Acetylaszonalenin biosynthesis in Neosartorya fischeri: Identification of the biosynthetic gene cluster by genomic mining and functional proof of the genes by biochemical investigation

Wen-Bing Yin, Alexander Grundmann, Jun Cheng and Shu-Ming Li *
Philipps-Universität Marburg, Institut für Pharmazeutische Biologie, Deutschhausstrasse 17A, D-35037 Marburg, Germany

Correspondence to: Shu-Ming Li (email: shuming.li@staff.uni-marburg.de), Tel: 0049-6421-2822461
Fax: 0049-6421-2826678

Running title: acetylaszonalenin biosynthesis in Neosartorya fischeri

Key words: acetylaszonalenin biosynthesis, aszonalenin, N-acetyltransferase, benzodiazepinedione, Neosartorya fischeri, prenyltransferase

ABSTRACT
Based on the structural information of acetylaszonalenin isolated from Neosartorya fischeri, a putative biosynthetic gene cluster was identified in the genome sequence of this fungus by genomic mining. This cluster consists of three genes coding for a putative non-ribosomal peptide synthetase (AnaPS), a prenyltransferase (AnaPT) and an acetyltransferase (AnaAT). The coding sequences of anaPT and anaAT were cloned in pQE70 and pQE60, respectively, and overexpressed in Escherichia coli. The soluble His6-fusion proteins were purified to near homogeneity and characterised biochemically. The structures of the enzymatic products were elucidated by NMR and MS analysis. AnaPT was found to catalyse the reverse prenylation of (R)-benzodiazepinedione at position C3 of the indole moiety in the presence of dimethylallyl diphosphate resulting in formation of aszonalenin. AnaAT was found to catalyse the acetylation of aszonalenin at position N1 of the indoline moiety in the presence of acetyl coenzyme A resulting in formation of acetylaszonalenin.

INTRODUCTION
Prenylated indole derivatives are hybrid natural products containing both aromatic and isoprenoid moieties and are found especially in fungi of the genera Aspergillus and Penicillium (1,2). Besides L-tryptophan, the precursor of indole rings, most of these substances contain a second amino acid, forming cyclic dipeptides with a diketopiperazine structure or a derivative thereof (1). Several members of this group carry reverse prenyl moieties (3´-(3´, 3´)-dimethylallyl = 3´-DMA) at position C3 of the indole ring and a ring system between the indole and the diketopiperazine ring (Fig. 1). For example, acetylaszonalenin (Fig. 1), also called LL-S490β, is a mycotoxin derived from tryptophan, anthranilic acid, dimethylallyl diphosphate and acetyl coenzyme A. Acetylaszonalenin was isolated initially from an unidentified Aspergillus species (3). Together with its non-acetylated form aszonalenin, it was also identified in various fungal strains, e.g. Aspergillus zonatus (4), A. fumigatus (5), A. carneus (6) and Neosartorya fischeri IFM52672 (7). Their stereoisomer epi-aszonalenins A (Fig. 1) and C were isolated from Aspergillus novofumigatus (8). Roquefortine C (Fig. 1) containing tryptophan and histidine moieties was isolated firstly from Penicillium roqueforti (9) and identified later in many Penicillium strains (10,11), including different strains of P. roqueforti used for cheese production (9,11). Roquefortine C is also the precursor of further prenylated indole alkaloids such as meleagrine (12), oxaline (12,13) and glandicolines (12). Amauromine (Fig. 1) from Amauroascus sp. (14) and epiamauromine (Fig. 1) from Aspergillus ochraceus (15) are prenylated diketopiperazines derived from two tryptophan molecules and carry two 3´-DMA moieties. Fructigenines A and B (Fig. 1) from Penicillium fructigenum (16), which were also identified as rugulosuvine B in Penicillium regulosum (17) and verrucofortin in Penicillium verucosum var. cyclopium (18),
respectively, are cyclic dipeptide derivatives of tryptophan and phenylalanine or tryptophan and leucine. In the structure of 5-N-acetylardeemin (Fig. 1), which was identified in Neosartorya fischeri var. brasiliensis (19), anthranilyl moiety is connected to the diketopiperazine ring consisting of tryptophan and alanine.

Despite the large number of the known structures in this group, little is known about their biosynthesis. Investigation on the biosynthesis of these compounds was limited to precursor feeding experiments in Penicillium roqueforti for roquefortine and in Aspergillus zonatus for acetylaszonalenin (1,20,21). It could be demonstrated that tryptophan was the biosynthetic precursor in the biosynthesis of both compounds (20,21). Histidine and mevalonate were involved in the formation of roquefortine C (22,23). No reports on biosynthetic genes or enzymes could be found in the literature. Availability of genome sequences of bacteria and fungi, which has been increased tremendously in the last years (24,25), accelerates the identification of genes involved in the biosynthesis of secondary metabolites (26-28). From the genome reference strain Neosartorya fischeri NRRL181, we have isolated a prenylated indole derivative and identified as acetylaszonalenin. Here we report on the identification and functional proof of the biosynthetic gene cluster of this compound.

**EXPERIMENTAL PROCEDURES**

**Chemicals.** Dimethylallyl diphosphate was prepared according to the method described for geranyl diphosphate by Woodside et al. (29). (R)-benzodiazepinedione was synthesized by condensation of D-tryptophan with isatoic anhydride in the presence of triethylamine according to the method described by Barrow and Sun (30).

**Computer-assisted sequence analysis.** FGENESH and the DNALIGN software package (version 2.1: Hitachi Software Engineering, San Bruno, CA) were used for exon prediction and sequence analysis, respectively. Sequence identities were obtained by alignments of amino acid sequences using the BLAST program “BLAST 2 SEQUENCES”.

**Bacterial strains, plasmids and cultural conditions for E. coli.** pGEMT easy vector was obtained from Promega (Mannheim, Germany), pQE60 and pQE70 from Qiagen (Hilden, Germany), respectively.

