Original Article

Gene mining and efficient biosynthesis of a fungal peptidyl alkaloid

Chen Zhang\(^1\), Bing-yu Liu\(^2\), Jia-wang Liu\(^3\), Dao-jiang Yan\(^3\), Jian Bai\(^3\), Ya-long Zhang\(^3\), Yan-hua Mou\(^4\), You-cai Hu\(^5\,\ast\)

\(^1\) State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China
\(^2\) School of Life Science and Biopharmaceutics, Shenyang Pharmaceutical University, Shenyang 110016, China

A R T I C L E   I N F O

Article info
Received 16 May 2020
Revised 14 July 2020
Accepted 1 August 2020
Available online 1 December 2020

Keywords:
anacine
biosynthesis
non-ribosomal peptide synthetase
peptidyl alkaloid

A B S T R A C T

Objective: Peptidyl alkaloids, a series of important natural products can be assembled by fungal non-ribosomal peptide synthetases (NRPSs). However, many of the NRPSs associated gene clusters are silent under laboratory conditions, and the traditional chemical separation yields are low. In this study, we aim to discovery and efficiently prepare fungal peptidyl alkaloids assembled by fungal NRPSs.

Methods: Bioinformatics analysis of gene cluster containing NRPSs from the genome of Penicillium thymicola, and heterologous expression of the putative gene cluster in Aspergillus nidulans were performed. Isolation, structural identification, and biological evaluation of the product from heterologous expression were carried out.

Results: The putative tri-modular NRPS AncA was heterologously-expressed in A. nidulans to give anacine (1) with high yield, which showed moderate and selective cytotoxic activity against A549 cell line.

Conclusion: Heterologous expression in A. nidulans is an efficient strategy for mining fungal peptidyl alkaloids.

© 2020 Tianjin Press of Chinese Herbal Medicines. Published by ELSEVIER B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Natural products play a pivotal role in drug discovery and fungi have been proved to be an important resource of bioactive natural products. Peptidyl alkaloids are a significant class of bioactive secondary metabolites with characteristic multicyclic, constrained rings in their chemical structures, and mainly produced in filamentous fungi (Walsh, Haynes, Ames, Gao, & Tang, 2013). Typical examples of peptidyl alkaloids included asperlicins as cholycystokin antagonist (Liesch, Hensens, Zink, & Goetz, 1988), ardeemins with function of obstruction the multiple drug resistance in tokinin antagonist (Liesch, Hensens, Zink, & Goetz, 1988), examples of peptidyl alkaloids included asperlicins as cholycystokin antagonist (Liesch, Hensens, Zink, & Goetz, 1988), fumiquinazolines with diverse bioactivities (Takahashi et al., 1995), as well as fumiquinazolines with diverse bioactivities (Takahashi et al., 1995).

Biiosynthetically, the core scaffold of peptidyl alkaloids is assembled by short non-ribosomal peptide synthetases (NRPSs), and then further modified via a series of tailoring enzymes to increase structural complexity and diversity, such as oxidation, acylation and pentenylation (Haynes, Ames, Gao, Tang, & Walsh, 2011; Yin, Grundmann, Cheng, & Li, 2009). It was reported that NRPSs have ability to select, activate, and incorporate the nonproteinogenic amino acid anthranilate (Ant) as building block in peptidyl alkaloids biosynthesis, exampled by fumiquinazolines (Gao et al., 2012; Haynes, Gao, Tang, & Walsh, 2013). Therein, fumiquinazolines F (FQF), containing a 6–6–6 tricyclic core, has the simplest modification in this family. Biochemical characterization has confirmed that FQF is a tripeptidyl skeleton assembled by anthranilic acid (Ant), L-tryptophan (L-Trp), and L-alanine (L-Ala) (Ames, Liu, & Walsh, 2010). As one of the most structurally complex quinazoline alkaloids, alanditrypinone was characterized to be biosynthesized by only three enzymes in the Aldp cluster, where a tri-modular NRPS, AldpA, catalyzed the generation of a tripeptide 14-epi-FQF, then an \( \alpha \)-KG-dependent oxygenase AldpC led to the formation of an iminium containing intermediate, and a mono-modular NRPS, AldpB, loaded an additional L-Trp moiety to construct the final structure (Yan et al., 2019).

