Characterization of the Azinomycin B Biosynthetic Gene Cluster Revealing a Different Iterative Type I Polyketide Synthase for Naphthoate Biosynthesis

Qunfei Zhao,1 Qingli He,1 Wei Ding,1,2 Mancheng Tang,1 Qianjin Kang,1,3 Yi Yu,1,3 Wei Deng,1 Qi Zhang,1 Jie Fang,1 Gongli Tang,1 and Wen Liu1,3,*

1State Key Laboratory of Bioorganic and Natural Product Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 354 Fenglin Road, Shanghai 200032, China
2School of Life Science, Lanzhou University, 222 South Tianshui Road, Lanzhou 730000, China
3Laboratory of Microbial Metabolism and School of Life Science & Biotechnology, Shanghai Jiaotong University, 154 Huashan Road, Shanghai 200030, China
*Correspondence: wliu@mail.sioc.ac.cn
DOI 10.1016/j.chembiol.2008.05.021

SUMMARY

Azinomycin B is a complex natural product containing densely assembled functionalities with potent antitumor activity. Cloning and sequence analysis of the aza gene cluster revealed an iterative type I polyketide synthase (PKS) gene, five nonribosomal peptide synthetases (NRPSs) genes and numerous genes encoding the biosynthesis of unusual building blocks and tailoring steps for azinomycin B production. Characterization of AziB as a 5-methyl-naphthoic acid (NPA) synthase showed a distinct selective reduction pattern in aromatic polyketide biosynthesis governed by bacterial iterative type I PKSs. Heterologous expression established the PKS-post modification route from 5-methyl-NPA to reach the first building block 3-methoxy-5-methyl-NPA. This proposed azinomycin B biosynthetic pathway sets the stage to investigate the enzymatic mechanisms for building structurally unique and pharmaceutically important groups, including the unprecedented azabicyclic ring system and highly active epoxide moiety.

INTRODUCTION

Azinomycin B (Figure 1), namely, carzinophilin A, originally isolated from Streptomyces sahachiroi (Hata et al., 1954), was subsequently reisolated from S. griseofuscus along with its naturally occurring analog azinomycin A (Yokoi et al., 1986). Azinomycin B contains a set of unusual functionalized moieties that are densely assembled in the architecture, including a naphthoate building block, a 2-amino-1,3-dicarbonyl group, an epoxide moiety, and an unprecedented aziridino[1, 2a]pyrrolidine (1-azabicyclo[3.1.0]hexane) ring system. As a consequence of these structural features, azinomycin B binds within the major groove of DNA and forms covalent interstrand crosslinks (ISCs) by electrophilic attack of C10 and C21 onto N7 positions of suitably disposed purine bases with apparent sequence selectivity (Armstrong et al., 1992; Hartley et al., 2000; Coleman et al., 2002; LePla et al., 2005). Recently, in vivo study on DNA damage in yeast supported the mode of ISC action of azinomycin B (Kelly et al., 2006), suggesting that the resulting DNA alkylation and crosslinking are biologically relevant to its potent antitumor activity.

Initial cytotoxicity studies and early phase clinical investigation on P388 leukemic mice showed that azinomycin B is highly active against leukemic cell line L5178Y at the nanomolar level and provides an increased lifespan (ILS) comparable to the clinically used DNA ISC drug mitomycin C (but with lower dosage) (Shimada et al., 1955; Ishizeki et al., 1987). Although poor availability and extreme instability prevented detailed screening of the scope of the biological activities and development in the clinical trial as the chemotherapeutic agents, extensive efforts on evaluation of azinomycin analogs have been performed by chemical synthesis. The total synthesis of azinomycin A was achieved in 2001 (Coleman et al., 2001), showing a drawback that the C12 hydroxyl group destabilizes the compound during isolation. A number of bioactive analogs that are structurally more stable than the natural products have also been synthesized (Alicar et al., 2005; Casely-Hayford et al., 2005; David-Cordonnier et al., 2006; Coleman et al., 2007), providing a structure-biological activity relationship to study for further understanding of the mechanisms of the natural products and pharmaceutical development by structural derivation.

Additionally, the unique structure, remarkable antitumor activity, and novel mode of action have attracted increased research interests toward the biosynthesis of azinomycins. The densely assembled functionalities on azinomycins offer a distinct opportunity to study the biosynthetic strategy for their unprecedented molecular scaffolds. However, the progress has been largely impeded by difficulties in culturing and securing a constant source of natural products. Recently, a methodology involving a nutrient-starved fermentation was developed and led to remarkable improvement of the azinomycin B production (Kelly et al., 2008), significantly facilitating the biosynthetic investigation in vivo. Previous isotope-labeled feeding experiments (Corre and Lowden, 2004; Corre et al., 2006; Liu et al., 2006), which were carried out either in fermentation or in a developed cell-free system of the producer S. sahachiroi, clearly revealed a polyketide origin of the naphthoate moiety and suggested the
incorporation of amino acids Valine (Val), glutamic acid (Glu) derivative, and Threonine (Thr) into the tripeptidyl backbone, indicating that a polyketide synthase (PKS)/nonribosomal peptide synthetase (NRPS) hybrid system might be involved in the biosynthesis of azinomycin B.

PKSs that are structurally type I but act in an iterative manner for aromatic polyketide biosynthesis are widely found in fungi (Shen, 2000). In contrast, the similar PKSs in bacteria comprises only five members to date (Figures 1 and 6): AviM (Gaisser et al., 1997) and CalO5 (Ahlert et al., 2002) catalyze the biosynthesis of monocyclic orsellinic acid (OSA) for avilamycin (AVI) and calicheamicin (CAL), respectively; ChlB1 (Jia et al., 2006) and MdpB (Van Lanen et al., 2007) catalyze the biosynthesis of 6-methylsalicylic acid (6-MSA) for chlorothricin (CHL) and maduropeptin (MDP), respectively; and NcsB (Sthapit et al., 2004; Liu et al., 2005) is the only representative that catalyzes the biosynthesis of bicyclic 2-hydroxyl-5-methyl-naphthoic acid (NPA) for neocarzinostatin (NCS). According to the structural similarity and polyketide origin revealed by isotope-labeled experiments, the biosynthesis of the naphthoate moiety of azinomycin B likely involves an iterative type I PKS.