Escherichia coli XLI Blue MRF’ (Stratagene, Amsterdam, Netherlands) was used for cloning and overexpression experiments and grown in liquid or on solid Luria-Bertani medium with 1.5 % (w/v) agar at 37 °C. Carbenicillin (50 µg·mL⁻¹) was used for selection of recombinant E. coli strains.

Neosartorya fischeri NRRL181 was kindly provided by ARS Culture Collection (Peoria, Illinois USA).

**Cultivation of N. fischeri and extraction of secondary metabolites.** N. fischeri mycelia from agar plate was inoculated and cultivated in 300 mL Erlenmeyer flasks containing 100 mL modified Czapek Dox liquid medium at 30 °C without shaking (31). After cultivation for three weeks, the cultures were extracted with ethyl acetate (100 mL). The ethyl acetate phase was evaporated to dryness on a rotary evaporator at 37 °C. The residue was dissolved in methanol (1 mL), analysed and purified on HPLC. 0.2 mg acetylaszonalenalin could be obtained from 100 mL culture.

**Cultivation of N. fischeri for DNA isolation.** For DNA isolation, mycelia of N. fischeri from plates were inoculated into 300 mL Erlenmeyer flask containing 100 mL YES media consisting of yeast extract (0.6 % (w/v)), sucrose 0.2 % (w/v) (pH 5.8) and cultivated at 30 °C and 170 rpm for 48 hrs.

**DNA propagation in E. coli.** Standard procedures for DNA isolation and manipulation were performed as described (32).

**DNA isolation from N. fischeri.** DNA isolation from N. fischeri was carried out according to the protocol described by Ausubel et al.(33).

**PCR amplification and gene cloning.** PCR amplification was carried out on an iCycler from BioRad (Munich, Germany). The entire coding sequence of AnaPT was obtained after 3 rounds PCR amplification by using genomic DNA as template and Expand High Fidelity Kit (Roche Diagnostics GmbH, Mannheim, Germany). The first round PCR is used for separate amplification of the two exons. The primers for the first exon were AnaPT_a1 (5’-GTCCGCAGTGTAGTAGTATACA-3’) at the 5’-end and AnaPT_a2 (5’-GTAGTGGTCTTCTGGAGGTTTTGAC AAGGTTTGGCATAGAATCCTT-3’) at the 3’-end. The primers for the second exon were AnaPT_b1 (5’-GCAGTACAAGAGAAGATCTCTATGCGC by guest on September 1, 2017 http://www.jbc.org/ Downloaded from
aacaaacctttgctgaacacctcgcaag-3') at the 5'-end and AnaPT_b2
(5'-CAATTGCTACCTATGGAAGCCA -3') at the 3'-end. Bold underlined letters represent
overlapping region of two exons. The PCR products of first and the second exon were
mixed in a molar ratio of 1:1 and used as template for a second round of PCR to get a
fragment consisting of the two exons with help of the overlapping region. A third round PCR
was then carried out by using a nested primer pair and the PCR product from the second round
as template. The nested primers are AnaPT_for (5´-CTGCATGCCCTCCTTGTCTATGCAA-3´)
at the 5´-end and AnaPT_rev (5´-GAAGATCCTGAGATTGCCCTTCATACC-3´)
at the 3´-end of the gene. Bold letters represent
mutations inserted in comparison to the original
genome sequence to give the underlined
restriction sites SphI located in the start codon in
AnaPT_for and BglII located in the predicted
stop codon in AnaPT_rev, respectively. A PCR
fragment of 1323 bp containing the entire coding
sequence of anaPT could be amplified after
three round PCR. The PCR fragment was cloned
into pGEMT easy vector resulting in plasmid
pWY16, which was subsequently sequenced
(Eurofins MWG Operon, Ebersberg, Germany)
to confirm the sequence. To create the
expression vector pWY22, pWY16 was digested
with SphI and BglII and the resulted SphI-BglII
fragment of 1313 bp was ligated into pQE70,
which had been digested with the same enzymes, previously.

For the amplification of anaAT, a PCR
fragment of 1540 bp containing the entire coding
sequence of NAIA_055310 was amplified from
the genomic DNA by using the primers
AnaAT_for (5´-ATTCCATGGCCTGACCATCAGTTTC-3´)
at the 5´-end and AnaAT_rev (5´-GAAGATCCTTCTTTCCATGCCAGAA
C-3´) at the 3´-end of the gene. Bold letters represent
mutations inserted in comparison to the original
genome sequence to give the underlined
restriction sites NcoI located at the start codon in
AnaAT_for and BglII located at the predicted stop codon in AnaAT_rev. The PCR
fragment was cloned into pGEM-T easy vector resulting in plasmid pWY18, which was subsequently sequenced to confirm the sequence. To create the expression vector pWY23, pWY18 was digested with NcoI and BglII and the resulted NcoI - BglII fragment of 1533 bp was
ligated into pQE60, which had been digested
with the same enzymes, previously.

Overproduction and purification of His_{6}-AnaPT and His_{6}-AnaAT. For anaPT
expression, E. coli XL1 Blue MRF' cells harbouring the plasmid pWY22 were cultivated
in 300 ml Erlenmeyer flasks containing 100 ml
liquid Luria-Bertani medium supplemented with
carbenicillin (50 µg ml^{-1}) and grown at 37 °C to
an A_{600} of 0.7. For induction, isopropyl thiogalactoside (IPTG) was added to a final
concentration of 0.5 mM and the cells were
cultivated for further 6 h at 37 °C before harvest.
The bacterial cultures were centrifuged and the
pellets were resuspended in lysis buffer (10 mM
imidazole, 50 mM NaH_{2}PO_{4}, 300 mM NaCl, pH 8.0) at 2-5 ml per gram wet weight. After
addition of 1 mg ml^{-1} lysozyme and incubation
on ice for 30 min, the cells were sonicated 6
times for 10 seconds each at 200 W. To separate
the cellular debris from the soluble proteins, the
lysate was centrifuged at 13,000 x g for 30
min at 4°C. One-step purification of the recombinant
His_{6}-tagged fusion protein by affinity
chromatography with Ni-NTA agarose resin
(Qiagen, Hilden, Germany) was carried out
according to the manufacturer’s instructions. The protein was eluted with 250 mM imidazole
in 50 mM NaH_{2}PO_{4}, 300 mM NaCl, pH 8.0. In
order to change the buffer, the protein fraction
was passed through a NAP-5 column (GE
Healthcare, Freiburg, Germany), which had been
equilibrated with 50 mM Tris-HCl, 15 % (v/v) of
glycerol, pH 7.5, previously. AnaPT was eluted
with the same buffer and stored frozen at - 80 °C
for enzyme assays.

