Characterization of the Bafilomycin Biosynthetic Gene Cluster from Streptomyces lohii


Bafilomycin B₁ is the archetypal compound of the plecomacrolide family of natural products, which include concanamycins, viramidomycins, hygrolidin, elaiophylin, and allied metabolites.[1] This compound was first isolated from Streptomyces griseus sp. sulphurus (TÜ 1922) in 1984.[2] The term “plecomacrolide” was later introduced to describe a subgroup of polyketides with a 16- or 18-membered macrolactone containing two conjugated diene units as well as a hemiacetal side chain.[3] These compounds have now been studied extensively as vacuolar ATPase (V-ATPase) inhibitors and, as such, have shown some promise as anti-osteoporotics.[4,5] In addition, synergistic antifungal activity of bafilomycin A₁ with the calcineurin inhibitor FK506 against the life-threatening fungal pathogen Cryptococcus neoformans[6] and inhibition of the release of β-amyloid by bafilo- mycin A₁,[7] together with the antitumor,[8] antiparasitic,[9] and immunosuppressant[10] activities of other bafilomycins have been reported. Despite continued efforts, general toxicity has kept these compounds from clinical application. Semisynthetic derivatives have provided some indication that target specificity can be enhanced.[11] However, extensive manipulation of the core structure for structure–activity relationship studies remains hampered by the complexity of total synthesis.[12] Decades of work in the field of polyketide biosynthesis have yielded tools to generate subtle changes in the core macrolide through genetic manipulation.[13,14] In order to use this knowledge to generate novel analogues, the native biosynthetic gene cluster sequence must be elucidated. When a bafilomycin-producing organism, Streptomyces lohii, was isolated during an ongoing drug discovery program in our laboratories, the characterization of the bafilomycin biosynthetic pathway was initiated.

To isolate the genes involved in bafilomycin biosynthesis, a fosmid library was constructed from S. lohii (ATCC BAA-1276, JCM 14114) genomic DNA. This library was probed for type I polyketide ketosynthase (KS) domains with a radiolabeled PCR product amplified from S. lohii gDNA by using degenerate primers. Colony hybridization experiments revealed 12 KS-positive fosmids. Initial restriction digestions and Southern hybridizations revealed two major groups of overlapping fosmids. One clone from each group was chosen for further analysis. BLAST analysis[15] of the sequence from a small number of subclones suggested that one of these representatives contained only type I polyketide synthase (PKS) genes. Complete sequencing of this clone showed that the domain organization of the PKS modules exactly matches with a section of the genetic architecture predicted by the structure of bafilomycin and the typical polyketide rules of colinearity. End sequencing of the other KS-positive fosmids identified those containing overlapping sequence. These were also sequenced to generate a map of the entire gene cluster.

The macrolactone core of bafilomycin is assembled by a type I PKS system. These genes are arranged as five open reading frames, bafAI–bafAV, spanning 59 kb (Scheme 1, Table 1). The domain organization of the PKS modules exactly follows the model suggested by the structure of the core macro lactone and the PKS pattern of colinearity. In addition, acyltransferase (AT) domains 2 and 6 contain the signature motifs associated with the incorporation of methoxymalonate,[16,17] whereas the remaining extender modules are specific for methylmalonate[18] again this matches the predicted motifs of the bafilomycin structure (see the Supporting Information). The AT domain of the loading module was expected to be specific for the incorporation of isobutyrate, as suggested by the metabolite structure and confirmed by precursor-feeding studies.[19] This domain aligns most closely with the isobutyrate loading domains from the lipomyein[20,21] and tautomycin[22] pathways. Downstream of the core open reading frames is bafH, which encodes a PKS type II thioesterase (TEII). Homologues of this enzyme have been found in a number of PKS clusters and are known to possess an editing role during biosynthesis.[23–25]

