Bioorganic Chemistry **27,** 155–171 (1999)
Article ID bioo.1998.1131, available online at http://www.idealibrary.com on IDEAL®

Bleomycin Biosynthesis in *Streptomyces verticillus* ATCC15003: A Model of Hybrid Peptide and Polyketide Biosynthesis

Ben Shen,[1] Liangcheng Du, Cesar Sanchez, Mei Chen, and Daniel J. Edwards

Department of Chemistry, University of California, One Shields Avenue, Davis, California 95616

Received July 31, 1998

The biosynthesis of bleomycins (BLMs) in *Streptomyces verticillus* (*Sv*) ATCC15003 was reviewed. Early biosynthetic studies by incorporation of isotope-labeled precursors and by isolation of biosynthetic intermediates and shunt metabolites were presented to support the hypothesis that (a) BLM is a hybrid peptide and polyketide metabolite and (b) the *blm* synthetase, which catalyzes the assembly of BLM from nine amino acids and an acetate, should bear the characteristics of both nonribosomal peptide synthetase (PTS) and polyketide synthase (PKS). After brief discussion of the cloning and characterization of the *blm* resistance genes from *Sv* ATCC15003 as well as from other microorganisms, emphasis was placed on our current efforts to clone the *blm* biosynthesis gene cluster from *Sv* ATCC15003. Four cloning strategies were discussed that included chromosomal walking from the *blmAB* resistance genes, cloning putative PKS genes by polymerase chain reaction (PCR), cloning putative PTS genes by PCR, and the combination of chromosomal walking from the *blm* resistance locus and heterologous probing with PTS probes. While at least three additional peptide and one polyketide biosynthesis gene clusters were identified from *Sv* ATCC15003, a putative 110-kb *blm* gene cluster has been cloned. Sequence analysis of 75 kb DNA provided strong support that the cloned gene cluster is responsible for BLM production.

I. INTRODUCTION

The bleomycins (BLMs)[2] are a family of glycopeptide-derived antibiotics originally isolated by Umezawa in 1966 from the fermentation broth of *Streptomyces verticillus* (*Sv*) (*1*). The naturally occurring BLMs differ structurally primarily at the C-terminus of the glycopeptide. The BLM structure was first proposed in 1972 (*2*), revised in 1978 (*3*), and confirmed by total synthesis in 1982 (*4,5*), as shown in Fig. 1. BLMs exhibit strong antitumor activity and are incorporated into current chemotherapy of several malignancies (*6*). The commercial product, Blenoxane, contains BLM A2, **1**, and BLM B2, **2**, as the principal constituents. However, wide application of this

[1] To whom correspondence and reprint requests should be addressed. Fax: (530) 752-8995. E-mail: shen@chem.ucdavis.edu.

[2] **Abbreviations used**: Ala, alanine; Asn, asparagine; BLM, bleomycin; Cys, cysteine; His, histidine; KS, β-ketoacyl synthase; MRSA, methicillin-resistant *Staphylococus aureus*; Met, methionine; ORF, open reading frame; PCR, polymerase chain reaction; PKS, polyketide synthase; PTS, peptide synthetase; Ser, serine; *Sv, Streptomyces verticillus*; Thr, threonine.

