin at 37 °C to OD600 0.6–0.8, cooled down on the ice for 10 min, and induced with 0.3 mM IPTG at 16 °C for 20 h. The ACP and ACP2 constructs were each co-transformed with pACYC Duet-sfp into BL21 (DE3) co-expression. The cells were grown in LB broth with 150 μg/mL ampicillin and 34 mg/L chloramphenicol at 37 °C to OD600 0.3–0.4, and the temperature was lowered to 16 °C for expression. Totally, 30 min prior to induction, cultures were supplemented with 2.5 mM calcium pantothenate, and once cells reached OD600 0.6–0.8, protein expression was induced with 0.6 mM IPTG at 16 °C for 20 h. All purification steps were carried out at 4 °C. Briefly, cells were collected by centrifugation at 3000 × g for 10 min, and resuspended in lyso buffer (20 mM Tris, 500 mM NaCl, pH 7.5). The cells were then lysed by a microfluidizer three times and centrifuged at 20,000 × g for 20 min. The supernatant was incubated with 10% and 10% equilibrated Nickel-resin (Thermo Scientific) for 1 h by shaking on ice. The resin was washed with 15 mL of lysis buffer, 15% of wash buffer (200 μL Tris, 500 μM NaCl, 20 mM imidazole, pH 7.5), and eluted with wash buffer containing 250 mM imidazole. During the purification of PKAL-1 for use in the kinetic assay, 1 mM ATP was included in the wash buffer. For PKAL-1, the eluted sample was concentrated and loaded onto an FPLC connecting to a Superdex 200 gel filtration column (GE healthcare) with buffer (200 μM Tris, 100 mM NaCl pH 7.5). Protein concentrations were determined by using Quick Start Dye reagent (Bio-Rad) with 2 mg/mL bovine serum albumin used as a standard (for PKAL-1) or by NanoDrop (for the carrier proteins). Purified proteins were flash-frozen in 10% glycerol and stored at −80 °C. Mass analysis of the carrier proteins expressed individually compared to co-expression with Sfp showed complete conversion of the carrier proteins from the apo to the holo form.
reaction was vortexed and centrifuged. A 5 μl supernatant was used for LC–MS analysis on an Agilent 6130 single quadrupole mass spectrometer in both positive and negative full-scan modes, mass range 150–1500, 125 V fragmentation voltage, 0.15 min peak width, and 2.20 s cycle length. Mobile phase A was water with 10 mM ammonium acetate, and mobile phase B was acetonitrile. The LC gradient was started from 95% A for 2 min and then ramped up to 100% B over 24 min.

MALDI-TOF MS analysis. A 15 μl reaction mixture contained 100 mM Tris buffer at pH 7.8, 10 mM MgCl₂, 1 mM DTT, 1 mM ATP, 100 μM fatty acid substrate, and 100 μM holo-ACP, or 100 μM holo-ACP. The reactions were initiated by adding 2 μl of 2 mg/ml purified PKAL-1 or PKAL-1(K488A) at 25 °C for 2 h. Samples were diluted 1:10 in ultrapure water and spotted onto a ground plate 1:1 with a matrix containing saturated sinapinic acid in 70% acetonitrile. A Bruker AutoFlex LRF MALDI-TOF (Bruker Daltonics) equipped with a Smartbeam-II UV laser was used to analyze the ACP masses, using the positive linear mode at a mass range of 5000–20,000 Da. Laser power was used at the threshold level required to generate the signal until suitable data were obtained. The instrument was calibrated with Protein Calibration Standard 1 (Bruker Daltonics) under the molecular weights of the samples (typically, mixtures of apo myoglobin and bovine serum albumin using doubly charged, singly charged, and dimer peaks, as appropriate). All data were analyzed using flexAnalysis software (Bruker). Every single spectrum was an average of 500 laser shots, and the final spectra were generated using the sum of at least three single spectra.

Enzyme-coupled continuous kinetic assay for PKAL-1. The enzyme kinetics of PKAL-1 were determined through an enzyme-coupled spectrophotometric assay. Each 100 μl reaction mixture contained 100 mM Tris buffer pH 7.4, 1 mM dithiothreitol, 10 mM MgCl₂, 4 mM ATP, 0.9 mM phosphoenolpyruvate, 0.3 mM NADH, 2.5 U pyruvate kinase, 3.5 U lactate dehydrogenase, 10 U adenylate kinase (Sigma M3003, prepared according to manufacturer’s protocol), 100 mM buffered hydroxyamine, and the tested substrates in DMSO (final volume 2.5%). The kinetic assay was initiated by the addition of an enzyme and run at 22 °C. No activity was detected in the assay when initiating with PKAL-1(K488A). GraphPad Prism was used to calculate the apparent kinetic constants.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The LC-MS/MS files for 3, 4, 5, 6, 7, and 8 have been deposited in the MetaboLights database under accession code MTBLS313. Source data are provided with this paper. All other data supporting the findings of this study are available within the article and its Supplementary Information files or from the corresponding author upon request. Source data are provided with this paper.

