Asperphenamate biosynthesis reveals a novel two-module NRPS system to synthesize amino acid esters in fungi†

Wei Li,‡abc Aili Fan,‡abc Long Wang,abc Peng Zhang,a Zhiguo Liu,a Zhiqiang And a and Wen-Bing Yin *abc

Amino acid esters are a group of structurally diverse natural products with distinct activities. Some are synthesized through an inter-molecular esterification step catalysed by nonribosomal peptide synthetase (NRPS). In bacteria, the formation of the intra-molecular ester bond is usually catalysed by a thioesterase domain of NRPS. However, the mechanism by which fungal NRPSs perform this process remains unclear. Herein, by targeted gene disruption in *Penicillium brevicompactum* and heterologous expression in *Aspergillus nidulans*, we show that two NRPSs, ApmA and ApmB, are sufficient for the synthesis of an amino acid ester, asperphenamate. Using the heterologous expression system, we identified that ApmA, with a reductase domain, rarely generates dipeptidyl alcohol. In contrast, ApmB was determined to not only catalyse inter-molecular ester bond formation but also accept the linear dipeptidyl precursor into the NRPS chain. The mechanism described here provides an approach for the synthesis of new small molecules with NRPS as the catalyst. Our study reveals for the first time a two-module NRPS system for the formation of amino acid esters in nature.

Introduction

Amino acid esters and their derivatives represent a huge group of pharmaceutically important natural products. These products include the cytotoxic asperphenamate (1),¹ anticancer taxol,² antibiotic daptomycin,³ and the recently identified antibiotic teixobactin, which kills bacteria with no detectable resistance (Fig. 1).⁴ ¹ is a rare linear amino acid ester derived from fungal nonribosomal peptide that exhibits antitumor activity towards a number of cell lines.⁵ ⁶ It was first isolated from *Aspergillus flavipes* in 1977⁷ and then identified from a wide range of *Aspergillus*⁸ and *Penicillium* species.° The structure of ¹ contains two subunits, N-benzoylphenylalanine (2) and N-benzoylphenylalaninol (3), which are connected by an inter-molecular ester bond. An experiment performed by feeding the labeled precursor [U-¹⁴C]-⁴ into *P. brevicompactum*

---

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc02396k

‡ These authors contributed equally.

---

Cite this: Chem. Sci., 2018, 9, 2589

Received 27th May 2017
Accepted 23rd January 2018
DOI: 10.1039/c7sc02396k
rsc.li/chemical-science

This journal is © The Royal Society of Chemistry 2018
proved that the benzoyl groups of 1 are derived from the precursor phenylalanine (4). However, no genetic study has been reported for the biosynthesis of 1. The motivation to study the biosynthesis of 1 is not only because of its biological activity but also because of the structural formation of the rare linear amino acid ester bond. The elucidation of the biosynthesis of 1 at the molecular level will provide an opportunity to produce 1 using an engineered microorganism. Moreover, it will provide new insights into the biosynthesis of related natural products.

NRPSs usually have modular structures that are responsible for peptide elongation. Typically, an NRPS consists of an adenylation (A) domain, a thiolation (T) domain or PCP (peptidyl carrier protein), and a condensation (C) domain. A previous study demonstrated that bacterial NRPSs utilize thioesterase (TE) or C domains to perform esterification reactions. For example, the NRPS subunit DptD with a termination TE domain is involved in the biosynthesis of daptomycin in Streptomyces roseosporus. In this process, the TE domain is required for the formation of the daptomycin lariat structure and releasing of the peptide from the synthetase.

In contrast, the NRPS subunit DptD with a termination TE domain is involved in the biosynthesis of daptomycin in Streptomyces roseosporus. In this process, the TE domain is required for the formation of the daptomycin lariat structure and releasing of the peptide from the synthetase. For example, the NRPS subunit DptD with a termination TE domain is involved in the biosynthesis of daptomycin in Streptomyces roseosporus. In this process, the TE domain is required for the formation of the daptomycin lariat structure and releasing of the peptide from the synthetase. For example, the NRPS subunit DptD with a termination TE domain is involved in the biosynthesis of daptomycin in Streptomyces roseosporus. In this process, the TE domain is required for the formation of the daptomycin lariat structure and releasing of the peptide from the synthetase.