0045-2068/99 $30.00

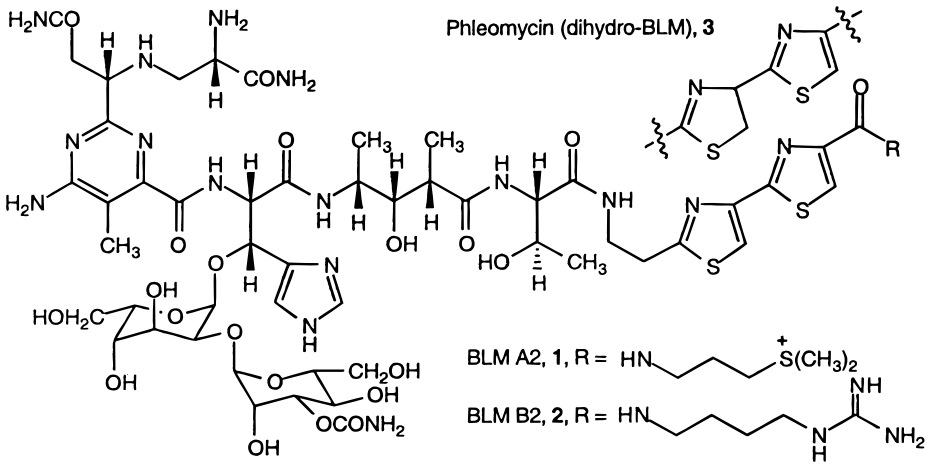

FIG. 1. Structures of bleomycins and phleomycins (dihydrobleomycins).

agent has been prevented by early development of drug resistance and dose-limiting pulmonary toxicity (*6,7*).

BLMs are thought to exert their biological effects through a metal-dependent oxidative cleavage of DNA or RNA in the presence of molecular oxygen (*8–10*). The pyrimidoblamic acid subunit, along with the adjacent β-hydroxyl histidine, provide the metal chelation coordination sites required for Fe(II) complexation and molecular oxygen activation responsible for the subsequent DNA cleavage. The C-terminal tri- or tetrapeptide subunit provide the majority of the BLM-DNA affinity and may contribute to polynucleotide recognition and the DNA cleavage selectivity. There have been continuing attempts to develop new BLM congeners to define the fundamental functional roles of the individual subunits and to search for drugs with better clinical efficacy and lower toxicity (*11–13*). Although total chemical synthesis of BLM has been accomplished (*4,5,13–17*) and semisynthesis and directed biosynthesis have produced a large number of BLM congeners (*6*), the former has very limited practical value and the latter has been problematic due to the heterogeneity and structural complexity of the naturally available BLMs and can only access a few functional groups.

The biosynthesis of BLM was extensively pursued in the late seventies and early eighties. During that time, high producing strains were selected, biosynthetic intermediates and minor metabolites were isolated, and primary biosynthetic precursors were identified. Fujii (*18*), Umezawa (*19*), and Takita (*20,21*) reviewed briefly these early studies, which were mainly driven by the prospect of isolating a highly productive strain for commercial production of BLM and of incorporating unnatural biosynthetic precursors into BLM to generate unnatural congeners for the development of a better BLM drug. After experiencing a relatively quiet period in the late eighties and early nineties, the biosynthesis of BLM has been rejuvenated in the past few years due to the belief that antibiotic production genes could be used to facilitate the characterization of biosynthetic enzymes and that novel congeners could be generated by manipulating

genes governing antibiotic biosynthesis. Attempts to clone genes for BLM biosynthesis were initiated by Davies and Sugiyama (*21*) and Schmidt (*22*), who cloned the *blm* resistance genes in 1994 from *Sv* ATCC15003.

Since we view BLM as a natural hybrid metabolite of peptide and polyketide biosynthesis, we are entering the field of BLM biosynthesis from a different perspective. Significant progress has been made recently in the chemistry, biochemistry, and genetics for peptide and polyketide biosynthesis in microorganisms. The seemingly unrelated peptide and polyketide metabolites are assembled in a remarkably similar fashion by repetitive addition of the extending units, being an amino acid (activated as an acyl adenylate) for the nonribosomal peptide synthetase (PTS) or a fatty acid (activated as an acyl CoA thioester) for polyketide synthase (PKS), to a growing chain (*24*). Both systems have evolved a modular organization to define the number, sequence, and specificity for the incorporation of the extending unit and utilized the same 4′-phosphopantetheine prosthetic group to channel the growing intermediate during the elongation process. It is very tempting to propose that a PTS-bound growing peptide intermediate could be further elongated by a PKS module or vice versa, resulting in the production of hybrid peptide and polyketide metabolites. Since BLM is a natural hybrid peptide and polyketide metabolite, we predict that the Blm synthetase, which catalyzes the assembly of the BLM backbone from nine amino acids and an acetate, should bear the characteristics of both PTS and PKS, on the assumption that the biosynthesis of BLM follows the paradigms of peptide and polyketide biosynthesis. BLM biosynthesis could, therefore, provide an excellent model to study the mechanism of how PTS and PKS can be hybridized into a functional system to make metabolites from amino acids and short fatty acids.

