EDGE ARTICLE

Comparative genome mining and heterologous expression of an orphan NRPS gene cluster direct the production of ashimides†

Jing Shi,[a] Ying Jie Zeng,[a] Bo Zhang,[a] Fen Li Shao,[a] Yan Chi Chen,[a] Xiang Xu,[a] Yang Sun,[a] Qiang Xu,[a] Ren Xiang Tan [ID] *[ab] and Hui Ming Ge [ID] *[a]

The ever-increasing bacterial genomic repositories reveal a great number of uncharacterized biosynthetic gene clusters, representing a tremendous resource for natural product discovery. Genome mining of the marine *Streptomyces* sp. NA03103 indicates the presence of an orphan nonribosomal peptide synthetase (NRPS) gene cluster (*asm*), to which there are no homologous gene clusters in the public genome databases. Heterologous expression of the *asm* gene cluster in the *S. lividans* SBT18 strain led to the discovery of two novel cyclopeptides, ashimides A and B (**1** and **2**), with **2** showing cytotoxic activity. In addition, we use bioinformatic analysis, gene inactivation and stable isotope labelling experiments, as well as *in vitro* biochemical assays, to present a coherent and novel assembly line for ashimide biosynthesis, featuring an unusual desaturation, halogenation and cyclization cascade catalyzed by a P450 monooxygenase and a FAD-dependent halogenase.

Nonribosomal peptides (NRPs) produced by actinobacteria constitute a large group of secondary metabolites with potent biological activities as exemplified by vancomycin, dactinomycin, and daptomycin. Nature developed modular biosynthetic machinery to synthesize NRPs, in which each module in the nonribosomal peptide synthetase (NRPS) assembly line activates and incorporates an amino acid into a growing peptide.[1–5] According to this rule, the amino acid building block composition of NRPs can correlate well with their corresponding NRPS architectures.[6,7] With the massive accumulation of bacterial genome data, it has become obvious that the biosynthetic potential of NRPs greatly exceeds the number of compounds that have been isolated.[8,9] Orphan gene clusters, which are gene loci to which no known homologous gene clusters can be found, represent a particularly rich and untapped source for novel natural product discovery. Genome mining targeting the orphan NRPS gene cluster is therefore a powerful approach to discover novel NRPs.

During our genome survey of sequenced strains, an orphan NRPS gene cluster (*asm*) was identified from a marine *Streptomyces* sp. NA03103. The *asm* gene cluster is unusual for three reasons: (i) no homologous gene clusters can be found in the public genome databases through ClusterBlast, a dereplication module in online software antiSMASH;[10] (ii) the upstream and downstream genes flanking the putative *asm* gene cluster are highly conserved (>93% identity) in a model *Streptomyces* strain *S. lividans* TK24 (Scheme 1); (iii) all four NRPSs (AsmC, AsmM, AsmN and AsmS) are small monomodule NRPSs, which are in sharp contrast to the canonical multimodule NRPS megaenzymes in bacterial gene clusters.[11] The putative *asm* gene cluster embedded in the conserved region is proposed to only span ~32 kb and consists of 19 genes (*asmA–asmS*) (Scheme 1 and Table S3†). In addition, several notable tailoring enzymes are also present including an FAD-dependent halogenase (*asmO*), two P450 monooxygenases (*asmI* and *asmK*), and a methyltransferase gene (*asmP*).

With the gene cluster in hand, we first aimed to identify its encoded natural products in strain NA03103. Somewhat to our surprise, no corresponding compounds could be isolated from metabolic extracts of this wild-type strain, even when cultured under a wide variety of growth conditions. Putatively, this gene cluster is silent in this strain under current laboratory conditions. To bypass the endogenous regulatory control and activate the cryptic pathway, we sought to clone the intact *asm* gene cluster into a heterologous host. Thus, we constructed a cosmid library of *S.* sp. NA03103 using a *Streptomyces* integrative vector pJTU2554 according to a standard procedure.[12] One cosmid, pHG4001 (Fig. S1†), containing the full length of the *asm* gene cluster was obtained and transferred into *S. lividans* SBT18 for heterologous expression. Metabolic extracts from the resulting *S.*

aState Key Laboratory of Pharmaceutical Biotechnology, Institute of Functional Biomolecules, School of Life Sciences, Nanjing University, 210023, China. E-mail: rxtan@nju.edu.cn; hmge@nju.edu.cn