Presently, we report access to the biosynthetic machinery of the azinomycin B in S. sahachiroi NRRL 2485 by cloning the iterative type I PKS gene aziB using our recently developed rapid PCR approach (Shao et al., 2006). The sequencing of the entire azi gene cluster allows for assignment of the functions to deduced gene products, providing insights into the biosynthetic pathway. A number of novel genes were found in this cluster to encode the highly diverse functionalities of unusual building blocks, which might be sequentially assembled to afford azinomycin B by an unusual NRPS system and tailoring enzymes. Heterologous expression of aziB in S. albus characterized an iterative type I PKS AziB for the biosynthesis of a bicyclic aromatic polyketide, 5-methyl-NPA, in a manner similar to previously found 2-hydroxyl-5-methyl-NPA synthase NcsB, but with distinct selective reduction pattern. Addition of aziB1 (encoding a P450 hydroxylase) and aziB2 (encoding an O-methyltransferase) to aziB in S. albus allowed for the production of 3-methoxy-5-methyl-NPA as the first building block for skeleton assembly of azinomycin B. The availability of the azi gene cluster and proposed biosynthetic pathway will pave the way for further studies regarding the unusual biochemistry and therefore inspire attempts to apply this knowledge for combinatorial biosynthesis.

RESULTS

Cloning and Sequencing of the azi Gene Cluster from S. sahachiroi NRRL 2485

Iterative type I PKSs for aromatic polyketide biosynthesis share high homology in both amino acid sequence and domain organization and are phylogenetically distinct from other groups of type I PKSs (Jenke-Kodama et al., 2008). To specifically probe the
gene encoding the naphthoate moiety of azinomycin B, we adopted a rapid PCR approach that was developed to access the 6-MSA synthase gene chiB1 according to the conserved motifs that are only unique in the iterative type I PKS family (Shao et al., 2006). With the genomic DNA of S. sahachiroi as the template, a distinct product with the expected size of 0.9 kb was readily amplified and subsequently cloned into pSP72 vector for sequencing. Analysis of randomly selected clones confirmed that most of them contained an identical PCR product P1, the deduced amino acid sequence of which is highly homologous to the known iterative type I PKSs. To determine the relevance of P1 to azinomycin B biosynthesis, we set out to inactivate the target alleles in S. sahachiroi. As expected, the P1 allele mutant strain AL1001 completely lost the ability to produce azinomycin B (Figures 4A2 and 4BII), confirming that this P1-containing iterative type I PKS gene is involved in the azinomycin B biosynthesis. Consequently, using the P1 product as a probe, approximately 6,000 clones of the S. sahachiroi genomic library were screened, resulting in 12 overlapping cosmids that span a 50 kb DNA region. To ensure full coverage of the entire azi gene cluster, additional chromosome walking from the left end of cosmid pAL1022 was carried out by using the 3.2 kb BamHI fragment P2 as a probe, eventually leading to a 80 kb contiguous DNA region on the chromosome, as exemplified by the inserts of cosmids pAL1022, pAL1023, pAL1024, pAL1025, and pAL1026 (Figure 2A).

The DNA region represented by cosmids pAL1023 and pAL1024 was selected for sequencing, yielding a 63,549 bp contiguous sequence with 71.48% of the overall GC content, characteristic of Streptomyces DNA. Bioinformatic analysis of the sequenced region revealed 47 ORFs (Figure 2B and Table 1), 34 of which, from aziA3 to aziD2, were proposed to constitute the azi gene cluster according to functional assignment of their deduced products. While an approximate 2.0 kb fragment, within which no apparent ORFs were found, physically separated the azi gene cluster from orf(−1) that encodes a putative sporulation protein at the left end, the relatively uncertain right-side boundary was judged by identification of orf4 encoding a putative transposase and a set of hypothetical genes (e.g. orf1−3) that lack significant similarity to those involved in secondary metabolite biosynthesis.

Genes Encoding the Biosynthesis of Naphthoate Moiety
aziB is the only typical PKS gene identified within the azi gene cluster, supporting its role in the naphthoate moiety biosynthesis as outlined in Figure 3A. The deduced product AziB consists of characteristic domains of type I PKSs, including a ketosynthase (KS), acyltransferase (AT), dehydratase (DH), ketoeductase (KR), and acyl carrier protein (ACP). It exhibits head-to-tail homology to the known iterative type I PKSs in bacteria (Figure 6), including NcsB, ChlB1 and MadB, and AviM and CalO5 (except the KR domain). In a mechanistic analogy, AziB could be envisioned to catalyze the assembly of a nascent linear hexaketide from one acetyl-CoA and five malonyl-CoAs by iterative decarboxylative condensation as NcsB; however, the KR domain in AziB may offer additional selective reductions at C3, C5, and C9 positions (Figure 6IV), which is apparently distinct from the actions of the KR domain of NcsB occurring at C5 and C9 positions and the KR domains of ChlB1 and MadB at C5 position only. The resulting hexaketide intermediate may then undergo two intramolecular aldol condensations (between C2–C7 and C6–C11, respectively) and dehydrations to produce the relatively hydrophobic aromatic compound 5-methyl-NPA.

To reach 3-methoxy-5-methyl-NPA (1) as the first building block for assembly of the azinomycin B skeleton, two genes, aziB1 and aziB2, were identified within the azi gene cluster and proposed to encode the PKS-post modifications on 5-methyl-NPA (Figure 3A). AziB1, with high sequence similarity to various bacterial P450 hydroxylases, might be responsible for introduction of a hydroxyl group at the C3 position, followed by methylation with the action of AziB2 that resembles a family of SAM-dependent O-methyltransferases, to render a methoxy functionality.