Here we will first review the early biosynthetic studies in such a way as to support our hypothesis that BLM is of hybrid peptide and polyketide biosynthesis. We will then discuss strategies available for cloning the BLM production genes. We will conclude by presenting a progress report on our current efforts to clone the putative *blm* gene cluster from *Sv* ATCC15003.

II. EARLY BIOSYNTHETIC STUDIES

Feeding of Isotope-Labeled Precursors

The wild-type BLM-producing *Sv* strain was designated as ATCC15003. The productivity of BLM was improved by selection of the highly productive strain among mutants that were generated upon chemical, UV, and ^{60}Co mutagenesis of *Sv* ATCC15003. One of the improved strains was designated as *Sv* ATCC31307, whose availability greatly facilitated the *in vivo* studies of BLM biosynthesis by feeding isotope-labeled precursors and the isolation and characterization of BLM and its biosynthetic intermediates.

BLM could be viewed as consisting of a linear peptide/polyketide/peptide backbone of P-6MO, **4** (Fig. 2), a terminal amine, and a disaccharide that is composed of 3-*O*-carbamoyl-D-mannose and L-glucose (Fig. 1). Fujii (*18*), Umezawa (*18*,*25*), Takita (*20*,*21*), and coworkers studied the biosynthesis of BLM by feeding experiments with ^{14}C- and ^{13}C-labeled compounds. These studies showed that the BLM aglycone, **5**, was derived from a serine (Ser), two asparagines (Asn), a histidine (His), an alanine

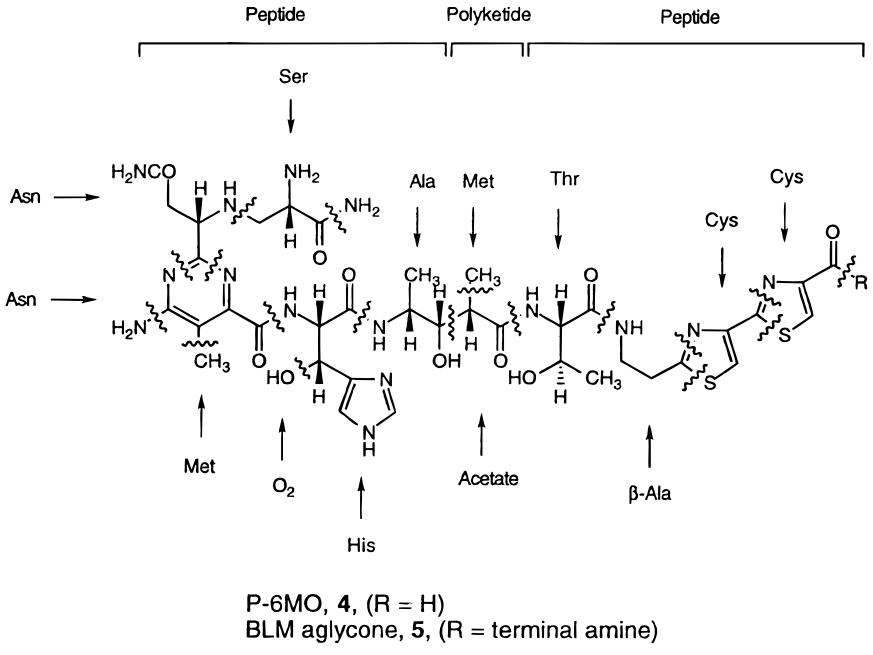

FIG. 2. Primary precursors for bleomycin biosynthesis in *Sv* ATCC15003.

(Ala), an acetate, a threonine (Thr), a β-alanine (β-Ala), and two cysteines (Cys), indicative of a hybrid peptide and polyketide biogenesis, with the two methyl groups from methionine (Met) and the β-hydroxy group of His from molecular oxygen (Fig. 2).