For anaAT expression, E. coli XL1 Blue
MRF' cells harbouring the plasmid pWY23
were cultivated in 300 ml Erlenmeyer flasks
containing 100 ml liquid Luria-Bertani medium
supplemented with carbenicillin (50 µg ml^{-1}) and
grown at 37 °C to an A_{600} of 0.6. For induction,
isopropyl thiogalactoside (IPTG) was added to a
final concentration of 0.7 mM and the cells were
cultivated for further 16 h at 37 °C before harvest. Protein purification was carried out in
analogy to AnaPT.

Protein analysis. Proteins were analysed by
SDS-PAGE according to the method of Laemmli
(34), and stained with Coomassie Brilliant Blue
G-250.

Determination of molecular mass of active
His_{6}-AnaPT and active His_{6}-AnaAT. The
molecular mass of the recombinant His_{6}-AnaPT
and His6-AnaPT was determined by gel filtration on a HiLoad 16/60 Superdex 200 column (GE Health Care, Freiburg, Germany), which had been equilibrated with 20 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl. The column was calibrated with dextran blue 2000 (2000 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa) (GE Health Care, Freiburg, Germany). The proteins were eluted with 20 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl. The molecular mass of the recombinant His6-AnaPT was determined as 162.2 kDa. This indicated that AnaPT acts likely as a tetramer. The molecular mass of the recombinant His6-AnaAT was determined as 43 kDa. This indicated that AnaAT acts as a monomer.

Assay for AnaPT and AnaAT activity.

For quantitative determination of the AnaPT activity, the reaction mixture (100 µL) contained 50 mM Tris-HCl (pH 7.5), 10 mM CaCl2, 1 mM (R)-benzodiazepinedione, 1 mM DMAPP, 1.5 % (v/v) glycerol, and 1.0 µg of purified recombinant AnaPT. The reaction mixtures were incubated at 37 °C and the reactions were terminated by addition of 100 µl methanol per 100 µl reaction mixtures. The protein was removed by centrifugation at 13,000 x g for 10 min. The enzymatic products were analysed by HPLC under conditions described below. For quantitative measurement of the enzyme activity, duplicate values were determined routinely. The assays for determination of the kinetic parameters (100 µl) of (R)-benzodiazepinedione contained 1 mM DMAPP, 0.5 µg of AnaPT and (R)-benzodiazepinedione at final concentrations of 0.0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, and 1.0 mM. The incubation time was 30 min. For determination of kinetic parameters of DMAPP, (R)-benzodiazepinedione at 1 mM and DMAPP at final concentrations of 0.0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, and 1.0 mM were used.

For quantitative determination of the AnaAT activity, the reaction mixtures (100 µL) contained 50 mM Tris-HCl (pH 7.5), 10 mM CaCl2, 1 mM aszonalenin, 1 mM acetylcoenzyme A, 1.5 % (v/v) glycerol, and 1.0 µg of purified recombinant AnaAT. The incubation and termination conditions were identical to those for AnaPT. The assays for determination of the kinetic parameters (100 µl) of aszonalenin contained 1 mM acetylcoenzyme A, 1 µg of AnaAT and aszonalenin at final concentrations of 0.0, 0.005, 0.01, 0.02, 0.05, 0.2. The incubation time was 30 min. For determination of kinetic parameters of acetylcoenzyme A, aszonalenin at 1 mM and acetylcoenzyme A at final concentrations of 0.0, 0.005, 0.01, 0.02, 0.05, 0.2, 0.5 and 1.0 mM were used.

Preparative synthesis of enzymatic products for structural elucidation. AnaPT assay for isolation of the enzymatic product (10 mL) contained DMAPP (1 mM), (R)-benzodiazepinedione (1 mM), CaCl2 (10 mM), Tris-HCl (50 mM, pH 7.5), glycerol 1.5 % (v/v) and AnaPT (0.8 mg). AnaAT assay for isolation of the enzymatic product (10 mL) contained acetylcoenzyme A (1 mM), aszonalenin (1 mM), Tris-HCl (50 mM, pH 7.5), glycerol 1.5 % (v/v) and AnaAT (0.8 mg). The reaction mixtures were incubated at 37°C for 16 h and extracted subsequently with ethyl acetate. After evaporation of the solvent, the residues were dissolved in methanol and purified on HPLC under condition described below. 2.8 mg of aszonalenin and 1.5 mg of acetylaszonalenin could be obtained from the AnaPT and AnaAT incubations, respectively. The isolated products were subjected to 1H-NMR, 13C-NMR, H-H-COSY, C-H-COSY as well as positive and negative electrospray ionization (ESI) mass spectrometry with a ThermoFinnigan TSQ Quantum. The mass spectrometer was coupled with an Agilent HPLC series 1100 equipped with a RP18 column (250 x 2 mm, 5 µm). For separation, the column was run with 10 % (v/v) solvent B (CH3OH) in solvent A (H2O) (each containing 0.1 % (v/v) HCOOH) for 5 min, followed by a gradient from 10 % (v/v) to 100 % B over 30 min. After washing with 100 % B for 10 min, the column was equilibrated with 10 % (v/v) B for 10 min. The flow rate was at 0.2 mL·min⁻¹.

HPLC conditions for analysis and isolation of enzymatic products of AnaPT and AnaAT. The enzymatic products of the incubation mixtures of AnaPT and AnaAT were analysed by HPLC on an Agilent series 1200 by using a LiChrospher RP 18-5 column (125 x 4 mm, 5 µm, Agilent) at a flow rate of 1 mL·min⁻¹. 50 % methanol in water (solvent A) and methanol (solvent B) were used as solvents. For analysis of enzymatic products, a linear gradient of 10 - 50 % (v/v) solvent B in 15 min was used. The column was then washed with 100 % solvent B for 5 min and equilibrated with 10 % (v/v) solvent A for 5 min. Detection was carried out by a Photo Diode Array detector and illustrated...
at 254 nm in the figures in this paper.