bState Key Laboratory Cultivation Base for TCM Quality and Efficacy, Nanjing University of Chinese Medicine, Nanjing 210023, China
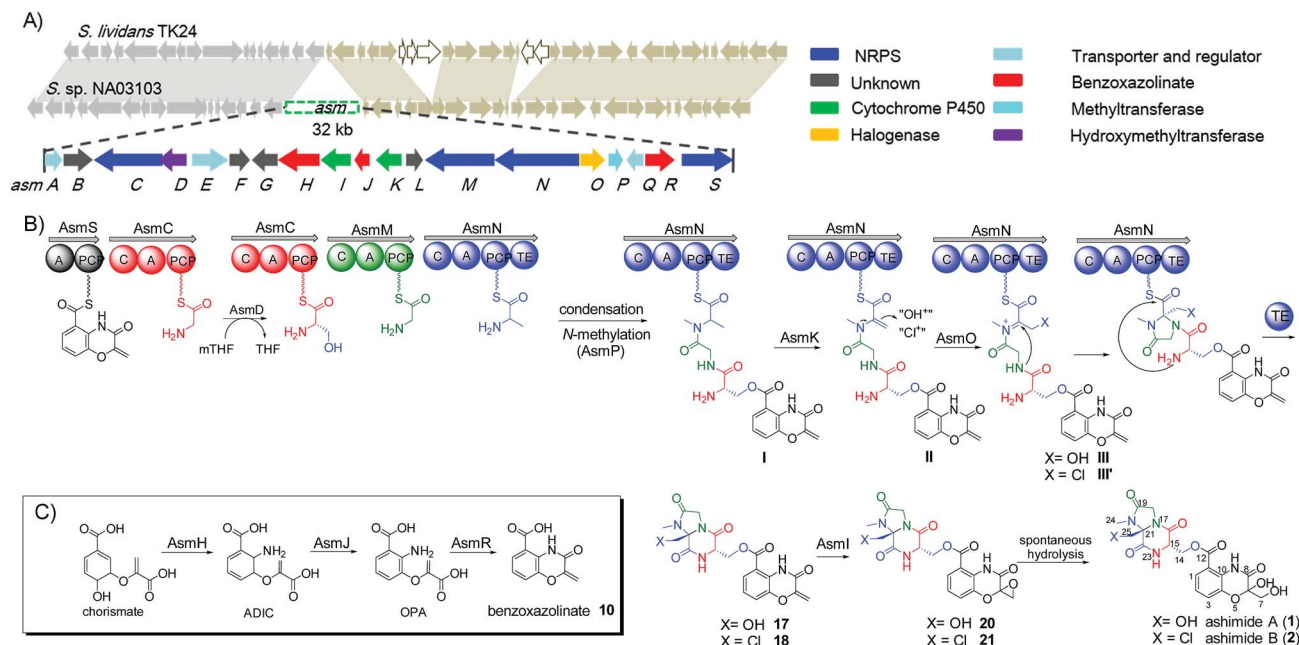
Scheme 1 Biosynthesis of ashimides. (A) The *asm* gene cluster and schematic comparison of *asm* neighbouring genes in *S.* sp. NA03103 with their homologs in *S. lividans* TK24; (B) proposed biosynthetic pathway of ashimides A and B; (C) proposed biosynthetic pathway of the benzoxazolinate moiety in ashimides. mTHF, 5,10-methylenetetrahydrofolate; THF, tetrahydrofolate; ADIC, 2-amino-2-desoxyisochorismate; OPA, 3-enolpyruvoylanthranilate.

lividans SBT18/pHG4001 strain were obtained and analyzed by HPLC and LC/MS. In comparison to the *S. lividans* harboring an empty cosmid vector, *S. lividans* SBT18/pHG4001 produced two major peaks (**1** and **2**) with molecular weights of 464 and 482 (Fig. 1), respectively. It is of note that **2** showed a characteristic ratio of monochlorinated isotopic signals ($[M + Na]^+ : [M + Na + 2]^+ = 3 : 1$) (Fig. S2†), in line with the presence of a putative FAD-dependent chlorinase (AsmO) encoded in the gene cluster. Hence, we carried out a large-scale (20 L) fermentation from which we isolated compounds **1–9** (Fig. S3†), the latter seven of which were barely observed in the HPLC profile.

Compound **1** (1.0 mg L^{-1}) was obtained as a pale yellow crystal with a molecular formula of $C_{19}H_{20}N_4O_{10}$ as determined by its HRESIMS and ^{13}C NMR data (Table S4†). The planar structure, along with the absolute configurations, was unambiguously determined by analyzing its NMR and single-crystal X-ray diffraction data with a flack parameter value of $-0.03(7)$ collected using Cu Kα at low temperature (Fig. S4†). Compound **2** (1.3 mg L^{-1}) has a molecular formula of $C_{19}H_{19}ClN_4O_9$. The ^1H and ^{13}C NMR spectra highly resembled those of **1** except for the absence of the proton signal for 25-OH and a chemical shift value for C-25 (δ_C 47.3 for **2** and 61.4 for **1**), which indicated that the OH group in C-25 was substituted by a chlorine atom. Further interpretation of 1D, 2D NMR and CD data established the structure of **2** (Table S5, Fig. S5†). Structures of **3–9** were characterized through extensive analyses of HRMS and NMR data (Tables S6–S12†). To assess the biological activity, we then evaluated the cytotoxicity of **1–9** against MCF-7 cell lines using the MTT assay. Only **2** exhibited cytotoxicity with an IC$_{50}$ value of 19.8 μM, whereas others are inactive. The flow cytometry
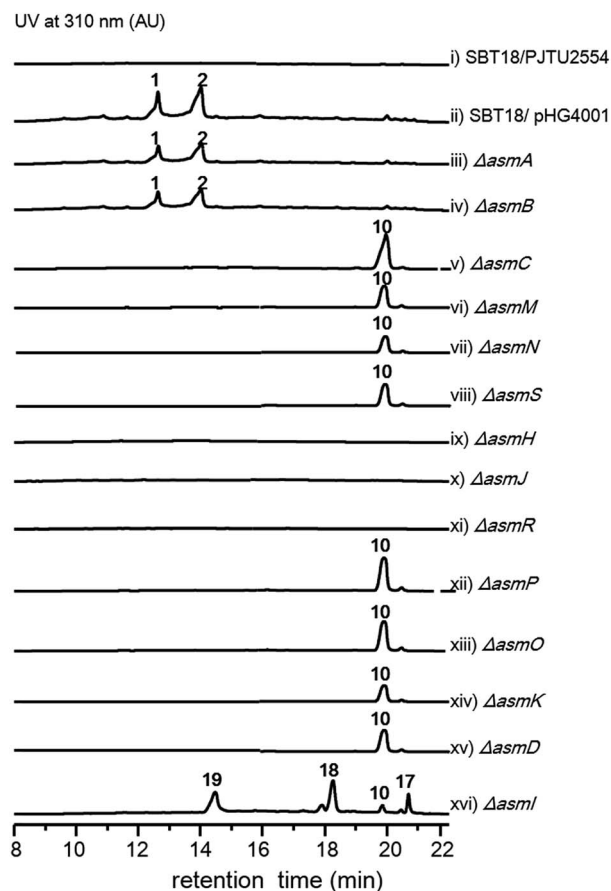