Genes Encoding the Biosynthesis of Unusual Building Blocks Derived from Amino Acids
aziC1 encodes a protein belonging to a family of branched chain amino acid aminotransferases, which catalyze the oxidative desamination of branched amino acids to form α-keto acids. Presumably, AziC1 might be responsible for the biosynthesis of α-ketoisovaleric acid (2: Figure 3B) that serves as the second building block to be incorporated into the peptidyl backbone of azinomycin B by NRPS AziA3 (described below).

At least ten genes, aziC2−C11, are assumed to be involved in the biosynthesis of the third building block, aziridinoo[1, 2α]pyrrolidiny amino acid (3) as shown in Figure 3C. AziC2 shares high sequence similarity to LysX that catalyzes the N-acetylation of 2-aminoacipate in the bacterial lysine (Lys) biosynthetic pathway, suggesting that the first step toward 3 starts with Glu by protection of the α-amino group. AziC3 and AziC4, respectively similar to various N-acetylglutamate kinases and phosphate reductases in the ornithine (Orn) biosynthetic pathways, may activate the γ-carboxyl group of N-acetylglutamic acid by phosphorylation and then reduce it to form N-acetylglutamyl-γ-semialdehyde. AziC5 and AziC6, both of which, respectively, share high homologies to the N- and C-terminal subunits of transketolases, could constitute a protein complex with thiamine pyrophosphate (TPP) as the cofactor for two-carbon unit extension on the C5 position. Subsequently, AziC7, resembling members of a pyridoxal-phosphate-dependent aminotransferase family, is likely responsible for transamination from Glu onto the 6-keto group, fulfilling all atoms requirement for the azabicyclic skeleton formation to reach 3.

To validate the role of AziC7 in the azinomycin B biosynthesis, we inactivated aziC7 by replacing it with an erythromycin resistance gene. As shown in Figure 4, the resultant mutant AL1003 completely lost the ability of producing azinomycin B. Further HPLC-MS analysis revealed an intermediate exhibiting an [M − H]− ion at m/z = 215.09, consistent with the molecular formula C13H12O3. With a standard as the control, we confirmed its identity to 3-methoxy-5-methyl-NPA (Figure S3IV), which was also accumulated in the wild-type strain fermentation with a higher yield (partially owing to the decomposition of azinomycin B, Figure 5).
candidates located in the azi gene cluster, it would be reasonable that the terminal hydroxyl can be activated (e.g., phosphorylation) by forming a good leaving group for the closure of the second aziridine ring. AziC9 resembles a P450 monooxygenase family, supporting its action at the C4 position for introduction of a hydroxyl group. Finally, AziC10, which belongs to a deacylase family, is assumed to catalyze the deacetylation reaction and release 3 as a key building block for assembly of the azinomycin B skeleton.

Noticeably, AziC11 exhibits high sequence homology to various discrete PCPs, raising the question of whether enzymatic reactions toward the formation of 3 occur on the free amino acid or acyl-PCP.

**Genes Encoding NRPSs for Assembly of the Azinomycin B Skeleton**

Five modular NRPS genes, aziA1–A5, were found within the azi gene cluster. As shown in Figure 3D, their deduced products constitute an atypical NRPS system. AziA1 is a bifunctional enzyme with two catalytic activities involving an acyl-CoA ligase (AL) and peptidyl carrier protein (PCP), presumably responsible for activating 3-methoxy-5-methyl-NPA 1. The naphthoate
moiety could then be transferred to AziA2, domains in which are arranged as condensation (C) domain -PCP-C, for initiation of the azinomycin B skeleton assembly. AziA3 shares high sequence homology to the N-terminal modules of CesA and CesB with an AL-ketoreductase (KR)-PCP organization in each (Magarvey et al., 2006), which were characterized to govern the α-keto acid activation, tethering, and reduction in the cereulide biosynthesis. Similarly, AziA3 could activate the α-ketosiovaleric acid 2 and transfer it onto the PCP domain, and the internal KR domain may perform the ketoreduction to form α-hydroxyisovaleryl moiety. Although structurally AzI4 is similar to typical NRPS modules composed of C-A-PCP domains, the A domain of AziA4 should be novel due to the recognition of the aziridino[1, 2a]pyrrrolidinyl amino acid 3 as the substrate. The last NRPS AziA5 is the only tetradomain protein in the azinomycin B assembly system, containing an additional reductase (RE) domain appended to C-A-PCP at the C terminus. Similar RE domains of NRPSs have been shown to release the resultant peptidyl intermediates as reductive products instead of the more commonly found thioesters. Although structurally AziA5-A resembles a family of proteins that were widely found in secondary metabolism and annotated as transmembrane transporters, such as Hyp19 in hygromycin biosynthesis (Palaniappan et al., 2006), AziE may function as a resistant protein to carry the synthesized product out of the cells.

To predict the substrates of the azinomycin B NRPS system, the eight specificity-conferring codes for each A or AL domain (except the AziA3-AL, which apparently lacks these codes) were identified as follows by sequence alignment with the A domain of PheA (Stachelhaus et al., 1999; Challis et al., 2000); GHTTGSOIK for AziA1-AL; DVFDFGGVCK for AziA4-A; and DFWSVGVMHK for AziA5-A. While all available models showed negative results for AziA1-AL and AziA4-A (consistent with an unpaired substrate incorporated in each described above), AziA5-A was supposed to recognize and activate Thr, which may serve as the fourth building block (4) for assembly. As a consequence, these azinomycin B NRPSs are hypothesized to sequentially condense building blocks 1, 2, 3, and 4, involving formation of one ester bond and two amide bonds, and finally release the putative key intermediate as an aldehyde, pre-azinomycin B, as shown in Figure 3D.

To support the NRPS function, we chose AziA3, which resides in the left end of the azi gene cluster, for inactivation. As expected, the resulting mutant AL1002 failed to produce azinomycin B (Figures 4A3 and 4BIII), confirming its need for azinomycin B biosynthesis. Surprisingly, several attempts failed to detect the intermediates such as 3-methoxy-5-methyl-NPA accumulated in the mutant AL1002 fermentation (Figure S3V).