The first evidence supporting the hypothesis that BLM biosynthesis involves a nonribosomal PTS mechanism came from the observation that addition of an amine to the fermentation medium suppressed the production of BLM congeners with other C-terminal amines and that ^{14}C-labeled amines were incorporated specifically at the C-terminal amine moiety of BLM (*26*). The assembly of **5** must have initiated at the β-aminoalaninamide moiety and concluded with the introduction of the terminal amine, consistent with the multiple carrier thiotemplate model of PTS effecting N to C stepwise assembly of the peptide product (*24*). In fact, many unnatural BLM congeners were prepared by fermentation in the presence of various unusual amines, some of which showed enhanced therapeutic properties and were developed into second- and third-generation BLM drugs for clinical application (*21*,*26*).

Isolation and Characterization of Biosynthetic Intermediates and Cell-Free Studies

The hybrid peptide and polyketide biosynthesis of **4** became apparent upon isolation and characterization of biosynthetic intermediates and shunt metabolites from *Sv* cultures. Thus, Fujii and coworkers painstakingly isolated intermediate peptide fragments from fermentation cultures made by using a high-potency strain in a high-productivity medium (*18*). The structures of these intermediates were established

primarily by NMR, with a few of the key structures confirmed by X-ray crystallographic analysis (*18*). Taking together these results and those from feeding experiments, Umezawa, Fujii, and Takita proposed a biosynthetic pathway for BLM, as shown in Fig. 3 (*21*). The biosynthesis starts with the formation of desmethyl pyrimidoblamic acid, **6**, from Ser, Asn, and Asn. Stepwise elongation of **6** with His and Ala yields P-3, **7**, and P-3A, **8**, respectively. These steps are presumably catalyzed by PTS enzymes. The isolation of P-3K, **9**, suggests that the conversion of **8** to P-4, **10**, is likely to proceed in the following sequence: condensation with malonyl CoA, methylation with *S*-adenosylmethionine, and β-ketone reduction—a characteristic reaction sequence catalyzed by a PKS module. Therefore, the peptide intermediate **8**, presumably in a PTS-bound form, must have been elongated by a PKS module to yield a PKS-bound form of **10**, implying the transfer of a growing intermediate from a PTS to a PKS enzyme. Further elongation of **10** with Thr results in P-5, **11**, at which the methyl group at the pyridine ring was introduced from *S*-adenosylmethionine, as evidenced by the isolation of P-5m, **12**. In contrast to the **8**-to-**10** conversion, the elongation of **10**, presumably in a PKS-bound form, by Thr to form **11** must have been catalyzed by a PTS enzyme, implying the transfer of a growing intermediate from a PKS to a PTS enzyme. Continuous elongation of **12** by PTS enzymes with β-Ala, Cys, and Cys to form P-6m, **13**, was consistent with the failure to incorporate ^{3}H-labeled 2′-(2-aminoethyl)-2,4′-bithiazole-4-carboxylic acid, **14**, into BLM *in vivo* (*21*). The conversion of **3** into **1** (*18*) by a cell-free preparation from *Sv* ATCC15003 in the presence of NAD^+ suggested that intermediates such as **15** could be involved in the **12**-to-**13** conversion. Hydroxylation at the β-position of the His residue of **13** yields P-6mo, **4**, which is finally aminated and glycosylated to yield BLM. The fact that the aglycon, **5**, was isolated from the fermentation broth suggested that the glycosylation most likely takes place after amination (*18,21*).