Fig. 1 HPLC analysis of metabolite extracts from mutant strains.

analysis showed that **2** could induce apoptosis in a dose-dependent manner in MCF-7 cells (Fig. S6†).

The two most abundant compounds, ashimides A and B (**1** and **2**), contain two rare structural features including the 5/6 bicyclic imidazo[1,2-a]diketopiperazine and benzoxazolinate moieties, the former of which has never been observed among natural products, hinting that a novel cyclization strategy may be involved in the biosynthesis of ashimides. Additionally, the benzoxazolinate moiety, which plays a crucial role in intercalating double-stranded DNA,[13] has only been reported in enediyne antitumor agents C-1027 and auromomycin.[14,15] These unusual structural features encouraged us to investigate the biosynthetic machinery of ashimides in depth. For this purpose, we systematically inactivated genes in the *asm* cluster through in-frame deletion (Fig. S7†), which was achieved through homologous recombination between the cosmid pHG4001 and a shuttle vector carrying the mutation (see the ESI†).[16] The resultant mutant strains were fermented, with the *S. lividans* SBT18/pHG4001 strain as a control, to investigate the effect of these genes on **1** and **2** biosynthesis.

Deletion of two boundary genes *asmA* and *asmB*, which are annotated to be a response regulator and a histidine kinase, respectively, did not affect the production of **1** and **2** (Fig. 1 and S8†). Inactivation of NRPS genes, *asmC*, *asmM*, *asmN* or *asmS*, as expected no longer produced **1** and **2** but instead clearly accumulated a new compound **10** (Fig. 1 and S8†). Isolation of **10** from the *ΔasmC* strain followed by NMR characterization (Table S13†) indicated that **10** is the benzoxazolinate compound (Scheme 1C). We proposed that the formation of **10** takes place prior to adenylation and the loading module AsmS may directly accept **10** as a substrate.

AsmH shares high sequence homology (52% identity) with PhzE, a biochemically characterized 2-amino-2-desoxyisochorismate (ADIC) synthase in the first step of phenazine biosynthesis,[17] whereas AsmJ and AsmR are homologous to SgcG (48% identity) and SgcD5 (48% identity),[18] respectively, two tandem enzymes involved in the conversion of ADIC to **10** in the biosynthesis of a benzoxazolinate motif of C-1027.[14] Therefore, these three enzymes are most likely sufficient for the biosynthesis of **10** (Scheme 1). Knocking out *asmH*, *asmJ* or *asmR* abolished the production of **1** and **2**, but did not accumulate any observed peaks (Fig. 1 and S8†). The production of **1** and **2** can be restored when feeding **10** into *ΔasmH*, *ΔasmJ* or *ΔasmR* mutants individually (Fig. S9†), confirming their involvement in the biosynthesis of **10** and also indicating that **10** is an on-pathway intermediate.

Deletion of the *asmP* gene, which encoded a methyltransferase (Table S3†), resulted in the complete abolition of **1** and **2** with only the accumulation of **10** (Fig. 1 and S8†). This indicates that *N*-methylation is essential for ashimide biosynthesis. In addition, supplementation of the putative *N*-methylated amino acid precursors including 3-chloro-*N*-methylalanine (**11**), *N*-methylserine (**12**), or *N*-methylalanine (**13**) into *ΔasmP* could not restore the production of **1** or **2** (Fig. S9†), suggesting that the methylation step most likely takes place after amino acid binding to the NRPS.[19,20] A similar scenario has also been observed in α-lipomycin biosynthesis.[21]