Three discrete TE genes, aziA6, -A7, and -A8, were identified within the azi gene cluster. aziA6 and A7 are genetically linked with the NRPS genes aziA3 and aziA4 (in particular, aziA6 is likely to be translationally coupled with these NRPS genes in a 12 kb large operon as judged by overlapping of their start and stop codons, shown in Figure 2B), supporting their functions associated with the peptide assembly of azinomycin B. Although their detailed roles remain to be determined, we suggest that AziA6, A7 and A8 may be involved in (1) improvement of the efficiency of product formation by regenerating misacylated NRPSs (Schwarzer et al., 2002), and (2) physiological competition with the action of AziA5-RE domain to release the intermediate as an acid, on which further modifications including decarboxylation at the C3 position might be carried out to produce azinomycin A.

Genes Encoding Tailoring Enzymes

To eventually produce azinomycin B, post-modifications (including acetylation and oxidations) are postulated to proceed with a set of tailoring enzymes encoded by aziD1, D2, and D3 as outlined in Figure 3D. AziD1, highly homologous to a family of O-acetyltransferases, may be responsible for acetylation of the hydroxyl group at C13 position. AziD2, sharing high sequence similarities to an unknown family of acyl-CoA dehydrogenases, might be a candidate for oxidation of the hydroxyl group to render the 2-amino-1,3-dicarbonyl group. In contrast, AziD3, which resembles members of the mitochondrial acyl-CoA dehydrogenase family (using FAD as the cofactor), is likely to be regiospecifically involved in the dehydrogenation at C20-21 to form a double bond, which could be further oxidized to form the highly active epoxide fragment in this molecule.

Genes Encoding Resistance and Other Functions

AziE resembles a family of proteins that were widely found in secondary metabolism and annotated as transmembrane transporters, such as Hyp19 in hygromycin biosynthesis (Palaniappan et al., 2006). AziE may function as a resistant protein to carry the synthesized product out of the cells.

Five genes within the azi gene cluster encode proteins whose functions cannot be assigned in the proposed azinomycin B biosynthetic pathway. AziF resembles SimX2-like proteins of unknown functions in a few antibiotic biosynthetic pathways, such as ChlH (46% identity) in CHL biosynthesis (Jia et al., 2006). AziG, belonging to a thioesterase superfamily, shares a hotdog fold region with members that have related catalytic activities, including PaaD (28% identity) involved in the benzoate degradation (Di Gennaro et al., 2007). A subset of genes, aziH1, aziH2, and aziH3, encode proteins that share high similarities to enzymes regarding the first two reaction steps of sulfur metabolism (e.g., CysD, CysN, and CysC in E. coli; Leyh et al., 1988).

Finally, three downstream genes of aziB, aziU1–U3, which are likely to be translationally coupled as judged by their overlapped start and stop codons, could not be predicted upon sequence analysis alone. Their involvement in the azinomycin B biosynthesis remains to be determined.

Establishment of the Naphthoate Pathway by Heterologous Overexpression

The involvement of AziB in azinomycin biosynthesis was confirmed by gene inactivation. To verify its function as a novel naphthoate synthase, a construct that carries aziB alone under the control of the constitutive promoter Perme* was introduced into the heterologous host S. albus, yielding the recombinant strain AL1004. With the S. albus strain AL1006 that carries the vector pTG-2 as a negative control, AL1004 was cultured and analyzed by HPLC (Figure 5II), revealing a distinct peak with a retention time at 19.6 min. To elucidate the structure, the resultant compound was extracted and purified. EI-HRMS analysis of the obtained powder exhibited an M+ ion at m/z = 186.0675, consistent with the molecular formula C12H10O2. In addition, 1H NMR spectrum analysis supported that this compound is 5-methyl-NPA. This finding unambiguously confirmed AziB as an iterative type 1 PKS for the biosynthesis of the bicyclic aromatic compound 5-methyl-NPA, other than NcsB for the biosynthesis of 2-hydroxyl-5-methyl-NPA.
### Table 1. Deduced Functions of orfs in the Azinomycin B Biosynthetic Gene Cluster