The identity of this gene cluster was established by targeted disruption of bafilomycin production using the REDIRECT system.[26] When bafAll replaced in-frame by the apramy-
Table 1. Predicted functions of the open reading frames shown in Scheme 1.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino acids</th>
<th>Putative function</th>
<th>Nearest homologue (enzyme, origin)</th>
<th>Identity/similarity [%]</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BafX</td>
<td>528</td>
<td>acyl CoA ligase</td>
<td>MoeA4 [Streptomyces ghanaensis]</td>
<td>72/83</td>
<td>ZP_06574361.1</td>
</tr>
<tr>
<td>BafY</td>
<td>514</td>
<td>amide synthetase</td>
<td>amide synthetase [S. tsukubaensis]</td>
<td>78/85</td>
<td>ZP_10067139.1</td>
</tr>
<tr>
<td>BafZ</td>
<td>414</td>
<td>5-aminolevulinic acid synthase</td>
<td>aminolevulinate synthase [S. nodosus subsp. asukaensis]</td>
<td>73/83</td>
<td>AAO62615.1</td>
</tr>
<tr>
<td>BafAI</td>
<td>4884</td>
<td>PKS loading module + modules 1-3</td>
<td>polyketide synthase [S. neyagawensis]</td>
<td>62/70</td>
<td>AAZ94387.1</td>
</tr>
<tr>
<td>BafAll</td>
<td>5145</td>
<td>PKS modules 4–6</td>
<td>BFA52 [S. tsukubaensis]</td>
<td>78/83</td>
<td>ZP_10067132.1</td>
</tr>
<tr>
<td>BafAllI</td>
<td>3968</td>
<td>PKS modules 7, 8</td>
<td>BFA52 [S. tsukubaensis]</td>
<td>79/85</td>
<td>ZP_10067131.1</td>
</tr>
<tr>
<td>BafAV</td>
<td>3511</td>
<td>PKS modules 9, 10</td>
<td>polyketide synthase [S. scabies]</td>
<td>62/73</td>
<td>YP_003493865.1</td>
</tr>
<tr>
<td>BafB</td>
<td>296</td>
<td>glycerol-ACP oxidase</td>
<td>methoxymalonate biosynthesis protein [S. tsukubaensis]</td>
<td>72/81</td>
<td>ZP_10067126.1</td>
</tr>
<tr>
<td>BafC</td>
<td>93</td>
<td>acyl carrier protein (ACP)</td>
<td>GalI [S. galbus]</td>
<td>69/80</td>
<td>ADE22334.1</td>
</tr>
<tr>
<td>BafD</td>
<td>363</td>
<td>acyl-ACP dehydrogenase</td>
<td>methoxymalonate biosynthesis protein [S. tsukubaensis]</td>
<td>85/93</td>
<td>ZP_10067124.1</td>
</tr>
<tr>
<td>BafE</td>
<td>365</td>
<td>glycerate ACP ligase</td>
<td>methoxymalonate biosynthesis protein [S. tsukubaensis]</td>
<td>83/90</td>
<td>ZP_10067123.1</td>
</tr>
<tr>
<td>BafF</td>
<td>220</td>
<td>O-methyl transferase</td>
<td>O-methyltransferase mdmC [S. clavuligerus]</td>
<td>73/81</td>
<td>ZP_05008783.1</td>
</tr>
<tr>
<td>BafG</td>
<td>609</td>
<td>AfsR homologue</td>
<td>putative regulator [S. neyagawensis]</td>
<td>54/64</td>
<td>AAZ94808.1</td>
</tr>
<tr>
<td>BafH</td>
<td>253</td>
<td>thiosterase [S. auratus]</td>
<td>thiosterase [S. auratus]</td>
<td>65/76</td>
<td>ZP_10551360.1</td>
</tr>
<tr>
<td>ORF1</td>
<td>117</td>
<td>LuxR homologue</td>
<td>LuxR family transcriptional regulator [S. tsukubaensis]</td>
<td>65/73</td>
<td>ZP_10067119.1</td>
</tr>
<tr>
<td>ORF2</td>
<td>320</td>
<td>putative malonyl transferase</td>
<td>acyl-carrier protein 5-malonyltransferase [S. tsukubaensis]</td>
<td>84/90</td>
<td>ZP_10067118.1</td>
</tr>
<tr>
<td>ORF3</td>
<td>332</td>
<td>putative CoA ligase</td>
<td>hypothetical protein [S. neyagawensis]</td>
<td>74/83</td>
<td>AAZ94384.1</td>
</tr>
</tbody>
</table>
cin resistance cassette (genotype confirmed by PCR analysis, see the Supporting Information), the mutant strain of *S. lohii* failed to produce bafilomycins, as demonstrated by HPLC analysis of culture extracts (Figure 1).