The substituted groups at C-25 of **1** and **2** are hydroxyl and chlorine, respectively, raising the question of their origin. Knocking out the FAD-dependent halogenase gene, *asmO*, eliminated the production of **1** and **2**, and accumulated **10**, excluding the possibility that **1** is the precursor of **2**, and *vice versa*. As the FAD-dependent halogenase/hydroxylase requires electron rich alkenes (such as phenols, pyrroles, or indoles) for halogenation/hydroxylation,[22] a putative alkene-containing intermediate is proposed to form during biosynthesis. In accordance with this hypothesis, a P450 monooxygenase AsmK, which shows moderate sequence homology (34% identity) to LtmK, a desaturase in lactimidomycin biosynthesis,[23] is encoded in the *asm* gene cluster. The mutant strain *ΔasmK* only accumulated **10** (Fig. 1 and S8†). Feeding of **11**, **13** and 3-chloroalanine (**14**) to *ΔasmO* cannot restore the production of **1** or **2** (Fig. S9†). Therefore, we reasoned that the AsmN may activate and accept alanine as a substrate instead of serine. After elongation, the alanine building block is desaturated by AsmK to form a 2-aminoacrylic acid containing intermediate. AsmO may play a bifunctional role in introducing either a hydroxyl or a chlorine group at the newly generated alkene position (Scheme 1), as the FAD-dependent halogenases can be viewed as variants of FAD-dependent hydroxylases.[22]

However, the A domain substrate specificity of AsmN predicted by NRPSpredictor2 is cysteine instead of alanine as hypothesized above (Table S20†).[24] In addition, A domains in AsmC, AsmM, and AsmS were predicted to incorporate cysteine, glycine, and glycine/alanine/valine, respectively (Table S20†),[24] which is not in full agreement with amino acid composition in ashimides. To determine the substrate specificity of the A domains, *asmC*, *asmM*, *asmN*, and *asmS* genes were individually overexpressed from *Escherichia coli* BL21(DE3), and the full-length NRPS proteins were purified (Fig. S10†). The adenylation activity towards various amino acid substrates was evaluated by monitoring pyrophosphate production through a fluorogenic pyrophosphate sensor.[25] As shown in Fig. 2, AsmS specifically accepts benzoxazolinate (**10**) as the substrate, which is
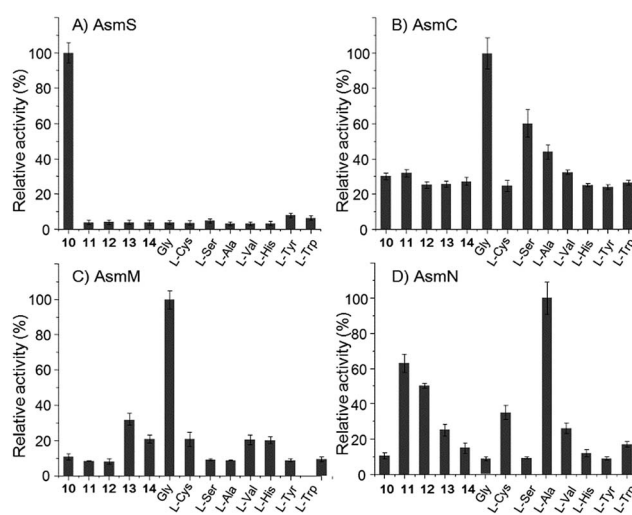

Fig. 2 The relative adenylation activities for (A) AsmS, (B) AsmC, (C) AsmM, and (D) AsmN. Error bars represent SEM (*n* = 3).

consistent with the structures of **1** and **2**, as **10** should be the first amino acid to be incorporated. Interestingly, both AsmC and AsmM showed clear preferences for glycine, although substrate specificity of AsmC was relatively broad. The last module, AsmN, prefers L-alanine as the substrate over 3-chloro-*N*-methylalanine (**11**), *N*-methylserine (**12**), *N*-methylalanine (**13**), 3-chloroalanine (**14**) or cysteine. These findings, in conjunction with the results of the *in vivo* knocking out

experiments, strongly suggested that benzoxazolinate, glycine, glycine and L-alanine are the substrates of AsmS, AsmC, AsmM and AsmN, respectively, in ashimide biosynthesis.