<table>
<thead>
<tr>
<th>Gene</th>
<th>Size&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Protein Homolog&lt;sup&gt;b&lt;/sup&gt; and Origin</th>
<th>Similarity/Identity, %</th>
<th>Proposed Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>orf(-1)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>143</td>
<td>SpoIIIE (CAF50935), from S. ambofaciens ATCC 23877</td>
<td>75/62</td>
<td>Stage II sporulation protein E (SpoIIIE)</td>
</tr>
<tr>
<td>aziA3</td>
<td>1263</td>
<td>CesA (ABK00751), from Bacillus cereus</td>
<td>60/41</td>
<td>NRPS (AL-KR-PCP)</td>
</tr>
<tr>
<td>aziA6</td>
<td>273</td>
<td>GrsT (YP_001106480), from Saccharopolyspora erythraea NRRL 2338</td>
<td>57/41</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>aziA4</td>
<td>1103</td>
<td>EndB (ABD65957), from S. fungicidicus</td>
<td>50/38</td>
<td>NRPS (C-A-PCP)</td>
</tr>
<tr>
<td>aziA7</td>
<td>250</td>
<td>TeLB (AAT45287), from S. tuberculosis</td>
<td>65/52</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>aziB2</td>
<td>345</td>
<td>Strop_0204 (YP_001157067), from Salinispora tropicala CNB-440</td>
<td>64/50</td>
<td>O-methyltransferase</td>
</tr>
<tr>
<td>aziB1</td>
<td>401</td>
<td>Haur_3696 (YP_001546460), from Herpetosiphon aurantiacus ATCC 23779</td>
<td>48/31</td>
<td>Cytochrome P450 hydroxylase</td>
</tr>
<tr>
<td>aziA1</td>
<td>627</td>
<td>NosC (AAF17280), from Nostoc sp. GSV224</td>
<td>53/37</td>
<td>NRPS (AL-PCP)</td>
</tr>
<tr>
<td>aziE</td>
<td>466</td>
<td>Hyp19 (ABC42556), from S. hygroscopicus</td>
<td>30/50</td>
<td>Export protein</td>
</tr>
<tr>
<td>aziA2</td>
<td>1115</td>
<td>NosD (AAF17281), from Nostoc sp. GSV224</td>
<td>45/30</td>
<td>NRPS (C-PCP-C)</td>
</tr>
<tr>
<td>aziH2</td>
<td>435</td>
<td>CyaN (NP_417231), from E. coli</td>
<td>47/63</td>
<td>Sulfate adenylyltransferase subunit 1</td>
</tr>
<tr>
<td>aziH1</td>
<td>270</td>
<td>CyaD (NP_417232), from E. coli</td>
<td>46/63</td>
<td>Sulfate adenylyltransferase subunit 2</td>
</tr>
<tr>
<td>aziH3</td>
<td>185</td>
<td>CysC (NP_417230), from E. coli</td>
<td>41/59</td>
<td>Adenylylsulfate kinase</td>
</tr>
<tr>
<td>aziG</td>
<td>133</td>
<td>Strop_1904 (YP_001158744), from S. tropicala CNB-440</td>
<td>66/55</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>aziF</td>
<td>62</td>
<td>SimX2 (AAK06794), from S. antibioticus</td>
<td>55/47</td>
<td>SimX2-like protein</td>
</tr>
<tr>
<td>aziD1</td>
<td>403</td>
<td>Smr6&lt;sup&gt;6&lt;/sup&gt; (CAM96572), from S. ambofaciens</td>
<td>56/42</td>
<td>O-acyltransferase</td>
</tr>
<tr>
<td>aziC11</td>
<td>87</td>
<td>PPA1287 (YP_055995), from Propionibacterium acnes KPA171202</td>
<td>58/36</td>
<td>PCP/ACP</td>
</tr>
<tr>
<td>aziC6</td>
<td>340</td>
<td>TktC (NP_561213), from Clostridium perfringens str. 13</td>
<td>61/41</td>
<td>Transketolase C-terminal subunit</td>
</tr>
<tr>
<td>aziC5</td>
<td>311</td>
<td>TM0954 (NP_228762), from Thermotoga maritima MSB8</td>
<td>61/44</td>
<td>Transketolase N-terminal subunit</td>
</tr>
<tr>
<td>aziC7</td>
<td>440</td>
<td>OleN2 (AD55458), from S. antibioticus</td>
<td>53/43</td>
<td>Aminotransferase</td>
</tr>
<tr>
<td>aziC4</td>
<td>307</td>
<td>ArgC (YP_604156), from Deinococcus geothermalis DSM 11300</td>
<td>55/45</td>
<td>N-acetyl-γ-glutamyl-phosphate reductase</td>
</tr>
<tr>
<td>aziC3</td>
<td>292</td>
<td>TTC1541 (YP_005510), from Thermus thermophilus HB27</td>
<td>55/36</td>
<td>N-acetylglutamate kinase</td>
</tr>
<tr>
<td>aziC10</td>
<td>376</td>
<td>Caur_1204 (YP_001634823), from Chloroflexus aurantiacus J-10-fl</td>
<td>54/40</td>
<td>N-acetylornithine/N-acetyl-lysine deacetylase</td>
</tr>
<tr>
<td>aziC8</td>
<td>272</td>
<td>ORF39 (ABD65959), from S. fungicidicus</td>
<td>63/47</td>
<td>Acyl-CoA dehydrogenase/reductase</td>
</tr>
<tr>
<td>aziA5</td>
<td>1541</td>
<td>SfmC (ABI22133), from S. lavendulae</td>
<td>30/42</td>
<td>NRPS (C-A-PCP-RE)</td>
</tr>
<tr>
<td>aziB</td>
<td>1779</td>
<td>NcsB (AAM77986), from S. carzinostaticus subsp. neocarzinostaticus</td>
<td>66/52</td>
<td>Iterative type I PKS (KS-DH-KR-ACP)</td>
</tr>
</tbody>
</table>
Further, the constructs harboring mutant aziB, in which site-specific mutations were performed for inactivating the KR and DH domains, respectively, were introduced into S. albus, yielding the recombinant strain AL1007 (to express the KR mutant AziB, G1398A within the NADPH-binding motif GxGxxG), AL1008 (to express the KR mutant AziB, Y1549F at the conserved active site), and AL1009 (to express the DH mutant AziB, H935F within the conserved motif HxxxGxxxxP). Upon HPLC-MS analysis (Figure 5), AL1007, AL1008, and AL1009 failed to produce 5-methyl-NPA, confirming that the reductive and dehydrating actions governed by the KR and DH domains of AziB are essential for the 5-methyl-NPA formation.