In order to discover more peptidyl alkaloids from fungi, we commenced to mining gene clusters with tri-modal NRPS homologous to AldpA. When analyzing the gene clusters for natural product biosynthesis from the sequenced Penicillium thymicola genome, we found a new gene cluster (named Anc) containing a tri-modal NRPS which was supposed to be responsible for a fungal quinazoline tripeptidyl alkaloid biosynthesis with unknown structure. Aspergillus nidulans has been established as a convenient expression host, which could increase the level of gene
expression by strong/inducible promoters, amenable to large-scale fermentation, and easily manipulated (Nielsen, Nielsen, Anyaogu, Holm, Nielsen, & Larsen, 2013; Ryan, Moore, & Panaccione, 2013; Yin et al., 2013), it is a perfect host to express fungal natural products biosynthetically.

Herein, we reported the genome mining, heterologous expression, and bioactivity of a fungal peptidyl alkaloid anacine, which encoded biosynthetic gene clusters. For transformant screening AncA-P1 CTTCATCCCCAGCATC
AncA-P2 GATTGTGAAGAAATGCCTCG
AncA-P3 GAAATAAAGTGGCACGAAAGT
AncA-P4 CTTGCTGGAATTGAGGAGAT
AncA-P5 GCCCACTTGCTCGTTAGG
AncA-P6 CTGCAGCCCGGGGGA

2.2. Sequencing and bioinformatic analysis

The genomic DNA (gDNA) of P. thymicola was prepared from mycelium stationary grown on PDA at 25 °C for 7 d. The mycelia (100 mg) was ground in liquid nitrogen, and then gDNA was extracted from the resulting cell powder using cetyltrimethylammonium bromide (CTAB) method. Shotgun sequencing was performed at Novogene Bioinformatics Technology Co., Ltd. (Tianjin, China) using Illumina Hiseq2000 platform. Gene clusters were predicted by antiSMASH (https://fungismash.secondarymetabolites.org) and 2ndFind (http://biosyn.nih.go.jp/2ndFind) analysis. The amino acid sequences of AncA (the target NRPS encoded by AncA in AncA cluster) were submitted to the web-based DOMAIN SEARCH PROGRAM for NRPS (http://www.nii.ac.in/~zeeshan/search_only_nrrps.html) to analyze the functional domain. The 10-residue specificity sequence (10AA code) of A domain in NRPS was exacted with Q5 High-Fidelity DNA Polymerase (New England Biolabs). The gene sequence of AncA was divided into three parts and amplified from the genomic DNA of P. thymicola, respectively. The plasmid pYTU was digested with PmeI and NdeI. All of the fragments and linearized vector were gel purified using AxyPrep DNA Gel Extraction Kit (Axygen). The vector pYTU-AncA was constructed in vivo yeast recombination. Three NRPS fragments and linearized vector pYTU were transformed into S. cerevisiae using yeast homologous recombination method as described by R Daniel Gietz subsequently (Gietz & Schiestl, 2007). All three DNA fragments for in vivo yeast recombination contained a minimum of 30 bp overlapping bases with the flanking fragments. The obtained yeast colonies were characterized by PCR. Yeast plasmids were isolated using Zymoprep™ (D2001) Kit (Zymo Research) and transformed into E. coli XL1-Blue cells. All plasmids were isolated using AxyPrep Plasmid Miniprep Kit (Axygen) and confirmed by restriction enzyme digestion and sequencing.

Table 1

| Strains/Plasmids | Description/Aim |
|------------------|-----------------|
| Anc-AncA | Overexpression AncA in A. nidulans |
| Penicillium thymicola | Genomic origin |
| Aspergillus nidulans | pyrG98, pyrA4, riboB2, heterologous host |
| Saccharomyces cerevisiae | Yeast homologous recombination host |
| Escherichia coli XL1-Blue | Plasmids multiplication |
| pYTU | URA3, AMA1, glaA::ApHpyG, Amp |
| pYPY | URA3, AMA1, amylB::ApHpyA, Amp |
| pYPY-AncA | AncA DNA in pANU under the control of glaA promoter |

2.3. Gene amplification and plasmid construction

The plasmids utilized in this work were shown in Table 1. The PCR primers were listed in Table 2. PCR reactions were performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs). The gene sequence of AncA was divided into three parts and amplified from the genomic DNA of P. thymicola, respectively. The plasmid pYTU was digested with PmeI and NdeI. All of the fragments and linearized vector were gel purified using AxyPrep DNA Gel Extraction Kit (Axygen). The vector pYTU-AncA was constructed in vivo yeast recombination. Three NRPS fragments and linearized vector pYTU were transformed into S. cerevisiae using yeast homologous recombination method as described by R Daniel Gietz subsequently (Gietz & Schiestl, 2007). All three DNA fragments for in vivo yeast recombination contained a minimum of 30 bp overlapping bases with the flanking fragments. The obtained yeast colonies were characterized by PCR. Yeast plasmids were isolated using Zymoprep™ (D2001) Kit (Zymo Research) and transformed into E. coli XL1-Blue cells. All plasmids were isolated using AxyPrep Plasmid Miniprep Kit (Axygen) and confirmed by restriction enzyme digestion and sequencing.