To further prove this, we next probed the origin of amino acid building blocks by feeding a stable isotope labeled amino acid to the culture. In contrast to the unlabeled ion (Fig. 3A(i)), LC-MS analysis of ^{15}N-labelled alanine clearly showed the increment of the ion peak at m/z 484.1 (**2**, $[M + H]^+$) (Fig. 3B(i)),
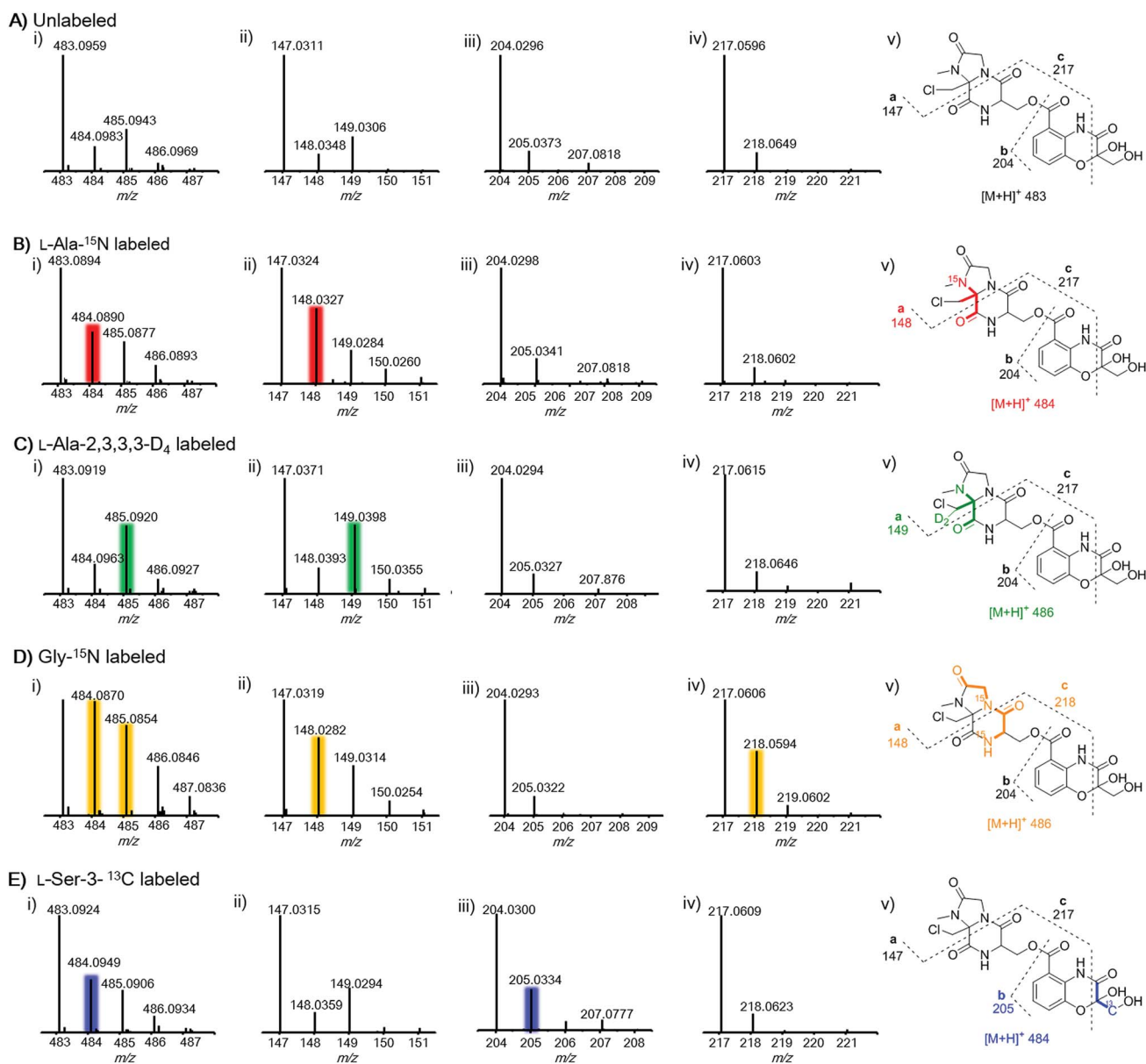