Results and discussion

To investigate the biosynthetic pathway of 1, we worked on the 1 producing strain P. brevicompactum. The genome of this strain was sequenced in 2013. First, compound 1 was isolated from a 3 day culture of P. brevicompactum on YES medium by repeated column purification and semi-preparative HPLC (see ESI† for further details). 1 was elucidated by NMR identification and the data corresponded well to the previously published data. Analysing the structure of 1 showed that it consisted of two subunits, 2 and 3. Therefore, we postulated that an NRPS containing at least two A-domains was required for catalysing its formation. By searching the genome of P. brevicompactum, we found 73 biosynthetic gene clusters (BGCs), including 14 NRPSs. The NRPS domain prediction was performed using the service NRPSpredictor2. The prediction indicated that 6 NRPSs containing two A-domains matched our previous predictions. Interestingly, cluster 1 contained two NRPSs harbouring four A-domains. It was also possible that 1 needed four A-domains for the synthesis because its structure was not symmetric. To narrow down the NRPS target, we assessed their expressions in transcriptional levels by reverse transcriptase PCR (RT-PCR). Five NRPSs, NRPS1.1, NRPS1.2, NRPS9, NRPS10, and NRPS49 were normally expressed under the tested conditions. These NRPSs are indicated as possibly involved in the biosynthesis of 1 (Fig. S1†).

To verify which NRPSs were involved in the biosynthesis of 1, we implemented a genetic deletion approach in P. brevicompactum. First, we examined the strain’s antibiotic resistance to hygromycin (hph) and determined the growth inhibition concentration of 100 µg ml⁻¹. Next, we made the protoplasts with modifications according to the previously reported method to improve the quality of the protoplasts. After that, the constructed NRPS deletion cassettes were transformed into P. brevicompactum protoplasts. Following the hygromycin selection and PCR verification, corrected transformants were created (Fig. S2†). HPLC analysis showed that 1 was completely abolished in both NRPS1.1 (renamed ApmA) and NRPS1.2 (renamed ApmB) deletion mutants. HPLC analysis also showed the essential roles of ApmA (KX443597) and ApmB (KX443596) in the biosynthesis of 1 (Fig. 2A). Therefore, these two NRPSs in cluster 1 (named the apm cluster) were shown to be involved in the biosynthesis of 1.

Next, we predicted the apm biosynthetic genes by bioinformatics analysis across the sequenced fungal genomes (JGI, http://genome.jgi.doe.gov). The apm cluster was conserved in the genomes of three fungi: P. brevicompactum, A. terreus, and A. aculeatus. The cluster consisted of two NRPS genes, apmA and apmB, one putative aldolase gene (apmC) and one putative epimerase/dehydratase gene (apmD) (Fig. 2B). The genes apmA-D from P. brevicompactum shared 68–80% sequence identities at the amino acid level with their orthologues from A. terreus; they...
shared 70–76% sequence identities with their orthologues from *A. aculeatus* (Table 1). Scanning the flanking regions of *apm*-D, no biosynthetic genes were found on either side. Furthermore, no homologues were found in the flanking regions of *A. aculeatus* or upstream of *A. terreus*. Though there were three homologues in the down stream of *A. terreus*, they were putative kinesin-like or zinc finger proteins (Table 1). In addition, neither 1 nor its derivative was reported from *A. terreus* or *A. aculeatus*. These results indicated that four genes in the cluster are possibly sufficient for the synthesis of 1. To address the question of whether *apmC* and *apmB* are involved in the biosynthesis of 1, the same strategy was used for gene deletion, and the extracts from the deletion mutants were analysed (Fig. S2†). HPLC analysis showed that 1 still existed in both *apmC* and *apmD* mutants, indicating that they are not essential for the production of 1 (Fig. S3†). Taken together, these results demonstrate that *apmA* and *apmB* would be sufficient for the biosynthesis of 1.

To probe the roles of *ApmA* and *ApmB* in the biosynthesis of 1, we identified the intermediates from the *apm* gene deletion mutants. By analysing ΔapmB strains, we found that a peak (named compound 3) was accumulated in 15.2 min (Fig. 2A, trace (iii)). We hypothesized that 3 is the product of *ApmA*. To identify the structure of 3, it was isolated by a scale-up fermentation of ΔapmB mutant. Its structure was elucidated as *N*-benzoylphenylenalaninol by NMR and MS analysis (see ESI† data), which corresponded well with the previously published data. Feeding 3 into the ΔapmA mutant restored 1 production, which confirmed that 3 was a product of *ApmA* (Fig. 2A, trace iv). It is rare for an NRPS to generate an alcohol. Using antiSMASH analysis, we identified *ApmA* as an unusual NRPS module consisting of A-T-C-A-T-R with a C-terminal reductase (R)-domain. Several studies have demonstrated the function of the R-domain in natural product biosynthesis. For example, Cox and co-workers revealed that the R domain of PKS-NRPS TENS functions as a Dieckmann cyclase in the formation of tenellin. Walsh and Liu demonstrated the role of the R-like domain of PKS-NRPS CpaS for releasing the product in the biosynthesis of cyclopiazonic acid. Recently, the R domain in PKS-NRPS was identified as an aryl-aldehyde generator from an aryl-acid in the formation of meroterpenoid LL-Z1272β and 2,4-dihydroxy 5,6-dimethyl benzaldehyde. Differences of R-domains in NRPSs from bacteria and fungi were demonstrated by a phylogenetic tree analysis (Fig. S4†). Thus, the hypothesis is that the R domain acts in the reductive release of the shunt product 3 in the two-modular NRPS *ApmA*. 4 and benzoic acid (5) are the possible substrates of *ApmA*.