The core macrolactone bafilomycin A₁, presumably formed by the activity of the TE domain in BafAV, acquires enhanced bioactivity by the addition of a fumarate moiety to the hydroxyl group on C21 to form bafilomycin C₁.[2] Further potency is provided by a 2-amino-3-hydroxy-cyclopent-2-enone (C₅N) ring attached to the pendant fumaryl group through an amide bond formed with the distal carboxylate to give bafilomycin B₁ (Scheme 1).

Notably, this C₅N moiety, which is present in a variety of secondary metabolites,[27] often serves as a hydrogen bond donor/acceptor pharmacophore. Its biosynthesis has long been proposed to be from 5-aminolevulinic acid (5-ALA) through an unusual cyclization process. This putative transformation was recently confirmed in the ECO-02301 pathway (Scheme 2).[28]

First, the pyridoxal 5'-phosphate (PLP)-dependent 5-ALA synthase (ALAS) ORF34 catalyzes the condensation of succinyl-CoA and glycine to yield 5-ALA, which is subsequently activated by the acyl-CoA ligase (ACL) ORF35 to form 5-ALA-CoA. Next, ORF34 acts again to cyclize 5-ALA-CoA to afford C₅N, avoiding nonenzymatic cyclization to the six-membered ring 2,5-piperidinedione. Finally, the ATP-dependent amide synthetase (AMS) ORF33 is responsible for linking C₅N to the carboxyl group. Interestingly, this enzyme trio is similarly encoded by a set of three genes located at the 5' end of the bafilomycin gene cluster, wherein BafX, Bafy, and BafZ are 62/72%, 40/55%, and 68/80% identical/similar to ORF35, ORF33, and ORF34, respectively, at the protein level. This strongly suggests that the conversion from bafilomycin C₁ to B₁ is likely mediated sequentially by BafZ → BafX → BafY → BafY (Scheme 2).

To provide evidence for this proposed tailoring pathway, the activity of BafZ was characterized in vitro. In addition, this confirmed the proposed functionality of BafZ and validated biochemically the identity of the bafilomycin cluster. It is also expected that BafZ might be useful in industrial fermentation to overproduce 5-ALA, which has been widely used as a photodynamic medicine to treat various cancers, as a selective biodegradable herbicide or insecticide, and as a precursor for the production of heme-containing enzymes, porphyrin, and vitamin B₁₂.[29, 30] Given the similarity between BafZ and ORF33, the BafZ-catalyzed condensation should be classified as a C₅ pathway (the Shemin pathway) for 5-ALA biosynthesis, which does not exist in *E. coli* (the normally selected industrial strain).[30]

To characterize the in vitro activity of BafZ, its encoding gene *bafZ* was PCR amplified from *S. lohii* and cloned into the expression vector pACYCDuet-1. The protein was overexpressed in *E. coli* BL21(DE3) and purified by Ni-NTA column chromatography (see the Supporting Information). Upon incubation of purified BafZ (35 μM) with the substrates glycine (5 mM) and succinyl-CoA (1 mM) at 28 °C for 2 h, significant formation of a new compound with retention time and molecular weight (calcld: 132.0655; found: 132.0658) matching those of an authentic standard of 5-ALA was detected during HR-LCMS analysis (Figure 2).

![Figure 1. HPLC chromatograms of extracts of wild-type *S. lohii* compared to three separate bafAIII gene-disruption mutants.](image1)

![Scheme 2. Putative pathway for 5-ALA biosynthesis and attachment. a) BafZ (5-ALA synthase) or ORF34; b) BafX (acyl-CoA ligase) or ORF35; c) BafZ or ORF34; d) BafY (amide synthetase) or ORF33 for ECO-02301 biosynthesis.](image2)