For isolation, the same HPLC equipment with a Multospher 120 RP-18 column (250 x 10 mm, 5 µm, C+S Chromatographie Service, Langenfeld, Germany) was used. A linear gradient of 30 - 70 % (v/v) solvent B in 12 min at a flow rate of 2.5 mL•min⁻¹ was used. The column was then washed with 100 % solvent B for 8 min and equilibrated with 30 % (v/v) solvent A for 5 min.

Positive ESI-MS data of aszonalenin m/z (intensity): 374.3 [m+1]⁺; ms² of [m+1]⁺: 306.2 (70), 318.2 (47), 198.2 (100)

Positive ESI-MS data of acetylaszonalenin m/z (intensity): 416.4 [m+1]⁺; ms² of [m+1]⁺: 374.1 (35), 348.1 (100), 306.1 (60)

Negative ESI-MS data of acetylaszonalenin m/z (intensity): 414.7 [m⁻1]⁻; ms² of [m⁻1]⁻: 414.3 (100), 372.5 (65), 259.5 (15)

Nucleotide sequence accession number: The nucleotide sequence of the genomic DNA from Neosartorya fischeri NRRL181 reported in this study is available at GenBank under accession number AAKE03000024.1. The coding sequence of anaPT is available at GenBank under accession number XM_001258077. The coding sequence of anaAT is available at GenBank under accession number XM_001258078.

RESULTS

Isolation and structural elucidation of acetylaszonalenin from N. fischeri NRRL181. In course of our investigations on prenylated indole derivatives, we have successfully identified genetic information for the biosynthesis of fumigaclavines and fumitremorgins by analysis of their structural features and the genome sequence of Aspergillus fumigatus (35,36). To expand the usage of this strategy, indole derivatives should be isolated and identified from fungal strains, whose genome sequences are available. HPLC analysis of the ethyl acetate extracts of three weeks old cultures of N. fischeri NRRL181 showed a dominant peak at 10.9 min (Fig. 2) with a typical UV spectrum of indole derivatives (data not shown). This compound was isolated on a preparative scale and subjected to ¹H-NMR and MS analysis. ¹H-NMR spectrum showed clearly signals for a reverse prenyl moiety at 5.89 (1H, dd, 17.4, 10.4 Hz), 5.14 (1H, d, 10.4 Hz), 5.13 (1H, d, 17.4, Hz), 1.19 (3H, s) and 0.99 (3H, s) ppm (Table 1), respectively (37). This was also confirmed by H-H-COSY analysis (data not shown). A singlet for three protons at 2.59 ppm could be assigned for an acetyl group. A three-spin system at 3.90 (1H, t, 8.4 Hz), 3.37 (1H, dd, 13.9, 8.5 Hz) and 2.44 (1H, dd, 13.9, 8.2 Hz) ppm was typical for the three protons at α- and β-positions of aromatic amino acids such as tryptophan or its dipeptides (37-39). In addition, two aromatic systems with four coupled protons each were also observed (Table 1). Literature search revealed that the NMR data of the isolated compound corresponded very well to those of acetylaszonalenin (3,4). Aszonalenin was previously isolated from another N. fischeri strain IFM52672 (7). MS data confirmed the structure of the isolated compound as acetylaszonalenin.

Putative biosynthetic gene cluster of acetylaszonalenin in sequenced genomes. Based on the structural feature of acetylaszonalenin, we tried to find its biosynthetic genes or gene cluster in genome sequences. Acetylaszonalenin is chemically a cyclic dipeptide derivative of tryptophan and anthranilic acid, which is acetylated at position N1 and prenylated at position C3 of the indoline ring. It could be expected that at least three genes/enzymes are involved in the biosynthesis of acetylaszonalenin, i.e. a non-ribosomal peptide synthetase for the formation of the cyclic dipeptide (40,41), a prenyltransferase for prenyl transfer (36,37) and an acetyltransferase for acetyl transfer. The formation of the ring system between the indoline and the diketopiperazine ring would require an additional enzyme or would be catalysed during the prenylation by the prenyltransferase.

Indeed, a 12.8 kb DNA segment consisting of three genes and spanning bp 130,585 –143,349 of AAKE03000024.1 could be identified on chromosome of N. fischeri NRRL181 by homologous search and sequence analysis (Fig. 3A). The putative genes NFIA_055290, NFIA_055300 and NFIA_055310 in this cluster coding for the three putative enzymes mentioned above are likely involved in the biosynthesis of acetylaszonalenin and therefore were termed anaPS, anaPT and anaAT, respectively. An orthologous cluster was also identified in the genome sequence of Aspergillus terreus NIH2624 (Fig. 3B). anaPS is predicted to encode a dimodular non-ribosomal peptide synthetase (NRPS) with two adenylation
domains (A), two peptidyl carrier domains (P) and two condensation domains (C). AnaPS from N. fischeri (EAW16180) shares an identity of 33 % on the amino acid level with brevianamide F synthetase in the biosynthesis of fumitremorgins (40) and 24 % with GlP in the biosynthesis of gliotoxin (41). anaPT is predicted to encode an indole prenyltransferase. Analysis by using FGENESH (see Experimental Procedures) and by comparing the sequence with those of known prenyltransferases revealed that anaPT from N. fischeri consists of two exons of 1187 and 127 bp, respectively, interrupted by an intron of 51 bp (Table 2). It spans bp 139,642 to bp 141,481 in AAKE03000024.1 (GenBank). AnaPT from N. fischeri (EAW16181) shares identities of 24 and 31 % on the amino acid level with FgaPT1 and FgaPT2 in the biosynthesis of fumigaclavine C (31,35), 29 and 30 % with FtmPT1 and FtmPT2 in the biosynthesis of fumitremorgins (36,42), respectively. According to the genome annotation, the putative acetyltransferase gene anaAT from N. fischeri would correspond to a segment of 1530 bp from bp 141,820 to bp 143,349 in AAKE03000024.1 (GenBank) and consist of only one exon. AnaAT from N. fischeri and A. terreus shares only very low sequence similarity with proteins of known functions in the databases, e.g. trichothecene 3-O-acetyltransferases (43). The three enzymes AnaPS, AnaPT and AnaAT from A. terreus NIH2624, i.e. EAU29302, EAU29303 and EAU29301, share sequence identities of 76 %, 72 % and 82 % with their orthologues EAW16180, EAW16181 and EAW16182 in N. fischeri, respectively (Table 2).