2.4. Transformation of A. nidulans

A. nidulans strain was used as the recipient host. Fungal protoplast preparation and transformation were performed as described from our laboratory (Yan et al., 2019). The plasmids pYTU-AncA, pYTR and pYPY were co-transformation into A. nidulans to own the AncA overexpression strain A.n-AncA. Co-Transformation with plasmids pYTU, pYTR and pYPY were used as control. Transformants were verified using diagnostic PCR with appropriate primers (Table 2).

2.5. Fermentation and LC-MS analysis

A. nidulans strain A.n-AncA was cultivated at 25 °C, 180 rpm in 5 ml liquid CD-ST medium (GMM liquid medium containing 20 g/L starch without glucose, 2% tryptone is added) for 5 d. The broth (2 ml) was extracted three times with ethyl acetate (EtOAc), the supernatant organic phase gave an oily residue (Labconco Corporation, Dry Evaporators, Concentrators & Cold Traps, MO, USA). The dried material was dissolved in 150 μL of acetonitrile (MeCN) and subjected to LC-MS analysis.

LC-MS were performed on a Waters ACQUITY H-Class UPLC-MS with QDA mass detector (ACQUITY UPLC® BEH, 1.7 μm, 50 mm × 2.1 mm, C18 column) using positive and negative mode electrospray ionization. LCMS grade MeCN and H2O (both with 0.02% formic acid, volume percentage) were used as the mobile phases. The sol-
vent gradient for LC-MS analyses changed as follows: 0–8 min 5%–99% MeCN, 8–12 min 99% MeCN, with a flow rate of 0.4 mL/min.

2.6. Extraction, purification and identification of product from heterologous expression

The desired product was purified from A. n.-AncA 7 d cultures in 3 L liquid CD-ST medium. The aqueous phase was filtered from the fermentation broth and then extracted with analytical grade EtOAc (1:1, volume percentage) for three times to offer the EtOAc extract. After evaporation of the organic phase, the EtOAc extract was dissolved with HPLC grade MeOH and concentrated by centrifugation to give the MeOH-soluble extract (452.5 mg). The MeOH-soluble extract was fractionated by semi-preparative HPLC with a DAD detector (SSI series 1500, CoMetro Technology Ltd., NJ, USA) and analyzed by LC-MS (Fig. 3A), where a new metabolite \( [M + H]^+ \) was observed compared to the control. Following PCR verification of the integration of pYTU-AncA, the strain A.n.-AncA was cultivated in CD-ST media, and the metabolites were extracted and analyzed by LC-MS (Fig. 3A), where a new metabolite with \( m/z = 343 \) \([M + H]^+\) was observed compared to the control.

Table 3

| Names (module) | Pos1 (235) | Pos2 (236) | Pos3 (239) | Pos4 (278) | Pos5 (299) | Pos6 (301) | Pos7 (322) | Pos8 (330) | Pos9 (331) | Pos10 (517) |
|---------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| NFIA_057960(A1) | G          | I          | I          | L          | G          | A          | A          | G          | I          | K           |
| AnaPS/NFIA_055290(A1) | G          | A          | L          | F          | F          | A          | A          | G          | V          | K           |
| ACLA_017890(A1) | G          | V          | I          | F          | L          | A          | A          | G          | V          | K           |
| ACLA_076770(A1) | G          | V          | I          | F          | V          | A          | G          | G          | V          | K           |
| ACLA_095980(A1) | G          | V          | I          | I          | L          | A          | G          | G          | L          | K           |
| TqaA (A1) | G          | V          | I          | F          | M          | A          | A          | G          | V          | K           |
| AFUA_6g12080(A1) | G          | V          | I          | L          | L          | A          | A          | G          | I          | K           |
| Aldp(A1) | G          | I          | I          | L          | G          | A          | A          | G          | I          | K           |
| CtgA (A1) | G          | I          | I          | L          | G          | A          | A          | G          | I          | K           |
| AncA (A1) | G          | I          | I          | L          | L          | A          | A          | G          | I          | K           |
| Consensus* | G          | X_h        | I/L         | X_h          | X_h        | A          | A/G          | G          | X_h        | K           |

* The abbreviation “X_h” stands for variable hydrophobic residues.