Fig. 3 MS and MS^2 fragmentation analysis of **2** when targeting the unlabeled (panel A) and after feeding with labeled amino acids including L-alanine-^{15}N (panel B), L-alanine-2,3,3,3-D$_4$ (panel C), glycine-^{15}N (panel D), and L-serine-3-^{13}C (panel E). Panel (A): parent MS signal for unlabeled **2** (i, m/z 483, $[M + H]^+$); MS^2 ion signals for fragments of a (ii, m/z 147), b (iii, m/z 204), and c (iv, m/z 217); (v) MS^2 fragmentation ion pattern for unlabeled **2**. Panel (B): parent MS signal for **2** after feeding L-alanine-^{15}N (i, m/z 484, $[M + H]^+$); MS^2 ion signals for fragments of a (ii, m/z 148), b (iii, m/z 204), and c (iv, m/z 217); (v) MS^2 fragmentation ion pattern for **2** after feeding L-alanine-^{15}N. Panel (C): parent MS signal for **2** after feeding L-alanine-2,3,3,3-D$_4$ (i, m/z 484, $[M + H]^+$); MS^2 ion signals for fragments of a (ii, m/z 148), b (iii, m/z 204), and c (iv, m/z 217); (v) MS^2 fragmentation ion pattern for **2** after feeding L-alanine-2,3,3,3-D$_4$. Panel (D): parent MS signal for **2** after feeding glycine-^{15}N (i, m/z 485, $[M + H]^+$); MS^2 ion signals for fragments of a (ii, m/z 148), b (iii, m/z 204), and c (iv, m/z 218); (v) MS^2 fragmentation ion pattern for **2** after feeding glycine-^{5}N. Panel (E): parent MS signal for **2** after feeding L-serine-3-^{13}C (i, m/z 484, $[M + H]^+$); MS^2 ion signals for fragments of a (ii, m/z 147), b (iii, m/z 205), and c (iv, m/z 217); (v) MS^2 fragmentation ion pattern for **2** after feeding L-serine-3-^{13}C.
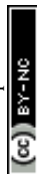
indicating that one alanine was incorporated into **2**. To ascertain which N atom was labeled by ^{15}N, 2 L scale fermentation was conducted with feeding labeled ^{15}N-alanine, leading to the purification of 0.2 mg of **2**. The ^{15}N–^{1}H HMBC spectrum clearly indicated that H-24 and H-25 showed correlations with N-20 (Fig. S11†), revealing that the 3-chloroalanine building block was indeed derived from alanine, which was reinforced by the observed MS/MS fragmentation of 148.0 (Fig. 3B(ii)). Supplementation of alanine-D$_4$ into the culture led to +2 Dalton mass shifts in **2** (Fig. 3C(i)) and a chlorine-containing fragment (*m/z* 149.0) (Fig. 3C(ii)), further supporting the fact that 3-chloroalanine originated from alanine. When the culture was supplemented with ^{15}N-glycine, a +2 Dalton mass shift was observed (Fig. 3D(i)), indicating that two molecules of glycine were incorporated into **2**. The relative abundance of ion peaks at *m/z* 148.0 (Fig. 3D(ii)) and 218.1 (Fig. 3D(iv)) in MS/MS analysis revealed that two glycines were incorporated in the second and third NRPS modules, respectively. To further verify if serine can be incorporated into **2**, ^{13}C$_3$-serine was added to the culture. The MS and MS/MS ion peaks with a +1 Dalton mass shift were observed only for the benzoxazolinate-containing part (Fig. 3E(iii)). This can be explained by the fact that the three-carbon unit in the benzoxazolinate moiety is derived directly from phosphoenolpyruvic acid (PEP), which originates from serine through the pyruvate intermediate (Fig. S12†).[26]

Thus, both the A domain substrate specificity assay and stable isotope labeling experiment clearly indicate that AsmC accepts glycine as the substrate, although ashimide requires a serine building block indeed. Interestingly, it appears that *asmC* and *asmD* are located in the same operon as the start codon of *asmD* overlaps with the stop codon of *asmC* (Scheme 1). AsmD is predicted to be a serine hydroxymethyltransferase that is a pyridoxal-5′-phosphate (PLP) dependent enzyme catalyzing the conversion of glycine to serine by transferring a hydroxymethyl group from 5,10-methylenetetrahydrofolate (mTHF) in primary metabolism.[27,28] The *ΔasmD* mutant abolished the production of **1** and **2** but only accumulated **10** (Fig. 1 and S8†), confirming the necessity of AsmD in ashimide biosynthesis. The co-translation of the genes in the same operon would yield equal amounts of enzymes for their interactions.[29] Thus, AsmD could recognize AsmC through a protein–protein interaction and converts AsmC bound glycine to serine. To examine the role of AsmD, we obtained the recombinant AsmD enzyme from *E. coli* (Fig. S10†) and chemically synthesized *N*-acetylcysteamine (SNAC) thioesters with L-serine (**15**) and glycine (**16**) as mimics of the nascent PCP-tethered amino acids. As mTHF is not commercially available, we carried out the AsmD-catalyzed reverse reaction.[28] AsmD was incubated with **15** in the presence of PLP and tetrahydrofolate (THF). Gratifyingly, a new peak with the expected molecular weight of **16** was observed (Fig. 4); in contrast, no reactions occurred when using boiled AsmD as the catalyst. These findings clearly indicate that AsmC activates glycine and recruits AsmD for hydroxymethylation to yield the required serine building block.

Finally, to investigate the function of the only remaining cytochrome P450 monooxygenase AsmI, we generated the *ΔasmI* strain. HPLC analysis showed that both final products **1** and **2** were eliminated, leaving **10** and **17–19** as the dominant
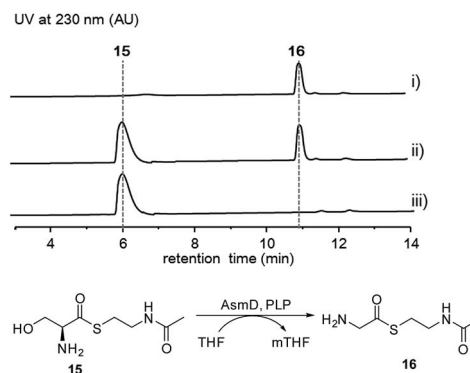

Fig. 4 HPLC analysis of the AsmD-catalyzed biochemical reaction *in vitro*. (i) Standard of **16**, (ii) **15** incubated with AsmD in the presence of PLP and tetrahydrofolate, and (iii) L-serine SNAC (**15**) incubated with boiled AsmD, PLP and tetrahydrofolate.