From the genetic information mentioned above, a putative biosynthetic pathway could be postulated for acetylaszonalenin (Fig. 3C). It could be expected that (R)-benzodiazepinedione is then formed by condensation of tryptophan and anthranilic acid catalysed by the NRPS AnaPS. (R)-benzodiazepinedione is then converted to aszonalenin by AnaPT alone in the presence of dimethylallyl diphosphate or with additional enzymes. The last step in the biosynthesis of acetylaszonalenin would be the acetylation of aszonalenin at position N1 catalysed by AnaAT.

Cloning of anaPT, overproduction and purification of His6-AnaPT. For the cloning of anaPT, we initially tried to amplify the coding region consisting of two exons from mRNA of N. fischeri NRRL181. For this purpose, the total RNA was isolated from the liquid culture and converted to cDNA by using different cDNA synthesis kits. The coding region of anaPT was amplified by PCR using cDNA as template and the product was cloned into pGEM-T easy vector. Restriction analysis and sequencing of the obtained recombinant plasmids showed, however, that the sequence still contained the predicted intron sequence (data not shown). The reasons for these unexpected results are unclear. It could be a result of incomplete splicing of the pre-mRNA, or very probably of contamination with genomic DNA during the mRNA isolation procedures. Unfortunately, we could not get desired plasmids without intron sequence from mRNA obtained from mycelia cultivated under different conditions including different media, cultivation time and temperatures. Subsequently, we tried to clone the first exon of anaPT by PCR using genomic DNA as template. The PCR product was cloned via cloning vector pGEM-T easy into the expression vector pQE70 for E. coli, to give the expression construct pWY15. Sequencing of pWY15 confirmed the predicted sequence and revealed neither a mutation nor a frame shift. However, no gene expression could be achieved under various conditions including different E. coli strains, IPTG concentrations, cultivation media and induction temperatures (data not shown). Therefore, we decided to amplify the coding region of anaPT by homologous recombinant DNA technique which was successfully used for gene cloning (44). For this purpose, the coding region of anaPT was amplified by 3 rounds PCR using three primer pairs. The PCR fragment consisting exclusively of the two exons was cloned via pGEM-T easy into pQE70 to create the expression construct pWY22 (see Experimental Procedures).

Soluble proteins were obtained from transformants of E. coli cells harbouring pWY22 after 6 h induction by 0.5 mM of IPTG at 37 °C. His6-AnaPT was purified with Ni-NTA agarose to near homogeneity as judged by SDS-PAGE (Fig. 4) and a protein yield of 20 mg of purified His6-tagged AnaPT per litre of culture was obtained. The observed molecular mass was 50 kDa and corresponded well to the calculated value of 49.6 kDa for His6-AnaPT.

Cloning of anaAT, overproduction and purification of His6-AnaAT. The coding region of anaAT was amplified using PCR from genomic DNA of N. fischeri NRRL181 and cloned into the pGEM-T easy vector to create the cloning construct pWY18 (see Experimental Procedures). For gene expression, the coding sequence of
anaAT was released from pGEM-T easy vector and cloned into the vector pQE60 resulting in the expression plasmid pWY23.

Soluble proteins were obtained from transformants of E. coli XL1 Blue MRF’-harbouring the expression construct pWY23 after overnight induction with 0.7 mM IPTG at 37°C and subjected to purification with Ni-NTA agarose. His6-AnaAT was purified to near homogeneity as judged by SDS-PAGE (Fig. 4). The observed molecular mass was 55 kDa and corresponded very well to the calculated value of 56.7 kDa for His6-AnaAT. A protein yield of 5 mg of pure, His6-tagged AnaAT per litre of culture was obtained.

**Enzymatic activity of AnaPT and identification of the enzymatic product.** By sequence comparison of AnaPT with those of known prenyltransferases and by analysis of the gene cluster of acetylaszonalenin mentioned above, it can be predicted that (R)-benzodiazipinedione is very likely the natural substrate of AnaPT. For determination of the enzymatic activity, (R)-benzodiazipinedione was synthesized from D-tryptophan and isatoic anhydride according to the method described by Barrow and Sun (30) and incubated with the purified His6-AnaPT in the presence of DMAPP and Ca2+. HPLC analysis showed that a product peak at 16.1 min was only detected in the incubation mixture with active recombinant enzyme (Fig. 5A), but not in the mixture with heat-denaturated His6-AnaPT (Fig. 5B). Product formation was strictly dependent on the presence of AnaPT, DMAPP and (R)-benzodiazipinedione (data not shown). Linear dependence of the product formation on the amount of protein was found up to 1 µg per 100 µL assay and on the reaction time up to 30 min at 37°C.