3. Results

3.1. Identification of Anc gene cluster containing a tri-modular NRPS by bioinformatic analysis

Table 3: Comparison of 10-residue specificity sequences for selecting NRPS adenylation domains.

Fig. 1. Gene cluster of Anc.

3.2. Heterologous expression of AncA in A. nidulans led to production of anacine

To investigate the product encoded by this tri-modular NRPS, we expressed the putative gene cluster in A. nidulans, an established heterologous expression system for production of cryptic natural products. The gene AncA was amplified from gDNA and then cloned directly into pYTU which contains an inducible strong glaA promoter using yeast recombination. Following PCR verification of the integration of pYTU-AncA, the strain A.n.-AncA was cultivated in CD-ST media, and the metabolites were extracted and analyzed by LC-MS (Fig. 3A), where a new metabolite with \( m/z = 343 \) \([M + H]^+\) was observed compared to the control.
The new metabolite (1) was purified as amorphous powder (452.5 mg) from the crude extract and fully characterized by NMR analyses (Fig. 4). The 1D NMR spectra data (Table 4) indicated the presence of one ortho-disubstituted benzene ring, three carbonyls, four aromatic protons, three methines, three methylenes, and two methyl groups. Also, the 13C NMR data (Table 4) also revealed the presence of two olefinic groups at δC 126.6, 127.0, 127.3, and 135.1. Comparison of the observed and reported spectroscopic data, compound 1 was identified as anacine, a known natural compound previously found in *P. aurantiogriseum* (Wang & Sim, 2001). Based on the specific optical rotation ([α]D + 142.9, c 0.02 in MeOH), the absolute configuration of 1 could be confirmed. The identified configuration of 1 was also consistent with the feature that no E domain in the tri-modular NRPS AncA. Although 1 was identified as known compound, heterogeneous expression was able to obtain sufficient amount of material for biological assay.

### 3.3. Biological evaluation of anacine (1)

The cytotoxicities of anacine (1) were evaluated against a series of cancer cell lines by MTT method, including PC3, MDA-MB-231 and MCF-7, A549, as well as HL-60 and THP-1, with doxorubicin as positive control. Anacine showed moderate cytotoxic activity.
against A549 cell (IC50 = 17.3 µmol/L), which was weaker than that of positive control (IC50 = 0.1 µmol/L). Anacine was inactive against other tested cancer cell lines.

4. Discussion and conclusion

Anacine was firstly identified as benzodiazepine alkaloids from *P. aurantiogriseum*. Later investigation indicated that anacine had a quinazoline instead of a benzodiazepine core, which was further confirmed by total synthesis (Wang & Sim, 2001). Biosynthetically, it can be synthesized by a tri-modular NRPS with three adenylation domains that incorporates anthranilate (Ant), L-glutamine (L-Gln) and L-leucine (L-Leu). The proposed biosynthetic mechanisms of anacine (1) were shown in Fig. 5. Non-ribosomal peptide synthetase AncA with three modules are responsible for the assembly of the tricyclic scaffold via a process from recognition of Ant to upload the L-Gln and L-Leu in turn. Nucleophilic attack from Ant-free amine to the thioester carbonyl contributes to achieve cyclization and release of the product. As the final step, there are two possible routs of cyclization catalyzed by C7 domain in the formation of anacine (1). In route A, the Ant amine attacks on the thioester carbonyl to yield a macrocyclic intermediate, which can later give rise to the ancine (1) through a bridging amide bond cyclization and dehydration. In route B, the L-Gln amide act as an initial nucleophile attacks on the thioester carbonyl to yield a diketopiperazine (DKP) intermediate. Subsequently, The Ant amine and the L-Leu amide carbonyl undergo intramolecular attack and dehydration to form anacine (1).

Anacine has been reported to be produced by strain *P. aurantiogriseum* with a yield of 0.4 mg/L (Boyes-Korkis et al., 1993; Xin
et al., 2007). Here, we obtained this compound with high yield (35.23 mg/L) through concise biosynthesis by only one single tri-modular NRPS. Our research also supported that heterologous biosynthesis is an efficient strategy for mining cryptic gene clusters from fungi and to obtain enough material of minor secondary metabolites.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported financially by the Drug Innovation Major Project (No. 2018ZX09711001-006) and the CAMS Initiative for Innovative Medicine (No. 2017-12M-4-004). We are grateful to the Department of Instrumental Analysis, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College for the spectroscopic measurements.