Table 1. Continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Size a</th>
<th>Protein Homolog b and Origin</th>
<th>Similarity/Identity, %</th>
<th>Proposed Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>aziU1</td>
<td>257</td>
<td>Orf(–12) (AAJ85502), from S. atroolivaceus</td>
<td>50/32</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>aziU2</td>
<td>221</td>
<td>Rxxyl_2597 (YP_645326), from Rubrobacter xylanophilus DSM 9941</td>
<td>47/30</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>aziU3</td>
<td>337</td>
<td>SMB02513 (NP_437034), from Sinorhizobium meliloti 1021</td>
<td>42/25</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>aziC2</td>
<td>287</td>
<td>Rcas_3469 (YP_001433537), from Roseiflexus castenholzii DSM 13941</td>
<td>61/44</td>
<td>Lysine biosynthesis enzyme LysX</td>
</tr>
<tr>
<td>aziC1</td>
<td>243</td>
<td>At5g27410 (Q9ASR4), from Arabidopsis thaliana</td>
<td>43/28</td>
<td>Branched-chain-amino-acid aminotransferase</td>
</tr>
<tr>
<td>AziC9</td>
<td>419</td>
<td>CypLB (AAT45286), from S. tubercidicus</td>
<td>58/39</td>
<td>Cytochrome P450 hydroxylase</td>
</tr>
<tr>
<td>aziA8</td>
<td>249</td>
<td>BorB (CAE45660), from S. parvulus</td>
<td>54/40</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>aziD3</td>
<td>382</td>
<td>MAV_1338 (YP_880582), from Mycobacterium avium 104</td>
<td>53/33</td>
<td>Acyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>aziD2</td>
<td>333</td>
<td>Acd_9 (YP_326591), from Natronomonas pharaonis DSM 2160</td>
<td>37/25</td>
<td>Acyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>orf1 e</td>
<td>371</td>
<td>Nfa26740 (YP_118885), from Nocardia farcinica IFM 10152</td>
<td>62/50</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>orf2 e</td>
<td>175</td>
<td>SACE_4220 (YP_001106414), from Sac. erythraea NRRL 2338</td>
<td>73/62</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>orf3 e</td>
<td>128</td>
<td>Franean1_6978 (YP_00151217), from Frankia sp. EAN1pec</td>
<td>63/52</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>orf4 e</td>
<td>207</td>
<td>Mmc1_2644 (YP_866546), from Magnetococcus sp. MC-1</td>
<td>54/40</td>
<td>Transposase</td>
</tr>
<tr>
<td>orf5 e</td>
<td>249</td>
<td>ECA0747 (YP_048859), from Erwinia carotovora subsp. Atroseptica SCRI1043</td>
<td>71/59</td>
<td>Short chain dehydrogenase</td>
</tr>
<tr>
<td>orf6 e</td>
<td>249</td>
<td>SACE_4949 (YP_001107140), from Sac. erythraea NRRL 2338</td>
<td>91/84</td>
<td>Short-chain dehydrogenase/reductase</td>
</tr>
<tr>
<td>orf7 e</td>
<td>199</td>
<td>SACE_4948 (YP_001107139), from Sac. erythraea NRRL 2338</td>
<td>88/83</td>
<td>TetR-family transcriptional regulator</td>
</tr>
<tr>
<td>orf8 e</td>
<td>151</td>
<td>StPA0005 (AAW49299), from S. Turgidiscabies</td>
<td>60/47</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>orf9 e</td>
<td>137</td>
<td>Orf2 (BAF46969), from S. griseus</td>
<td>78/63</td>
<td>Transcriptional regulator</td>
</tr>
<tr>
<td>orf10 e</td>
<td>286</td>
<td>SAML0296 (CAJ89283), from S. ambofaciens ATCC 23877</td>
<td>62/45</td>
<td>Dehydrogenase</td>
</tr>
<tr>
<td>orf11 e</td>
<td>493</td>
<td>SCO2878 (NP_627106), from S. coelicolor A3(2)</td>
<td>47/30</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>orf12 e</td>
<td>198</td>
<td>SAV_5747 (NP_826924), from S. avermitilis MA-4680</td>
<td>89/81</td>
<td>Two-component system response regulator</td>
</tr>
</tbody>
</table>

a Numbers are in amino acids.
b NCBI accession numbers are given in parentheses.
c orfs beyond the azi gene cluster.

Further, the constructs harboring mutant aziB, in which site-specific mutations were performed for inactivating the KR and DH domains, respectively, were introduced into S. albus, yielding the recombinant strain AL1007 (to express the KR mutant AziB, G1398A within the NADPH-binding motif GxGxxG), AL1008 (to express the KR mutant AziB, Y1549F at the conserved active site), and AL1009 (to express the DH mutant AziB, H935F within the conserved motif HxxxGxxxxP). Upon HPLC-MS analysis (Figure 5), AL1007, AL1008, and AL1009 failed to produce 5-methyl-NPA, confirming that the reductive and dehydrating actions governed by the KR and DH domains of AziB are essential for the 5-methyl-NPA formation. To establish the entire biosynthetic pathway for the building block 1, a construct that carries aziB along with aziB (encoding
Figure 3. Proposed Biosynthetic Pathway of Azinomycin B

(A–C) Pathways for building blocks (A) 3-methoxy-5-methyl-NPA (1); (B) α-ketoisovaleric acid (2); and (C) aziridino[1, 2a]pyrrolidinyl amino acid (3).

(D) A mode for NRPS-directed skeleton assembly and tailoring modifications.
a putative P450 hydroxylase) and aziB2 (encoding a putative O-methyltransferase) was introduced into S. albus, yielding the recombinant strain AL1005. Again with AL1006 as a negative control, HPLC-MS analysis revealed a distinct compound with an [M−H]− ion at m/z = 215.09 (Figure SIV), and the identity of 3-methoxy-5-methyl-NPA was confirmed by EI-HRMS analysis (M+ ion at m/z = 216.0782) and using the synthesized compound as a standard. These results not only confirmed the assigned functions of aziB1 and aziB2 upon sequence analysis, but also strongly supported the building block 3-methoxy-5-methyl-NPA as the substrate of NRPS AziA1 for skeleton assembly of azinomycin B.

DISCUSSION

Polyketide metabolites are biosynthesized by PKSs in a manner resembling fatty acid syntheses. In bacteria, reduced polyketide (i.e., macrolide, polyene, and polyether) biosynthesis usually requires noniterative type I PKSs, which are giant multifunctional enzymes that function following a linearity rule. Recently, the emergence of many novel type I PKS systems has become apparent in publications (reviewed in Shen, 2003; Walsh, 2004; Wenzel and Muller, 2005), revealing a rich variety of chemistry and architecture of PKSs beyond our previous understanding. For example, a specific module in PKSs could be iteratively used (e.g., borrelidin and aureothin biosynthesis), and single module-contained PKSs can act iteratively for complex structural moieties biosynthesis (e.g., enediyne cores and unsaturated fatty acids).