For structural elucidation, the enzymatic product was subsequently isolated on a preparative scale and subjected to NMR and MS analysis. 1H- and 13C-NMR are given in Table 1. Analysis of the 1H-NMR data of the isolated product (Table 1) revealed that the signals of a reverse dimethylallyl moiety could be clearly observed at 6.11 (1H, dd, 17.3, 10.7 Hz), 5.14 (1H, d, 10.7 Hz), 5.11 (1H, d, 17.3 Hz), 1.14 (3H, s) and 1.06 (3H, s) ppm, respectively. In comparison to the signal for H-2 of (R)-benzodiazipinedione at 7.19 ppm, the signal for H-2 in the spectrum of the isolated product of AnaPT was strongly upfield shifted to 5.58 ppm. The signals of C2 and C3 in the 13C-NMR spectrum of the isolated product were also upfield shifted from 124.8 and 110.6 ppm in CD3OD to 81.7 and 60.7 ppm in CDCl3, respectively. This indicates that the double bond between C2 and C3 was disappeared. However, only one proton at C2, but no proton at C3 could be detected in the 1H-NMR spectrum of the isolated product. This indicates that the reverse prenyl moiety is very likely attached to position C-3 of the indoline ring, which is subsequently connected via N-12 to the diketopiperazine ring. Indeed, the signal of NH-12 of the substrate at 6.00 ppm was disappeared in the spectrum of the enzymatic product. The NMR data of the isolated compound corresponded well to those of aszonalenin described by Kimura et al. (4). Therefore, the enzymatic product of AnaPT could be unequivocally identified as aszonalenin. This was also confirmed by detection of the [M+1]+ ion at m/z 374.3 in positive ESI-MS spectrum and the [M-1] ion at m/z 372.5 in negative ESI-MS spectrum. Assignments of NMR signals in Table 1 were also confirmed by H-H-COSY and H-C-COSY (data not shown).

**Enzymatic activity of AnaAT and identification of the enzymatic product.** HPLC analysis of the incubation mixture of aszonalenin and acetylcoenzyme A showed a significant product peak with the same retention time of acetylaszonalenin isolated from N. fischeri at 10.9 min (Figs. 2 and 5C). This peak was not detected in the incubation mixture with heat-denaturated enzyme (Fig. 5D). Product formation was strictly dependent on the presence of His6-AnaAT, aszonalenin and acetylcoenzyme A (data not shown). Linear dependence of the product formation on the amount of protein was found up to 3 µg per 100 µL assay and on the reaction time up to 90 min at 37°C.

For structural elucidation, the enzymatic product of aszonalenin was isolated on a preparative scale and its structure was elucidated by NMR and MS analysis. Comparison of the 1H-NMR data of the enzymatic product with that of aszonalenin revealed the presence of an additional singlet for three protons at δ 2.59 ppm, which could be assigned to the methyl protons of an N-acetyl group. The 1H-NMR data of the enzymatic product of AnaAT corresponded well to those reported previously (4) and to those of the isolated acetylaszonalenin described above (Table 1). Therefore, the enzymatic product of AnaAT could be unequivocally identified as acetylaszonalenin. This was also confirmed by detection of the...
[M+1]$^+$ ion at m/z 416.4 in positive ESI-MS spectrum and the [M-1]$^-$ ion at m/z 414.7 in negative ESI-MS spectrum. Assignments of NMR signals in Table 1 were also confirmed by H-H-COSY (data not shown).

Biochemical properties of AnaPT and AnaAT. An important parameter of prenyltransferases is the dependence of their enzymatic activities on metal ions, especially on the presence of Mg$^{2+}$ or Ca$^{2+}$ (35,36,45-47). Therefore, the enzymatic activity of AnaPT was assayed with (R)-benzodiazepinedione and DMAPP in the presence of different metal ions at a final concentration of 5 mM. Incubations with EDTA at a final concentration of 5 mM and without additives were used as controls. Our results (data not shown) showed that product formation was independent of the presence of metal ions. Even in the presence of EDTA, a chelating agent for divalent metal ions, comparable enzymatic activity was detected to that of incubation without additives. However, Ca$^{2+}$, Mg$^{2+}$ and Co$^{2+}$ enhanced the enzymatic activity of AnaPT slightly and showed 134 %, 125 % and 116 % of relative activity, respectively, compared to that of the incubation without additives. This finding corresponded to those of known prenyltransferases e.g. DMATS from C. purpurea (48), FgaPT1, FgaPT2 and FtmPT1 from A. fumigatus (31,35,36).

Kinetic parameters of AnaPT and AnaAT. The AnaPT reaction apparently followed Michaelis-Menten kinetics (Figs. 6A and 6B). By using Lineweaver-Burk plots, the $K_m$ values were determined at 232 µM for (R)-benzodiazepinedione and 156 µM for DMAPP, respectively. The average maximum reaction velocity of AnaPT was calculated as 445 nmol $\cdot$ min$^{-1}$$ \cdot$ mg$^{-1}$ protein. The average turnover number was determined as 1.5 s$^{-1}$.

The AnaAT reaction also followed Michaelis-Menten kinetics (Figs. 6C and 6D). By using Lineweaver-Burk plots, the $K_m$ values were determined at 61 µM for aszonalenin and 96 µM for acetylcoenzyme A, respectively. The average maximum reaction velocity of AnaAT was calculated as 147 nmol $\cdot$ min$^{-1}$$ \cdot$ mg$^{-1}$ protein. The average turnover number was determined as 0.14 s$^{-1}$.

DISCUSSION

In this study, we have identified a putative gene cluster for acetylaszonalenin biosynthesis by isolation of the natural product from the producer and by genomic mining. The functional proof of the gene cluster was provided by identification of enzymatic products and biochemical characterisation of two transferases, i.e. the prenyltransferase AnaPT and the acetyltransferase AnaAT. The biosynthetic pathway described in this study could also be plausible for other dipeptide derivatives with prenyl moieties at position C3 of the indole rings as given in Fig. 1, i.e. formation of dipeptides catalysed by NRPS, prenylation and cyclization catalysed by a prenyltransferase and acetylation, if necessary, by an acetyltransferase.

Feeding experiments have shown that L-tryptophan was incorporated into aszonalenin (20,21), indicating that the original (S)-configuration of the $\alpha$-atom in L-tryptophan was converted to (R)-configuration at position C11 of aszonalenin. Prenylation and acetylation steps led to formation of aszonalenin and acetylaszonalenin with the same configuration at C11 as in (R)-benzodiazepinedione or D-tryptophan. This indicated that in the biosynthesis of aszonalenin, L-tryptophan was converted to D-tryptophan before or during formation of the cyclic dipeptide (R)-benzodiazepinedione. It is plausible that this took place during the dipeptide synthesis catalysed by an epimerase domain of the non-ribosomal peptide synthethase AnaPS from N. fischeri. It could be speculated that an epimerase domain in AnaPS from a producer of epi-aszonalenins (Fig. 1) would not be necessary. However, no epimerase domain could be detected in AnaPS from N. fischeri and A. terreus.