References

Ames, B. D., Liu, X., & Walsh, C. T. (2010). Enzymatic Processing of Fumiquinazoline F: A Tandem Oxidative-Acylation Strategy for the Generation of Multicyclic Scaffolds in Fungal Indole Alkaloid Biosynthesis. Biochemistry, 49(39), 8564–8576.
Ames, B. D., & Walsh, C. T. (2010). Anthranilate-Activating Modules from Fungal Nonribosomal Peptide Assembly Lines. Biochemistry, 49(15), 3351–3365.
Gao, X., Haynes, S. W., Ames, B. D., Wang, P., Vien, L. P., Walsh, C. T., & Tang, Y. i. (2012). Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. Nature Chemical Biology, 8(10), 823–830.
Gietz, R. D., & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nature Protocols, 2(1), 31–34.
Haynes, S. W., Ames, B. D., Gao, X., Tang, Y. i., & Walsh, C. T. (2011). Unraveling Terminal C-Domain-Mediated Condensation in Fungal Biosynthesis of Imidazoindolone Metabolites. Biochemistry, 50(25), 5668–5679.
Haynes, S. W., Gao, X., Tang, Y. i., & Walsh, C. T. (2013). Complexity Generation in Fungal Peptidyl Alkaloid Biosynthesis: A Two-Enzyme Pathway to the Hexacyclic MDR Export Pump Inhibitor Ardeemin. ACS Chemical Biology, 8(4), 741–748.
Hochlovsikis, J. E., Mullally, M. M., Spanton, S. G., Whittern, D. N., Hill, P., & Mcalpine, J. B. (1993). 5-N-Acetylardeemin, a novel heterocyclic compound which reverses multiple drug resistance in tumor cells. II. Isolation and elucidation of the structure of 5-N-acetylardeemin and two congeners. Journal of Antibiotics, 46(3), 380–386.
Liesch, J. M., Hensens, O. D., Zink, D. L., & Goetz, M. A. (1988). Novel cholecystokinin antagonists from Aspergillus alliaceus. II. Structure determination of asperlicins B, C, D, and E. Journal of Natural Products, 56(10), 1707–1717.
Nielsen, M. T., Nielsen, J. B., Anyaogu, D. C., Holm, D. K., Larsen, T. O., et al. (2013). Heterologous reconstitution of the intact geodin gene cluster in Aspergillus nidulans through a simple and versatile PCR based approach. PLoS One, 8(8), e7871.
Rottig, M., Medema, M. H., Blin, K., Weber, T., Rausch, C., & Kohlbacher, O. (2011). NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Research, 39(Web Server issue), W362-327.
Ryan, K. L., Moore, C. T., & Panaccione, D. G. (2013). Partial reconstruction of the ergot alkaloid pathway by heterologous gene expression in Aspergillus nidulans. Toxins (Basel), 5(2), 445–455.
Takahashi, C., Matsushita, T., Doi, M., Minoura, K., Shingu, T., Kumeda, Y., et al. (1995). Fumiquinazolines A-G, novel metabolites of a fungus separated from a Pseudolabrus marine Fish. Journal of the Chemical Society, Perkin Transactions, 1, 2345–2353.
Walsh, C. T., Haynes, S. W., Ames, B. D., Gao, X., & Tang, Y. (2013). Short pathways to complexity generation: Fungal peptidyl alkaloid multicyclic scaffolds from anthranilate building blocks. ACS Chemical Biology, 8(7), 1360–1382.
Wang., H., & Sim., M. M. (2001). Total solid phase syntheses of the quinazoline alkaloids: Verrucines A and B and anacine. Journal of Natural Products, 64(12), 1497–1501.

Xin, Z. H., Fang, Y., Du, L., Zhu, T., Duan, L., Chen, J., et al. (2007). Aurantiomides A–C, quinazoline alkaloids from the sponge-derived fungus Penicillium aurantiogriseum SP0-19. Journal of Natural Products, 70(5), 853–855.

Yan, D., Chen, Q., Gao, J., Bai, J., Liu, B., Zhang, Y., et al. (2019). Complexity and diversity generation in the biosynthesis of fumiquinazoline-related peptidyl alkaloids. Organic Letters, 21(5), 1475–1479.

Yin, W., Chooi, Y. H., Smith, A. R., Cacho, R. A., Hu, Y., White, T. C., et al. (2013). Discovery of cryptic polyketide metabolites from dermatophytes using heterologous expression in Aspergillus nidulans. ACS Synthetic Biology, 2(11), 629–634.

Yin, W., Grundmann, A., Cheng, J., & Li, S. (2009). Acetylaszonalenin biosynthesis in Neosartorya fischeri. Journal of Biological Chemistry, 284(1), 100–109.