To confirm this, *apmA* was expressed in both *Saccharomyces cerevisiae* and *A. nidulans* systems. HPLC analysis of *ApmA*-expressing *A. nidulans* showed a new peak at 12.2 min. This peak was isolated and the structure was elucidated as 3. The final titer was 20 mg L⁻¹ by the scale-up fermentation using 4 and 5 as substrates (Fig. 3A and ES†). LC-MS analysis of crude samples showed that compounds 3 and 4 are candidates for the biosynthesis of 1.

### Table 1 Asperphenamate (1) biosynthetic genes in *P. brevicompactum* and their homologues in *A. terreus* and *A. aculeatus*

<table>
<thead>
<tr>
<th><em>P. brevicompactum</em></th>
<th><em>A. terreus</em></th>
<th>Coverage/identity</th>
<th><em>A. aculeatus</em></th>
<th>Coverage/identity</th>
<th>Putative function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF_06</td>
<td>ATEG_09 071</td>
<td>—</td>
<td>ORF_07</td>
<td>—</td>
<td>C6 zinc finger protein</td>
</tr>
<tr>
<td>ORF_07</td>
<td>ATEG_09 070</td>
<td>—</td>
<td>ORF_08</td>
<td>—</td>
<td>RTA-like protein</td>
</tr>
<tr>
<td>ORF_08</td>
<td>ATEG_09 069</td>
<td>—</td>
<td>ORF_09</td>
<td>—</td>
<td>HHE domain protein</td>
</tr>
<tr>
<td><em>apmA</em></td>
<td>ATEG_09 068</td>
<td>99/71</td>
<td>ORF_10</td>
<td>99/71</td>
<td>NRPS</td>
</tr>
<tr>
<td><em>apmC</em></td>
<td>ATEG_09 067</td>
<td>100/80</td>
<td>ORF_11</td>
<td>100/75</td>
<td>Phospho-2-dehydro-3-deoxyheptonate aldolase</td>
</tr>
<tr>
<td><em>apmD</em></td>
<td>ATEG_09 065</td>
<td>39/80</td>
<td>ORF_13</td>
<td>96/76</td>
<td>NAD dependent epimerase</td>
</tr>
<tr>
<td><em>apmB</em></td>
<td>ATEG_09 064</td>
<td>96/68</td>
<td>ORF_14</td>
<td>94/70</td>
<td>NRPS</td>
</tr>
<tr>
<td><em>ORF_14</em></td>
<td>ATEG_09 063</td>
<td>—</td>
<td>ORF_15</td>
<td>86/65</td>
<td>Kinesin-like protein klpA</td>
</tr>
<tr>
<td><em>ORF_15</em></td>
<td>ATEG_09 062</td>
<td>—</td>
<td>ORF_16</td>
<td>99/72</td>
<td>Corrin-binding protein</td>
</tr>
<tr>
<td><em>ORF_16</em></td>
<td>ATEG_09 061</td>
<td>—</td>
<td>ORF_17</td>
<td>90/45</td>
<td>Zinc finger protein</td>
</tr>
</tbody>
</table>
extracts from apmA expressed strains in yeast and in A. nidulans revealed its molecular weight as 256.1395 [M + H]+, confirming 3 production (Fig. S5†). Two additional peaks were also observed with masses of 270.2131 [M + H]+ and 254.0875 [M + H]+ by LC-MS analysis. These molecular weights suggested the formations of the acid form of 2 (molecular formula C16H15NO2, calculated molecular weight 269.1052) and the aldehyde form of N-benzoylphenylalanine (6, molecular formula C16H15NO2, calculated molecular weight 253.1103) (Fig. S5†). These data indicate that ApmA catalyses the carrier protein-bound thioester-ester bond formation and release of the shunt product. Then, the terminal R domain of ApmA catalysed the reduction reaction to get the shunt product 3 through the acid form 2 and the aldehyde form 6 (Fig. 5). Subsequently, ApmB activated the same substrates as did ApmA and waited for the substrate, $^{13}$C,$^{15}$N-2 (Fig. S10†) was synthesized chemically and used as a substrate for the feeding experiments. LC-MS analysis of the crude extract clearly demonstrated that $^{13}$C,$^{15}$N-1 was the major product with a determined molecular weight of 509.2282 [M + H]+ (Fig. S9†). To characterise the structure of $^{13}$C,$^{15}$N-1, large scale fermentations were conducted by feeding labelled $^{13}$C,$^{15}$N-2 and 3 as substrates, and the structure was elucidated by NMR analysis (Fig. S12–S13†). This result suggested that ApmB activates and accepts the linear dipeptide 2 as a substrate. Finally, we coexpressed apmA and apmA genes in A. nidulans. A clear final product peak 1 was observed by feeding the substrates of 4 and 5 in the transformant (Fig. 3C).