From the results obtained by feeding experiments, it was speculated that the reverse prenyl moiety at position C3 of the indole ring in roquefortine could be a result of the rearrangement of a derivative with a regular prenyl moiety at position N1 (1,22,23). A diketopiperazine derivative with a reverse prenyl moiety at position C2 could be involved in this rearrangement. In this study, we detected only one dominant enzymatic product after incubation of (R)-benzodiazepinedione in the presence of AnaPT and DMAPP (Fig. 3C and Fig. 5A), i.e. aszonalenin. It seems, therefore, the prenylation at C3 position of tryptophan-containing dipeptides with a diketopiperazine ring was very likely directly catalysed by prenyltransferases such as AnaPT. The ring formation between C2 of the indole...
ring and an N-atom of the diketopiperazine ring took very likely during the prenylation, in which a carbonium ion could be involved (Fig. 7). According to this hypothesis, H-2 of L-tryptophan would be retained in roquefortine or aszonalenin. This hypothesis was supported by the observation that feeding of L-[2,4,5,6,7-2H5]-tryptophan in *Penicillium roqueforti* and *Aspergillus zonatus* resulted in formation of roquefortine and aszonalenin with significant incorporation of deuterium at all of the five positions (20,21). This indicated that the H-2 was retained during the prenylation reaction.

AnaPT catalysed the transfer of a dimethylallyl moiety to C3 of the indole ring of a cyclic dipeptide and meanwhile the formation of a ring system between the indole and the diketopiperazine ring. Although AnaPT catalyses two different reaction steps, it has a comparable size in the range of 430 to 470 amino acids to two different reaction steps, it has a comparable sequence similarities (see Results).

AnaAT catalysed the N-acetylation of aszonalenin. Sequence analysis showed that AnaAT shares low sequence similarity with putative trichothecone-3-O-acetyltransferases from different genome sequences such as *Aspergillus* and *Fusarium* strains. To best of our knowledge, no functional proof of such genes was reported previously.

Our previous studies showed that indole prenyltransferases from *Aspergillus fumigatus* have broad flexibility towards the aromatic substrates. A large number of simple tryptophan derivatives were accepted by the dimethylallyltryptophan synthases FgaPT2 and 7-DMATS and converted to 4- or 7-prenylated derivatives, respectively (38,39). All of the tryptophan-containing cyclic dipeptides were converted by CdpNPT and FtmPT1 to N1 prenylated or C2 prenylated derivatives, respectively (36,37). It would be interesting to test the substrate specificity of AnaPT towards cyclic dipeptides.

REFERENCE


FOOTNOTES
This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SPP1152 to S.-M. Li). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The abbreviations used are DMA: dimethylallyl; DMAPP: dimethylallyl diphosphate

FIGURE LEGENDS
Figure 1: Examples of at position C3 reversely prenylated indole alkaloids from fungi.
Figure 2: HPLC analysis of extracts from Neosartorya fischeri. The substances were detected with a Photo Diode Array detector and illustrated at 254 nm.
Figure 3: Putative ana gene cluster from Neosartorya fischeri (A) and Aspergillus terreus (B) and hypothetical biosynthetic pathway for acetylaszonalenin (C).
Figure 4: Analysis of the overproduction and purification of Hisα-AnaPT and Hisα-AnaAT. The proteins were separated on a 12 % SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue G-250. Lanes 1 and 5: Molecular mass standard; 2: Total protein for AnaPT overproduction before induction; 3: Total protein for AnaPT overproduction after induction; 4: Purified Hisα-AnaPT; 6: Total protein for AnaAT overproduction before induction; 7: Total protein for AnaAT overproduction after induction; 8: Purified Hisα-AnaAT.
Figure 5: HPLC chromatograms of enzyme assays with Hisα-AnaPT and Hisα-AnaAT. Incubation of (R)-benzodiazepinedione and DMAPP with active Hisα-AnaPT (A); Incubation of (R)-benzodiazepinedione and DMAPP with heat-inactivated Hisα-AnaPT by boiling of the enzyme solution for 20 min (B); Incubation of aszonalenin and acetylcoenzyme A with active Hisα-AnaAT (C); Incubation of aszonalenin and acetylcoenzyme A with heat-inactivated Hisα-AnaAT (D). The
substances were detected with a Photo Diode Array detector and illustrated at 254 nm.

Figure 6: Determination of kinetic parameters of AnaPT for (R)-benzodiazepinedione (A) and DMAPP (B) as well as of AnaAT for aszonalenin (C) and acetylcoenzyme A (D). For conditions please see the description in EXPERIMENTAL PROCEDURES.

Figure 7: Hypothetical reaction mechanism of aszonalenin formation catalysed by AnaPT
### Table 1: $^1$H-NMR and $^{13}$C-NMR data of the reported compounds