Received: 9 October 2020; Accepted: 30 June 2021;
Published online: 13 August 2021

References

1. Torres, J. P. & Schmidt, E. W. The biosynthetic diversity of the animal world. J. Biol. Chem. 294, 17684–17692 (2019).
2. Ganley, J. G. & Derbyshire, E. R. Linking genes to molecules in eukaryotic systems: a comprehensive resource for the genome mining of non-adenylating enzyme. Nat. Chem. Biol. 14, 730–737 (2018).
3. Trivedi, O. A. et al. Enzymic activation and transfer of fatty acids as acyl-CoA thioesters from non-adenylating enzyme. Nat. Chem. Biol. 14, 730–737 (2018).
4. Cao, J. et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357, 661–667 (2017).
5. Hutter, H., Ng, M. P. & Chen, N. GExplore: a web server for integrated queries of protein domain genomes, expression and mutant phenotypes. BMC Genomics 10, 529 (2009).
6. Perez, E. et al. Molecular dissection of the role of two methyltransferases in the biosynthesis of phthiocerol dimeric mycosporine in the Mycobacterium tuberculosis complex. J. Biol. Chem. 279, 42584–42592 (2004).
7. Purwantini, E., Daniels, L. M. & Mukhopadhyay, B. FAD/FH4 is required for phthiocerol dimericmycosporine synthesis in mycobacteria. J. Bacteriol. 198, 2020–2028 (2016).
8. Wang, N. et al. Natural separation of the acyl-CoA ligase reaction results in a more efficient non-adenylating enzyme.
9. Nature 428, 441–445 (2004).
10. Arora, P. et al. Mechanistic and functional insights into fatty acid activation in Mycobacterium tuberculosis. Nat. Chem. Biol. 5, 166–173 (2009).
11. Leger, M. et al. The dual function of the Mycobacterium tuberculosis FadD32 is required for mycolic acid biosynthesis. Chem. Biol. 16, 510–519 (2009).
12. Liu, Z., Zoeger, T. R., Wang, F. & Sacchettini, J. C. Structures of Mycobacterium tuberculosis FadD10 protein reveal a new type of adenylation–forming enzyme. J. Biol. Chem. 288, 18473–18483 (2013).
13. Galica, T., Hrouzek, P. & Marses, J. Genome mining reveals high incidence of putative lipopeptide biosynthesis NRPS/PKS clusters containing fatty acyl-AMP ligase genes in biofilm-forming cyanobacteria. J. Phycol. 53, 985–998 (2017).
14. Harris, N. C. et al. Biosynthesis of isonicotinyl lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria. Proc. Natl Acad. Sci. USA 114, 7025–7030 (2017).
15. Lambalot, R. H. et al. A new enzyme superfAMILY—the phosphaesthetinyltransferases. Chem. Biol. 3, 923–936 (1996).
16. Drake, E. J. et al. Structures of two distinct conformations of halo-nonribosomal peptide synthetases. Nature 529, 235–238 (2016).
17. Estrada, P. et al. The pemoly-CoA synthetase BioW defines a new fold for adenylation–forming enzymes. Nat. Chem. Biol. 13, 668–674 (2017).
18. Jorgensen, H. et al. Biosynthesis of the macrocyclic lactam E8641 involves two distinct PKS systems and amino acid processing enzymes for generation of the aminocyclo starter unit. Chem. Biol. 16, 1109–1121 (2009).
19. Jorgensen, H. et al. Insights into the evolution of macrocyclic biosynthesis through cloning and comparative analysis of the biosynthetic gene cluster for a novel macrocyclic lactam, ML-449. Appl. Environ. Microbiol. 76, 283–293 (2010).
Acknowledgements

We thank Patrick McGrath and Yi Sheng for kindly providing plasmids for CRISPR-Cas9 and Steven Bruner for providing the pACYCDuet-sfp plasmid. We thank Qingyao Shou for advice and Manasi Kamat for help with mass spectrometry. Mass spectrometry analysis was performed at the University of Florida’s Mass Spectrometry Research and Education Center, which is supported in part by the NIH (S10 OD01758-01A1). We acknowledge the CGC, which is supported by the NIH Office of Research Infrastructure Programs (P40 OD010440), for providing some worm strains. This work was supported by the NIH (GM118775) and the NSF (CHE-1553050).

Author contributions

R.A.B. and L.F. designed the study, analyzed the data, and wrote the paper, which was reviewed by all authors; L.F. and M.T.G. performed the experiments; L.F., Y.L., and K.B.B. performed the LC-MS/MS analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-24682-9.

Correspondence and requests for materials should be addressed to R.A.B.

Peer review information Nature Communications thanks the anonymous reviewers for their contributions to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021