III. CLONING AND CHARACTERIZATION OF THE *blm* RESISTANCE GENES

Cloning of the Blm Resistance Genes

BLM resistance is widely spread among non-BLM producing Gram-negative and Gram-positive bacteria (in particular those of clinical importance), in spite of its lack of therapeutic use as an antibacterial drug (*27*). Two *blm* resistance genes, one from transposon Tn5 (*27,28*) and the other from *Staphylococcus aureus* plasmid pUB 110 (*29*), were characterized before the cloning of any *blm* resistance gene from a BLM producer. (Another *blm* resistance gene, encoding the BLM hydrolase, was considered to be irrelevant to BLM biosynthesis (*30*) and, therefore, is not included here.) Tiraby and coworkers cloned a *blm* resistance gene from *Streptoalloteichus hindustanus* ATCC31158, which produces phleomycin, **3**, an analog of BLM, and named it the Sh *ble* gene (*31*). More recently, Sugiyama and coworkers noticed that the methicillin-resistant *S. aureus* (MRSA) strains isolated from their university hospital were BLM-resistant even after treatment with a plasmid cutting agent. They reasoned that there must have been a chromosome-born *blm* resistance element in these strains. Using primers designed according to the *blm* resistance gene from the *S. aureus* pUB 110,

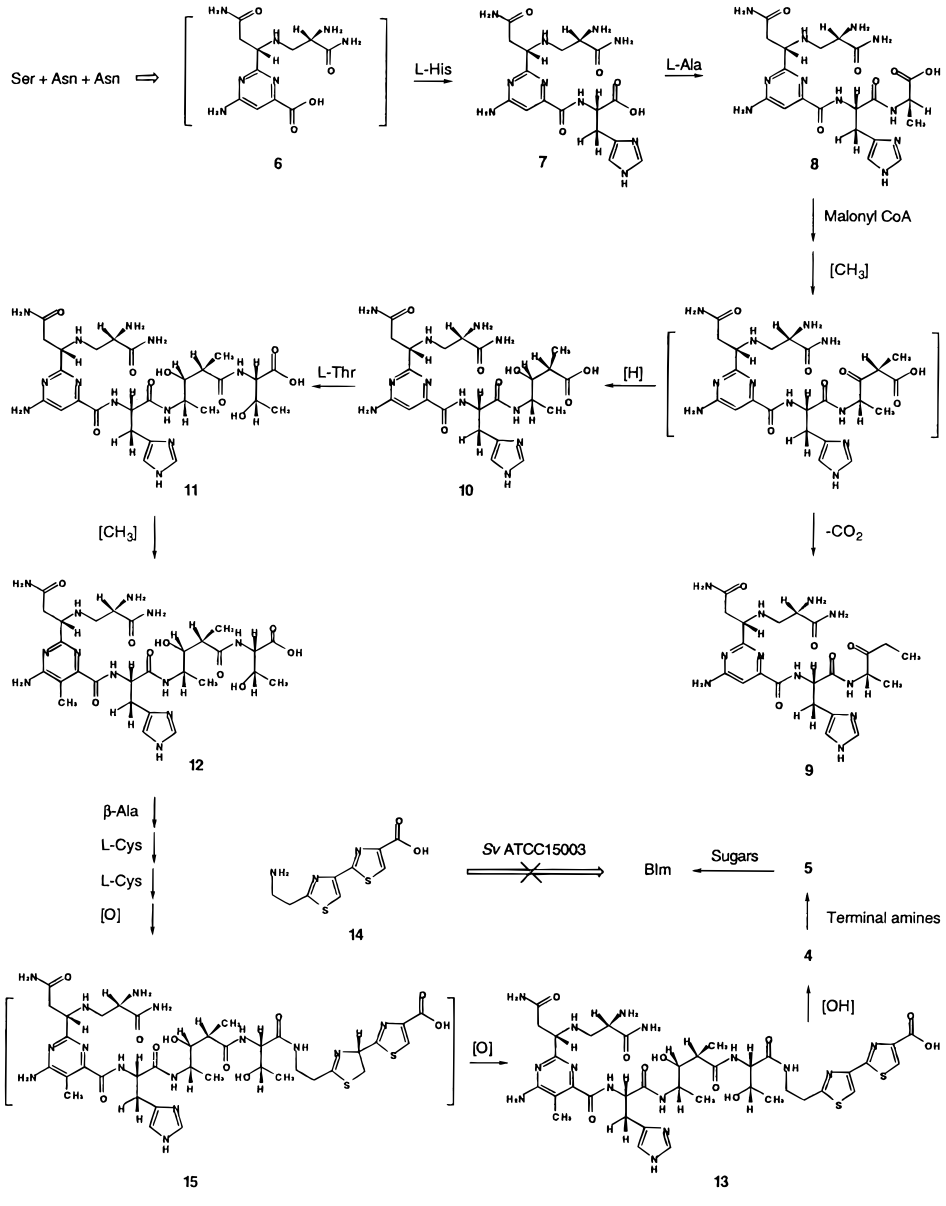

FIG. 3. Biosynthetic pathway of bleomycin in *Sv* ATCC15003 (Intermediates except those in the brackets were identified). [H], reduction; [CH_3], methylation; [OH], hydroxylation; [O], oxidation.

they indeed amplified by polymerase chain reaction (PCR) a *blm* resistance gene from the chromosomal DNA of a MRSA strain. Surprisingly, sequence analysis revealed that the chromosomal *blm* resistance gene was identical to that from pUB 110, both of which were clustered with the kanamycin resistance gene (*32*). They

hypothesized that this *blm* resistance gene was originally present in pUB 110 and was later integrated into the chromosome of the MRSA strains.