products (Scheme 1, Fig. 1 and S8, and Tables S17–S19†). Supplementation of **17** and **18** to the *ΔasmS* mutant was sufficient to restore the production of **1** and **2**, respectively, indicating that both compounds are on-pathway biosynthetic intermediates (Fig. S9†). However, **19** is a shunt metabolite or a hydrolyzed artefact as feeding of **19** to *ΔasmS* could not restore the production of **2** (Fig. S9†). Hence, AsmI is directly involved in the oxidation of the alkene group of benzoxazolinate. However, whether the diol group is directly generated by AsmI or through an epoxide intermediate by spontaneous hydrolysis remains uncertain. To examine this, AsmI was overproduced and purified to homogeneity with an N-terminal his$_6$-tag from *E. coli* BL21(DE3) (Fig. S10†). The native ferredoxin reductase (FDR) and the ferredoxin (FDX) partner cloned from the same strain *S.* sp. NA03103 were also overproduced and purified (Fig. S10†). Compound **17** was thus incubated with AsmI in the presence of FDR, FDX and electron donor NADPH for 1 h (Fig. 5A). The production of **1** and a putative epoxide compound **20** with an observed $[M + H]^+$ ion at *m/z* 447.1145 in LC-MS was observed (Scheme 1, Fig. S13†). Attempts to purify **20** were not successful as it rapidly converted to **1**. To further understand this conversion, we carried out a time-course analysis for the AsmI-catalyzed reaction (Fig. 5A). The production of **20** initially increased but was finally converted to **1**, suggesting that the epoxide intermediate **20** could be the real natural product, which can be spontaneously hydrolyzed to **1**. In addition, AsmI can accept **18** to produce **2** through a putative epoxide intermediate **21** as indicated by its HRESIMS (Scheme 1, Fig. 5B and S14†). Kinetic analysis showed that AsmI displayed similar catalytic activities toward both substrates with K_m 2.224 ± 0.2718 mM and k_{cat} 0.166 ± 0.053 min^{-1} for **17**; and K_m 2.035 ± 0.2937 mM and k_{cat} 0.171 ± 0.058 min^{-1} for **18** (Fig. S15 and S16†). In addition, AsmI has no effect on **10** (Fig. S17†).

Taken together, our data revealed that four monomodule NRPSs successively incorporate benzoxazolinate, glycine, glycine and alanine into the assembly line. After conversion of PCP-tethered glycine to serine at the second module by AsmD, the tetrapeptide framework is assembled based on the canonical NRPS rule, followed by *N*-methylation catalyzed by AsmP.
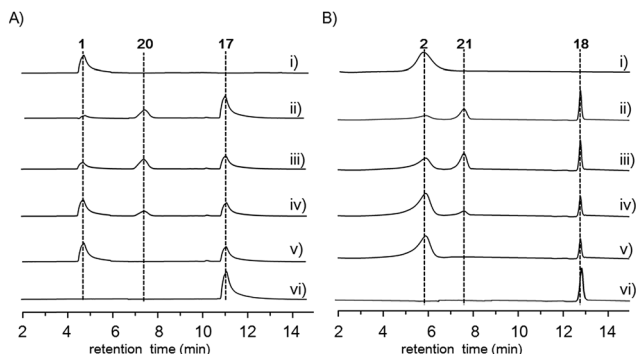
A)

B)

Fig. 5 HPLC analysis of the AsmI-catalyzed reaction. (A) *In vitro* assays of AsmI using **17** as the substrate. (i) Standard of **1**; **17** was incubated with AsmI, FDR, FDX, and NADPH for (ii) 1 h; (iii) 1.5 h; (iv) 2 h; (v) 3 h; (vi) **17** was incubated with boiled AsmI, FDR, FDX and NADPH for 2 h. (B) *In vitro* assays of AsmI using **18** as the substrate. (i) Standard of **2**; **18** was incubated with AsmI, FDR, FDX, and NADPH for (ii) 1 h; (iii) 1.5 h; (iv) 2 h; (v) 3 h; (vi) **18** was incubated with boiled AsmI, FDR, FDX and NADPH for 2 h.

The cytochrome P450 AsmK is expected to desaturate the C21–C25 position to afford 2-aminoacrylic moiety intermediate **II**,[30] which further undergoes hydroxylation/halogenation (both *via* Fl-4a-OOH as the electrophilic donor) to afford iminium ion intermediates **III** or **III'** (Scheme 1). The transient iminium ion can then be intramolecularly captured by the amine –NH to yield imidazolidin-4-one ring.[31,32] The TE domain catalyzes the offloading of the tetrapeptide from AsmN by forming an amide bond to produce **17** and **18**, which were finally epoxidized by AsmI to afford the expected final natural products **20** and **21** (Scheme 1). However, both **20** and **21** are prone to be spontaneously hydrolyzed to the isolatable **1** and **2**, respectively.