Upon sequence comparison and heterologous expression in S. albus (Figure 5), the finding of AziB as a 5-methyl-NPA synthase not only adds a new member to the pool of bacterial evolved feature from the ancestor to intrinsically program the chain elongation in an iterative manner. Sequence analysis showed that AziB is genetically closer to NcsB, consistent with the fact that both of them catalyze the biosynthesis of bicyclic polyketides from one acetyl-CoA and five malonyl-CoAs. Based on a model for prediction of the stereochemistry of hydroxyacyl intermediates (Wu et al., 2005; Keatinge-Clay, 2007), similar to KRs of NcsB, ChlB1 and MdpB, the AziB-KR domain lacks the H residue at the active site groove and characteristic L-D-D motif (Figure S4), implicating that it falls into a A1 type KR family and acts on the keto groups to yield L-hydroxy configurations exclusively. Subsequently, cis-double bonds are formed with the dehydration activity of the paired DH domain, facilitating the folding of the resulting polyketide intermediate to furnish the aromatic structure by aldol condensations. In contrast, the KR domains of the iterative type I PKSs for assembling aliphatic chain contain the L-D-D motif (Wenzel et al., 2005), consistent with the predicted D-configuration of the cryptic hydroxyl groups. Further dehydrations catalyzed by its paired DH domain may lead to the formation of the conjugated trans-double bonds. The KR domain of AziB, which acts on the C3, C5, and C9 positions (Figure 6), offers a regiospecifically distinct and relatively rich reduction pattern from those of known iterative type I PKSs, raising an interesting question regarding how these enzymes biochemically control the selective reductions in the assembling process of aromatic polyketides. Previous biochemical studies on the fungal 6-MSA synthases indicated that the NADPH-dependent reduction and dehydration occur on the triketoite intermediate (Shen, 2000; Moriguchi et al., 2008). Either in vitro bioassay in the absence of NADPH or in vivo expression of the KR mutant 6-MSA synthases led to the formation of a same shunt product, triacetic acid lactone (TAL),
Azinomycin B is a complex natural product containing densely assembled functionalities with potent antitumor activity. The availability of the azi biosynthetic gene cluster described here provides an excellent opportunity to understand the complex azinomycin B biosynthetic machinery. The finding of AziB as a 5-methyl-NPA synthase adds a new member to the pool of bacterial iterative type I PKSs and shows the fourth selective reductive pattern in aromatic polyketide biosynthesis governed by this family. Heterologous expression established the PKS-post modification route to reach the first building block, 3-methoxy-5-methyl-NPA. The proposed azinomycin B biosynthetic pathway sets the stage to investigate the enzymatic mechanisms for building the structurally unique and pharmaceutically important groups, including the unprecedented azabicyclic ring system and highly active epoxide moiety. Since many of proposed enzymes and reactions are novel, characterization of these will not only enrich the current knowledge on biosynthetic machineries to make complex products in nature, but also contribute to the general field of combinatorial biosynthesis by expanding the repertoire of novel genes.

**SIGNIFICANCE**

Azinomycin B biosynthetic gene cluster
dense functionalities with potent antitumor activity. The availability of the azi biosynthetic gene cluster described here provides an excellent opportunity to understand the complex azinomycin B biosynthetic machinery. The finding of AziB as a 5-methyl-NPA synthase adds a new member to the pool of bacterial iterative type I PKSs and shows the fourth selective reductive pattern in aromatic polyketide biosynthesis governed by this family. Heterologous expression established the PKS-post modification route to reach the first building block, 3-methoxy-5-methyl-NPA. The proposed azinomycin B biosynthetic pathway sets the stage to investigate the enzymatic mechanisms for building the structurally unique and pharmaceutically important groups, including the unprecedented azabicyclic ring system and highly active epoxide moiety. Since many of proposed enzymes and reactions are novel, characterization of these will not only enrich the current knowledge on biosynthetic machineries to make complex products in nature, but also contribute to the general field of combinatorial biosynthesis by expanding the repertoire of novel genes.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains, Plasmids, and Reagents**

Bacterial strains and plasmids used in this study are summarized in Table S1. Biochemicals, chemicals, media, restriction enzymes, and other molecular biological reagents were purchased from standard commercial sources.

**DNA Isolation, Manipulation, and Sequencing**

DNA isolation and manipulation in E. coli and Streptomyces were carried out according to standard methods (Kieser et al., 2000; Sambrook and Russell, 2001). PCR amplifications were carried out on an Authorized Thermal Cycler (Eppendorf AG) using either Taq DNA polymerase or PhuUltra High-Fidelity DNA polymerase. Primer synthesis and DNA sequencing were performed at the Shanghai Invitrogen Biotech Co., Ltd., and Chinese National Human Genome Center.

**Genomic Library Construction and Screening**

A genomic library of S. sahachiroi NRRL 2485 was constructed in pOJ446 according to the previous protocol (Kieser et al., 2000). E. coli LE392 and
The Packagene Lambda DNA Packaging System (Promega) were used for library construction according to the manufacturers' instructions. For the PCR product P1 as a probe to screen the library, a 0.9 kb fragment was amplified by using the primers 5'-GCG GAC GGC TAC GGS MGNGGN GAR GG-3' and 5'-CGA GCC GTG GCC SGA RAA NAC CCA NAC-3', and confirmed by sequencing. The genomic library (6.0×10^3 colonies) was screened by colony hybridization with P1 and P2 (a 3.2 kb BamHI DNA fragment from cosmid pAL1022) as probes, and resultant positive clones were further confirmed by southern hybridization.

**Sequence Analysis**

The sequence of the azi gene cluster reported in this paper has been deposited into GenBank under the accession number EU240558. The orfs were deduced from the sequence with the FramePlot 3.0 beta program (http://watson.nih.gov/cgi-bin/frameplot-3.0b.pl/). The corresponding deduced proteins were compared with other known proteins in the databases by available BLAST methods (http://www.ncbi.nlm.nih.gov/blast/). Amino acid sequence alignments were performed by the CLUSTALW method and the DRAWTREE and DRAWGRAM methods, respectively, from BIOLOGYWORKBENCH 3.2 software (http://workbench.sdsc.edu/). Prediction of amino acid specificities of individual A and AL domains was performed by using the BLAST server provided at http://www.ncbi.nlm.nih.gov/blast/submit.html.