<table>
<thead>
<tr>
<th>Structure</th>
<th>Solvent</th>
<th>CDCl$_3$</th>
<th>CD$_3$OD</th>
<th>CDCl$_3$</th>
<th>CDCl$_3$ (from N. fischeri)</th>
<th>CDCl$_3$ (enzymatic product)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R)-benzodiazepinedione</td>
<td>NH-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>H-2</td>
<td>7.19, d, 2.5</td>
<td>7.13, br s</td>
<td>124.82</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>H-4</td>
<td>7.53, dd, 7.9, 1.0</td>
<td>7.42, d, 7.9</td>
<td>119.83</td>
<td>5.58, s</td>
<td>81.7</td>
</tr>
<tr>
<td></td>
<td>H-5</td>
<td>7.11, td, 8.2, 1.0</td>
<td>6.94, t, 7.6</td>
<td>118.84</td>
<td>-</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td>H-6</td>
<td>7.22, td, 7.6, 1.0</td>
<td>7.05, t, 7.9</td>
<td>122.29</td>
<td>-</td>
<td>7.19, t, 7.3</td>
</tr>
<tr>
<td></td>
<td>H-7</td>
<td>7.38, d, 7.9</td>
<td>7.30, d, 7.2</td>
<td>112.34</td>
<td>6.62, d, 7.9</td>
<td>7.71, d, 7.6</td>
</tr>
<tr>
<td></td>
<td>C-9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C-10</td>
<td>3.53, dd, 15.2, 5.6</td>
<td>3.38, dd, 15.1, 6.0</td>
<td>24.95</td>
<td>3.46, dd, 14.2, 7.6</td>
<td>33.4</td>
</tr>
<tr>
<td></td>
<td>C-11</td>
<td>4.15, dt, 8.2, 8.4</td>
<td>4.09, dd, 8.8, 5.7</td>
<td>54.58</td>
<td>3.99, brt, 8.1</td>
<td>57.0</td>
</tr>
<tr>
<td></td>
<td>NH-12</td>
<td>6.00, d, 4.4</td>
<td>-</td>
<td>170.97</td>
<td>-</td>
<td>166.8</td>
</tr>
<tr>
<td></td>
<td>NH-16</td>
<td>7.63, s</td>
<td>-</td>
<td>128.37</td>
<td>7.91, s</td>
<td>126.8</td>
</tr>
<tr>
<td></td>
<td>C-15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>138.08</td>
<td>149.0</td>
</tr>
<tr>
<td></td>
<td>C-17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>173.72</td>
<td>169.8</td>
</tr>
<tr>
<td></td>
<td>H-18</td>
<td>6.96, dd, 7.6, 1.0</td>
<td>7.12, d, 8.0</td>
<td>122.47</td>
<td>6.90, d, 8.2</td>
<td>120.4</td>
</tr>
<tr>
<td></td>
<td>H-19</td>
<td>7.50, td, 7.9, 1.3</td>
<td>7.53, dd, 8.0, 1.6</td>
<td>134.15</td>
<td>7.44, td, 7.9, 1.0</td>
<td>132.6</td>
</tr>
<tr>
<td></td>
<td>H-20</td>
<td>7.27, td, 7.6, 1.3</td>
<td>7.24, td, 7.7, 1.0</td>
<td>125.82</td>
<td>7.22, t, 7.6</td>
<td>124.9</td>
</tr>
<tr>
<td></td>
<td>H-21</td>
<td>7.92, dd, 7.9, 1.3</td>
<td>7.78, dd, 7.9, 1.3</td>
<td>131.67</td>
<td>7.83, d, 7.9</td>
<td>131.2</td>
</tr>
<tr>
<td></td>
<td>H-‘1’</td>
<td>5.14, d, 10.7</td>
<td>-</td>
<td>114.3</td>
<td>5.14, d, 10.7</td>
<td>114.3</td>
</tr>
<tr>
<td></td>
<td>H-‘2’</td>
<td>6.11, dd, 17.3, 10.7</td>
<td>-</td>
<td>143.7</td>
<td>5.89, dd, 17.4, 10.4</td>
<td>5.89, dd, 17.4, 10.4</td>
</tr>
<tr>
<td></td>
<td>3H-4’</td>
<td>1.06, s</td>
<td>-</td>
<td>22.7</td>
<td>0.99, s</td>
<td>0.99, s</td>
</tr>
<tr>
<td></td>
<td>3H-5’</td>
<td>1.14, s</td>
<td>-</td>
<td>22.5</td>
<td>1.19, s</td>
<td>1.19, s</td>
</tr>
<tr>
<td></td>
<td>COCH$_3$</td>
<td>-</td>
<td>-</td>
<td>41.4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Chemical shifts are given in p.p.m. and coupling constants in Hz.
Table 2 ORFs of the putative *ana* cluster in *N. fischeri* NRRL181 and *A. terreus* NIH2624

<table>
<thead>
<tr>
<th>Gene</th>
<th><em>N. fischeri</em> NRRL181</th>
<th><em>A. terreus</em> NIH2624</th>
<th>Identity of gene products</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ORF</td>
<td>Exons (bp)</td>
<td>Introns (bp)</td>
<td>Protein (aa)</td>
</tr>
<tr>
<td>anaPS</td>
<td>NFIA_055290</td>
<td>6630/450</td>
<td>52</td>
<td>EAW16180 (2359)</td>
</tr>
<tr>
<td>anaPT</td>
<td>NFIA_055300</td>
<td>1187/127</td>
<td>51</td>
<td>EAW16181 (437)</td>
</tr>
<tr>
<td>anaAT</td>
<td>NFIA_055310</td>
<td>1530</td>
<td>-</td>
<td>EAW16182 (509)</td>
</tr>
</tbody>
</table>
Figure 1

- acetylaszonalenin
- epi-aszonalenin A
- amuromine
- epiamauromine
- 5-N-acetylardeemin
- roquefortine C
- fructigenine A (rugulosuvine B)
- fructigenine B (verrucofortine)
Figure 2

[mAU]

acetylaszonalenin

[min]
Figure 5

A

(B)-benzodiazepinedione

[mAU]

[min]

B

(B)-benzodiazepinedione

[mAU]

[min]

C

aszonalenin

[mAU]

[min]

D

aszonalenin

[mAU]

[min]
Figure 6

A

K_m = 232 μM
k_cat = 2.50 s⁻¹

B

K_m = 156 μM
k_cat = 0.64 s⁻¹

C

K_m = 61 μM
k_cat = 0.21 s⁻¹

D

K_m = 96 μM
k_cat = 0.06 s⁻¹
Acetylaszonalenin biosynthesis in Neosartorya fischeri: Identification of the biosynthetic gene cluster by genomic mining and functional proof of the genes by biochemical investigation

Wen-Bing Yin, Alexander Grundmann, Jun Cheng and Shu-Ming Li

J. Biol. Chem. published online November 10, 2008

Access the most updated version of this article at doi: 10.1074/jbc.M807606200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2008/11/10/jbc.M807606200.citation.full.html#ref-list-1