![Figure 2. Extracted ion chromatograms (m/z 132.06). A) 5-ALA authentic standard; B) the BafZ-catalyzed synthesis of 5-ALA from succinyl-CoA and glycine; C) the negative control with boiled BafZ added to the reaction mixture. The high-resolution mass spectrum of 5-ALA is shown in the Supporting Information.](image3)
Sequencing > 15 kb upstream and downstream of the defined \textit{baf} cluster identified no additional ORFs that encode an enzyme catalyzing attachment of the fumarate moiety. Clustering of all genes involved in the biosynthesis of a given product is the paradigm for \textit{Streptomyces} secondary metabolite production. However, separation of the genes producing 2-amino-3-hydroxy-cyclopent-2-enone from those involved in the production of the core polysaccharide in the moenomycin pathway illustrates one of several exceptions to this rule.\textsuperscript{[13]} Therefore, this could be the case with the bafilomycin cluster in \textit{S. lohii}. Alternatively, it is possible that the amide synthetase \textit{BaFy} also catalyzes ester bond formation between fumarate and bafilomycin A\textsuperscript{1}; this is supported by the recent report that a nonribosomal peptide synthetase condensation enzyme \textit{SgcC5} is capable of catalyzing both amide and ester bond formation.\textsuperscript{[30]} To test these hypotheses, additional work will be necessary, and is currently ongoing in our laboratories.

The incorporation of glycerol-derived precursors has been inferred from feeding studies conducted more than 20 years ago.\textsuperscript{[23, 26] Recently, a number of biosynthetic gene clusters utilizing the glycerol-derived methoxymalonate extender molecule have been sequenced. Associated with these clusters is an operon responsible for the biosynthesis of this uncommon extender unit.\textsuperscript{[17, 34, 36]} Downstream of \textit{bafAI–AV} are \textit{bafB–F}, five open reading frames predicted to encode enzymes responsible for methoxymalonate biosynthesis (Table 1). \textit{BaF}c is a discrete acyl carrier protein that is loaded with glycerate, presumably oxidized from glycerol during primary metabolism by the enzyme \textit{BaFe}. The subsequent steps in the generation of methoxymalonate are each proposed to be catalyzed as a \textit{BaFc}-bound intermediate. First, \textit{BaFb} is proposed to oxidize the 3-hydroxy group into an aldehyde, which is subsequently converted to a carboxylic acid. This conversion is catalyzed by \textit{BaFd}, an acyl-ACP dehydrogenase.\textsuperscript{[67]} Finally, the \(\alpha\)-hydroxyl group is methylated by \textit{BaFf}, a putative O-methyltransferase. Although following a genetic organization conserved within several pathways, each of these genes is most closely related to their discontinuously distributed counterparts found in the concanamycin biosynthetic gene cluster from \textit{Streptomyces neagawaensis} ATCC 27449.\textsuperscript{[17]}

Based on the results of feeding studies, Schumann et al. have proposed that methoxymalonate is selected by specific AT domains directly from the dedicated ACP (\textit{BaFc}).\textsuperscript{[19]} Specific protein interaction was inferred due to low incorporation of an \(N\)-acyethylcysteamine thioester (\textit{SNAC}) of methoxymalonate. However, it is possible that methoxymalonate is transferred to coenzyme A, which then acts as the substrate for the AT. This scenario is more consistent with the classical mechanisms of PKS systems. The enzyme encoded by ORF3 could possibly fulfill this function. It is predicted to be a CoA ligase and shares high homology with an enzyme encoded as part of the concanamycin biosynthetic gene cluster.\textsuperscript{[17]}

Downstream of the biosynthesis ORFs are a number of genes encoding putative regulatory elements. \textit{BaFG} is a homolog of \textit{AfsR}, which has been shown to be a pleiotropic transcriptional activator involved in the regulation of secondary metabolite biosynthesis in the model organism \textit{Streptomyces coelicolor}.\textsuperscript{[38]} Homologues of this protein are associated with a number of secondary metabolite biosynthesis pathways within the \textit{Streptomyces} genus, including the closely related concanamycin pathway.\textsuperscript{[17, 38]} ORF1 encodes a member of the LuxR family of transcriptional repressors. Repression is relieved by the binding of a quorum-sensing molecule, most often a homoserine lactone.\textsuperscript{[40]} LuxR family proteins are associated with a number of secondary metabolite biosynthesis clusters, including the concanamycin system.\textsuperscript{[17]}

In summary, sequencing and analysis of the bafilomycin biosynthetic gene clusters has elucidated a new 16-membered pleomacrolide pathway identified after the elaiophylin gene cluster.\textsuperscript{[41]} Recently, the discovery of more bafilomycin derivatives\textsuperscript{[42–44]} from diverse \textit{Streptomyces} strains and additional putative bafilomycin biosynthetic genes\textsuperscript{[45]} has suggested broad distribution of this biosynthetic system. Thus, our work provides an important reference for the identification and analysis of analogous gene clusters, and crucial information for the future engineered biosynthesis of bafilomycin analogues. Moreover, the in vitro characterization of \textit{BaFz} expands the pool of 5-ALA synthases, and this might benefit industrial production of 5-ALA.