Conclusions

In conclusion, in spite of the great biosynthetic potential of natural products revealed by recent advances in microbial genomics, dereplication and activation gene clusters remain a key issue in obtaining novel secondary metabolites. Our work presented here underpins the importance of combining bioinformatics analysis and heterologous expression of the orphan and neglected gene cluster for novel natural product discovery. The biosynthetic pathway for ashimides A and B isolated in this work was established through gene inactivation, biochemical analysis and stable isotope-labelling experiments, revealing a new desaturation, halogenation/hydroxylation and cyclization cascade to generate structural complexity.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 M. A. Marahiel, *Nat. Prod. Rep.*, 2016, **33**, 136–140.
2 C. T. Walsh, *Nat. Prod. Rep.*, 2016, **33**, 127–135.
3 K. Bloudoff, C. D. Fage, M. A. Marahiel and T. M. Schmeing, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 95–100.
4 F. Kopp and M. A. Marahiel, *Nat. Prod. Rep.*, 2007, **24**, 735–749.
5 R. D. Süssmuth and A. Mainz, *Angew. Chem., Int. Ed.*, 2017, **56**, 3770–3821.
6 M. Strieker, A. Tanovic and M. A. Marahiel, *Curr. Opin. Struct. Biol.*, 2010, **20**, 234–240.
7 C. T. Walsh, *Science*, 2004, **303**, 1805–1810.
8 J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H. C. Tai, U. I. Zakai and D. A. Mitchell, *Nat. Chem. Biol.*, 2017, **13**, 470–478.
9 X. Zhang, T. T. Wang, Q. L. Xu, Y. Xiong, L. Zhang, H. Han, K. Xu, W. J. Guo, Q. Xu, R. X. Tan and H. M. Ge, *Angew. Chem., Int. Ed.*, 2018, **57**, 8184–8188.
10 T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano and M. H. Medema, *Nucleic Acids Res.*, 2015, **43**, 237–243.
11 M. A. Fischbach and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3468–3496.
12 S. Huang, Y. Zhao, Z. Qin, X. Wang, M. Onega, L. Chen, J. He, Y. Yu and H. Deng, *Process Biochem.*, 2011, **46**, 811–816.
13 L. Yu, S. Mah, T. Otani and P. Dedon, *J. Am. Chem. Soc.*, 1995, **117**, 8877–8878.
14 W. Liu, S. D. Christenson, S. Standage and B. Shen, *Science*, 2002, **297**, 1170–1173.
15 Y. Kumada, T. Miwa, N. Naoi, K. Watanabe, H. Naganawa, T. Takita, H. Umezawa, H. Nakamura and Y. Iitaka, *J. Antibiot.*, 1983, **36**, 200–202.
16 M. D. Lalioti and J. K. Heath, *Nucleic Acids Res.*, 2001, **29**, 14–19.
17 Q. A. Li, D. V. Mavrodi, L. S. Thomashow, M. Roessle and W. Blankenfeldt, *J. Biol. Chem.*, 2011, **286**, 18213–18221.
18 S. G. Van Lanen, S. Lin and B. Shen, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 494–499.
19 J. Liu, B. Wang, H. Li, Y. Xie, Q. li, X. Qin, X. Zhang and J. Ju, *Org. Lett.*, 2015, **17**, 1509–1512.
20 K. M. J. de Mattos-Shipley, C. Greco, D. M. Heard, G. Hough, N. P. Mulholland, J. L. Vincent, J. Micklefield, T. J. Simpson, C. L. Willis, R. J. Cox and A. M. Bailey, *Chem. Sci.*, 2018, **9**, 4109–4117.
21 C. Bihlmaier, E. Welle, C. Hofmann, K. Welzel, A. Vente, E. Breitling, M. Müller, S. Glaser and A. Bechthold, *Antimicrob. Agents Chemother.*, 2006, **50**, 2113–2121.
22 C. T. Walsh and T. A. Wencewicz, *Nat. Prod. Rep.*, 2013, **30**, 175–200.

23 J. W. Seo, M. Ma, T. Kwong, J. Ju, S. K. Lim, H. Jiang, J. R. Lohman, C. Yang, J. Cleveland, E. Zazopoulos, C. M. Farnet and B. Shen, *Biochemistry*, 2014, **53**, 7854–7865.

24 M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher, *Nucleic Acids Res.*, 2011, **39**, 362–367.

25 J. Kottur and D. T. Nair, *Nucleic Acids Res.*, 2018, **46**, 5875–5885.

26 P. J. Large, D. Peel and J. R. Quayle, *Biochem. J.*, 1962, **82**, 483–488.

27 M. Vivoli, F. Angelucci, A. Ilari, V. Morea, S. Angelaccio, M. L. di Salvo and R. Contestabile, *Biochemistry*, 2009, **48**, 12034–12046.

28 A. Muliandi, Y. Katsuyama, K. Sone, M. Izumikawa, T. Moriya, J. Hashimoto, I. Kozone, M. Takagi, K. Shin-ya and Y. Ohnishi, *Chem. Biol.*, 2014, **21**, 923–934.

29 S. Alt and B. Wilkinson, *ACS Chem. Biol.*, 2015, **10**, 2468–2479.

30 C. T. Walsh, S. W. Haynes, B. D. Ames, X. Gao and Y. Tang, *ACS Chem. Biol.*, 2013, **8**, 1366–1382.

31 S. W. Haynes, X. Gao, Y. Tang and C. T. Walsh, *J. Am. Chem. Soc.*, 2012, **134**, 17444–17447.

32 S. W. Haynes, X. Gao, Y. Tang and C. T. Walsh, *ACS Chem. Biol.*, 2013, **8**, 741–748.