Because of the unusual function of ApmB, it is logical to ask whether ApmB can use linear dipeptides such as 2 and 3 as substrates. This question was addressed by feeding 2 and 3 into the apmB-expressing A. nidulans strain. 1 was detected by LC-MS analysis (Fig. 4). To exclude the residual 4 in the medium as substrate, $^{13}$C,$^{15}$N-2 (Fig. S10–S11†) was synthesized chemically and used as a substrate for the feeding experiments. LC-MS analysis of the crude extract clearly demonstrated that $^{13}$C,$^{15}$N-1 was the major product with a determined molecular weight of 509.2282 [M + H]+ (Fig. S9†). To characterise the structure of $^{13}$C,$^{15}$N-1, large scale fermentations were conducted by feeding labelled $^{13}$C,$^{15}$N-2 and 3 as substrates, and the structure was elucidated by NMR analysis (Fig. S12–S13†). This result suggested that ApmB activates and accepts the linear dipeptide 2 as a substrate. Finally, we coexpressed apmA and apmA genes in A. nidulans. A clear final product peak 1 was observed by feeding the substrates of 4 and 5 in the transformant (Fig. 3C).

Based on the gene disruption results and feeding studies in heterologous hosts, we proposed a biosynthetic pathway for 1 (Fig. 5). Using 4 and 5 as substrates, ApmA catalysed amide bond formation and tethered the intermediate into the NRPS cluster. Then, the terminal R domain of ApmA catalysed the reduction reaction to get the shunt product 3 through the acid form 2 and the aldehyde form 6 (Fig. 5). Subsequently, ApmB activated the same substrates as did ApmA and waited for the
attack of 3. After the attack of 3 on the peptidyl-S-N-benzoyl-phenylalanine, the inter-molecular ester bond was formed and the final product 1 was released by ApmB (Fig. 5). The whole process was completed by the coordination of two NRPSs, ApmA and ApmB. This is the first case of a simple compound formation by the machinery of two NRPSs shown in nature.

Conclusions

In summary, we identified the biosynthetic gene cluster of asperphenamate (1) and characterised the two NRPSs function using a heterologous expression approach in A. nidulans. We demonstrated that the NRPS enzyme ApmA had an R-domain, which was responsible for the formation of alcohol reduced from the acid through the aldehyde intermediate. Meanwhile, we also elucidated another NRPS enzyme, ApmB, with functions for ester bond formation and the release from the synthetase to form the final product. More interestingly, ApmB was able to accept the linear dipeptide to form the final product 1. Therefore, our findings not only provide new insights into fungal NRPS machinery for the formation of amino acid esters, but also demonstrate a synthetic approach to develop new small molecules catalysed by NRPS.

Materials and methods

The strains, plasmids and primers used in this work are described in the ESI Tables S1, S2 and S3.† The HPLC analyses of the feeding experiment in apmB-expressed A. nidulans are listed in the ESI Fig. S7 and S8.† The LC-MS spectra of feeding experiment in apmB-expressed A. nidulans with 13C,15N-2 and 3 as substrates are listed in the ESI Fig. S9.† The NMR spectra for compound 13C, 15N-1, 13C,15N-2, 3 and 2-SNAC are listed in the ESI Fig. S10–S15†. Detailed descriptions of experimental procedures and reagents are provided in detail in the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Drs Jinwei Ren, Wenzhao Wang and Haining Lyu (Institute of Microbiology, CAS) for NMR and MS data collection. We thank Dr Xuebing Li for generously providing N-(3-dimethylaminopropyl)-N0-ethylcarbodiimide hydrochloride and 1-hydroxysuccinimide. We thank Dr Georgina Salazar at the University of Texas Health Science Center at Houston for her careful and critical reading of the manuscript. This work was supported by the National Natural Science Foundation of China (31470178 to W. B. Y. and 81502964 to Z. L.) and the State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College (Grant GTZK201506), and the Welch Foundation (Grant AU00024 to Z. A.). W. B. Y. is a scholar of “the 100 Talents Project” of CAS.

References

3 V. Miao, M. F. Coeffet-Legal, P. Brian, R. Brost, J. Penn, A. Whiting, S. Martin, R. Ford, I. Parr, M. Bouchard,