### Experimental Section

Genomic DNA was isolated from \textit{S. lohii} by using standard \textit{Streptomyces} protocols.\textsuperscript{[46]} A genomic library was constructed in \textit{E. coli} Epi300 by using a Fosmid (Epiconcentre) vector. Initial candidates were selected from a library of 1152 clones by probing with a radio-labeled ketosynthase gene fragment. This probe was a heterologous gene fragment generated by PCR with degenerate KS targeting primers 4U and S1L (see the Supporting Information) from \textit{S. lohii} gDNA as the template.

The sequence was assembled by using Seqman software (DNASTar). Translation analysis was carried out by using the MacVector open reading frame analysis function. Functional assignment was accomplished by BLAST analysis\textsuperscript{[13]} against amino acid sequences contained within the NCBI and EMBL databases. Polyketide synthase domain organization was facilitated by the PKS/NRPS database maintained by the Indian National Institute of Immunology.\textsuperscript{[45]}

Genes in \textit{S. lohii} were disrupted by using a modified version of the REDIRECT system.\textsuperscript{[20]} A suicide knockout vector was created by using fosmid BD3 from the cluster sequencing effort by replacing the gene of interest with a selectable disruption cassette. This fosmid was transferred by interspecies conjugation from \textit{E. coli} ET12567 into \textit{Streptomyces}. Genetic confirmation of the disruption was accomplished by PCR amplification with genomic DNA as template.

Metabolite profiling of the genetically confirmed knockout strains and one wild-type strain was accomplished by extracting cultures grown (12 mL) for 72 h. The cultures were centrifuged to remove cells, and the supernatant was extracted twice with an equal volume of \(\text{CH}_2\text{Cl}_2\). The organic extracts from each culture were pooled and dried in vacuo, and redissolved in DMSO. Extracts were analyzed by LC-MS on a C-18 column eluted with a gradient of acetonitrile in water with 0.1% formic acid. The mass and retention times were validated by comparison to an authentic standard purchased from Sigma Aldrich.
The bafZ gene was amplified from the gDNA of S. lohii by using the following primers: forward, cgccggtgtgatcagcctctctc (BamHI site italicized); reverse, caaaggggtcatcggctcgccgttccgagcactcgtt (HindIII site underlined), and inserted into pACYCDuet-1 plasmid between the restriction sites of BamHI and HindIII to obtain the expression vector pACYCDuet-1-bafZ. To prepare BaFZ protein, E. coli BL21(DE3) cells carrying pACYCDuet-1-bafZ were grown at 37 °C in isosoygeny broth containing chloramphenicol (25 μg mL⁻¹) until OD₆₀₀ = 0.4 was reached. Then, isopropyl-β-D-thiogalactopyranoside (IPTG; 0.1 mM final concentration) was added to initiate protein overproduction. The cells were cultured at 16 °C for another 16 h, collected by centrifugation (5000 g, 10 min, 4 °C), resuspended in lysis buffer (30 mL, 20 mM HEPES, pH 8.0, 0.5 mM NaCl, 5 mM imidazole), and lysed by sonication on ice. After a high-speed centrifugation (40 000 g, 30 min, 4 °C) to remove cell debris, Ni-NTA agarose resin (1 mL, Qiagen) was added to the supernatant. The authors are grateful to PNG BioNet and the University of Papua New Guinea Department of the Environment and Conservation for permission to collect research samples.

Drug Discovery Today


Acknowledgements

This work was supported by funding from the “Recruitment Program of Global Experts, 2012” (to S.L.) and the NIH grant GM076477 (to D.H.S.). The authors are grateful to PNG BioNet and the University of Papua New Guinea Department of the Environment and Conservation for permission to collect research samples.

Keywords: bafilomycins • biosynthesis • plecomarolides • polyketides • secondary metabolites

CHEM B I O C H E M  COMMUNICATIONS

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

ChemBioChem 2013, 14, 301 – 306 305


Received: November 29, 2012
Published online on January 29, 2013