Davies and coworkers found that a *Sv* ATCC15003 cell-free extract inactivated BLM in the presence of acetyl CoA (*22*). They subsequently attributed the inactivation of BLM to a BLM acetyltransferase, which catalyzes the formation of the inactive derivative BLM acetate. By screening BLM resistance *in vivo* in *Streptomyces lividans* and BLM inactivation *in vitro* to form BLM acetate, they isolated from the chromosomal DNA of *Sv* ATCC15003 a 6-kb fragment. Further characterization of the 6-kb fragment led them to conclude that this fragment in fact harbored two *blm* resistance elements, only one of which carried the BLM acetyltransferase activity. Sequence analysis indeed revealed two open reading frames (ORFs), *blmA* and *blmB*. *blmA* encodes an acidic protein of 122 amino acids that is highly homologous to Sh Ble (60% similarity). *blmB* encodes a protein of 301 amino acids that shows no significant sequence similarity to proteins in the databases. Concurrently, with *blmAB* as probes, Calcutt and Schmidt isolated cosmids from a genomic library of *Sv* ATCC15003 and confirmed the presence of the *blmAB* genes by sequencing a 7.2-kb DNA fragment (*23*).

Overexpression and Characterization of the BlmAB Proteins

The mechanism for BlmA to confer BLM resistance by drug sequestering is well established. Both the Sh *ble* (*33*) and *blmA* (*34*) genes as well as the *blm* resistance genes from transposon Tn5 (*28*), *S. aureus* pUB 110 (*29*), and *S. aureus* MRSA (*34*) were overexpressed in *E. coli*, and the resulting proteins were purified. Tiraby first established that Sh Ble binds to BLM reversibly in a one-to-one ratio without enzymatically modifying BLM, and the resulting complex can no longer cleave DNA. This was supported by the fact that either heating or treating the Sh Ble-BLM complex with proteinase K restored BLM its full DNA-cleaving potency, presumably by inactivation of the Sh Ble protein (*33*). A similar conclusion was reached by Sugiyama and coworkers who carried out the same experiment with the BlmA-BLM complex (*22*). Additional evidence supporting reversible binding came from immunological characterization of BlmA. Sugiyama and coworkers raised a monoclonal antibody against BlmA and reported that the DNA-cleaving ability of BLM can be restored by addition of anti-BlmA to a BlmA-BLM complex. Surprisingly, no cross reactivity was detected between anti-BlmA and Sh Ble or vice versa (*34*), suggesting that BlmA and Sh Ble may adopt different three-dimensional structures despite the fact that the two proteins show 60% similarity.