**Production, Isolation, and Analysis of Azinomycin B in S. sahachiroi**

Sporulation and fermentation of *S. sahachiroi* wild-type and mutant strains were carried out according to the methods described previously (Kelly et al., 2008). For Azinomycin B isolation, each 200 ml culture was centrifuged at 4°C and 3,800 rpm for 10 min. After removal of the cell pellets, the resulting supernatant was extracted twice with an equal volume of methylene chloride. The combined extract was immediately dried over anhydrous magnesium sulfate, concentrated in vacuum, stored at −80°C, and resolved in 200 μl methanol for analysis.

High-performance liquid chromatography–mass spectrometry (HPLC-MS) analysis was carried out on an Agilent Rp18 column (4.6×250 mm, part number 880975-902, S/N USCL021611). The column was equilibrated with 50% solvent A (H2O) and B (CH3CN) and developed with the following program: 0–10 min, 80% A/20% B; 10–35 min, a linear gradient from 80% A/20% B to 20% A/80% B; 35–37 min, a linear gradient from 20% A/80% B to 80% A/20% B; and 37–40 min, constant 80% A/20% B. This was carried out at a flow rate of 0.5 ml/min and UV detection at 218 nm using an Agilent 1100 HPLC system (Agilent Technologies). The compound with a retention time at 21.4 min showed a [M + H]^+ ion at m/z = 624.4 on a mass spectrometer (Agilent Technologies), consistent with the molecular formula of azinomycin B C31H33N3O11. The identity of this compound to azinomycin B was further confirmed by HPLC-MS/MSn analysis (see the Supplemental Data).

**Production, Isolation, and Analysis of 5-Methyl-NPA and 3-Methoxy-5-Methyl-NPA in S. albus**

Each 500 μl spore suspension (cfu 1.0×10^8 cells/ml) of the *S. albus* strain was inoculated into 100 ml R5A medium (Kieser et al., 2000) in a 500 ml flask and incubated at 30°C and 240 rpm for 5 days. The fermentation culture was filtered, adjusted pH to 2–3, fragmented by ultrasound (10S/50S) for 15 min, and then centrifuged to remove the precipitate. The supernatant was extracted twice with an equal volume of ethyl acetate. After being concentrated in vacuum, the extract was resolved in 1 ml methanol for HPLC-MS analysis on a Phenomenex C18 column (4.6×250 mm, part number 00F-3300-E0, S/N 115575-1). The column was equilibrated with 90% solvent A (H2O, 0.1% TFA) and 10% solvent B (CH3CN, 0.1% TFA), and an analytic method was developed with the following program at a flow rate of 1 ml/min and UV detection at 218 nm using an Agilent 1100 HPLC system: 0–5 min, a linear gradient from 90% A/10% B to 80% A/20% B; 5–25 min, a linear gradient from 80% A/20% B to 30% A/70% B; 25–26 min, a linear gradient from 30% A/70% B to 5% A/95% B; and 26–29 min, constant 5% A/95% B, and 29–30 min, a linear gradient consistent with the molecular formula of azinomycin B C31H33N3O11.

The identity of this compound to azinomycin B was further confirmed by HPLC-MS/MSn analysis (see the Supplemental Data).

**Figure 6. Iterative Type I PKSs for Aromatic Polyketide Biosynthesis in Bacteria**

I, AviM and CaIO5 for OSA biosynthesis; II, ChlB1 and MadB for 6-MSA biosynthesis; III, NcsB for 2-hydroxy-5-methyl-NPA biosynthesis; and IV, AziB for 5-methyl-NPA biosynthesis.
from 5% A/95% B to 90% A/10% B. The compound with a retention time at 19.6 min showed a [M − H]− ion at m/z = 185.07, consistent with the molecular formula of 5-methyl-NPA C12H10O2; the compound with a retention time at 20.5 min showed a [M − H]− ion at m/z = 215.09, consistent with the molecular formula of 3-methoxy-5-methyl-NPA C12H9O2.

For purification of 5-methyl-NPA, the extract was subjected to chromatography on a silica column (Petroleum ether-ethyl acetate, 4:1) followed by further preparation on a Venuel XBP-C18 HPLC column (10 × 250 mm, catalog number VX925510, S/N XBP-1025070001) under the condition described above at a flow rate of 3 ml/min. S-methyl-NPA was obtained as a white powder that exhibited an M+ ion at m/z = 186.0675 upon EI-HRMS analysis, consistent with the molecular formula C12H10O2. The 1H NMR spectrum was measured on a Varian Mercury 300 (300 MHz) spectrometer. 1H NMR (CD3OD, δ = 8.1 Hz), and 8.67 (d, 1H, J = 9.6 Hz).

Bioassay of Azinomycin B

To detect the biological activities of azinomycin B against M. luteus, each 20 μl methanol extract described above was added to stainless steel cylinders on LB agar plates that were preseeded with an overnight M. luteus culture at a concentration of 1% (vol/vol). The plates were incubated at 37°C for 24 hr, and the biological activity was estimated by measuring the sizes of the inhibition zones.

SUPPLEMENTAL DATA

Supplemental Data include bacterial strains and plasmids, gene inactivation and heterologous expression, chemical synthesis, HPLC and HPLC-MS/MS analyses, and sequence alignments and can be found with this article online at http://www.chembiol.com/cgi/content/full/15/7/693/DC1.

ACKNOWLEDGMENTS

We thank Cesar Sanchez, Univeridad de Oviedo, Spain, for providing of S. albus J1047 and Johnathan Zandrew Cheng, Massachusetts Institute of Technology, for reading of the manuscript and comments. This work was supported in part by grants from the National Natural Science Foundation of China (30770035, 30252001, 90713012, 20321202, and 20402022), the Ministry of Science and Technology of China (2006AA020304 and 2006AA022185), the Chinese Academy of Science (KJCKX2-YW-H08 and KJCX2-YW-GQ15), and the Science and Technology Commission of Shanghai Municipality (074319115-3 and 04DZ14901).

Received: March 13, 2008
Revised: May 20, 2008
Accepted: May 27, 2008
Published: July 18, 2008

REFERENCES