◦ Dumas and coworkers determined the crystal structure of the Sh Ble protein at 2.3 A resolution by X-ray crystallographic analysis (*35*). The most prominent feature of the structure is that it is a dimer composed of two identically folded halves with a crevice located at the dimer interface. Analytical centrifugation and light scattering analysis showed that this dimeric form exists in solution as well, establishing that one dimer binds two BLM molecules. Intrigued by the fact that Sh Ble and BlmA are highly homologous and functionally the same, yet display significant difference in immunological response (*34*), Sugiyama and coworkers have very recently succeeded in generating structural data for BlmA by x-ray crystallographic analysis (*36*).

It would be very interesting to compare the BlmA structure, when it becomes available, with that of Sh Ble.

The inactivation of BLM by BlmB was established by overexpression of *blmB* in both *E. coli* and *S. lividans* (*37*), inactivation of BLM with the purified BlmB *in vitro* by covalent modification of BLM to form the inactive BLM acetate (*22*), and characterization of the resulting BLM acetate by NMR analysis (*38*). Sugiyama and coworkers prepared BLM A2 acetate *in vitro* by incubating BLM A2 with a recombinant BlmB in the presence of acetyl CoA. NMR analysis of the purified BLM A2 acetate showed a characteristic signal for the N-acetyl group at 2.059 ppm, identical with that for the chemically synthesized standard. These results unambiguously established that the acetylation occurs at the β-aminoalaninamide moiety of BLM (*38*). BlmB is very specific for **1**, with K_m and V_{max} values of 13 μM and 3.4 nmol min^{-1} ml^{-1}, respectively; neither **3** nor bleomycinic acid, which lacks the terminal amine moiety of BLM, can be served as a substrate for BlmB (*37,39*).

IV. CLONING OF THE *blm* GENE CLUSTER

Chromosomal Walking from the blmAB Locus

Since antibiotic production genes commonly occur as a cluster of structural, resistance, and regulatory genes, it is a sensible approach to clone the *blm* gene cluster by chromosomal walking from the *blmAB* resistance genes. Using *blmAB* as probes, Calcutt and Schmidt isolated cosmid clones spanning the resistance region from a *Sv* ATCC15003 genomic library and sequenced 7.2 kb of the cloned DNA (*23*). Flanking the *blmAB* genes, five additional ORFs were identified (Fig. 4A), none of which, however, showed homology to known PTS or PKS genes. The deduced products of ORF4 and ORF6 had no similarity to any sequences in the databases. The other three ORFs encoded for a putative amino acid dioxygenase (ORF1), a protein containing three ankyrin-like repeats (ORF3), and a putative ATP-binding cassette (ABC) transport protein (ORF7) (*23*).

Calcutt and Schmidt further proposed that *blmORF1* could encode for the His dioxygenase catalyzing the β-hydroxylation of **13** to **4** and suggested that *blm* resistance and structural genes are clustered. To test this hypothesis, we overexpressed *blmORF1* in both *E. coli* and *S. lividans* and purified the recombinant BlmORF1 protein. Modeled on α-ketoglutarate-dependent dioxygenase (*40,41*), the BlmORF1 protein was assayed by incubating **13** with α-ketoglutarate, O_2, Fe^{2+}, ascorbate, and catalase, in 50 mM Tris–HCl, pH 7.5, at 30°C for 2 h. The reaction mixture was monitored by HPLC analysis with authentic **13** and **4** as standards. Under all conditions tested, however, no *in vitro* transformation of **13** to **4** was detected, failing to support BlmORF1 as the putative histidinyl β-hydroxylase involved in BLM biosynthesis.

We subsequently sequenced the 8.2 kb upstream region of the *blmORF1* gene and identified seven additional ORFs, ORF8-15 (Fig. 4A). Based on amino acid sequence similarity to enzymes of known function, these ORFs could be tentatively assigned to encode for a putative oxidase (ORF8), an ADP-heptose synthase homolog (ORF9), a small protein homologous to the peptidyl carrier protein domain of PTS (ORF10), a putative carbamoyltransferase (ORF11), a heptosyltransferase homolog (ORF12), as well as unknown functions for the remaining three ORFs (ORF13–15). Since the