The Blasticidin S Biosynthesis Gene Cluster from *Streptomyces griseochromogenes*: Sequence Analysis, Organization, and Initial Characterization

Martha C. Cone,[a] Xihou Yin,[a] Laura L. Grochowski,[a] Morgan R. Parker,[a] and T. Mark Zabriskie*[a, b]

Blasticidin S is a potent antifungal and cytotoxic peptidyl nucleoside antibiotic from Streptomyces griseochromogenes. *The mixed biosynthesis of the compound is evident from the three distinct structural components: a cytosine base, an amino deoxyglucuronic acid, and* N*-methyl* β*-arginine. The blasticidin S biosynthesis gene cluster was cloned from* S. griseochromogenes *and the pathway heterologously expressed in* S. lividans *from a cosmid harboring a 36.7-kb fragment of* S. griseochromogenes *DNA. The complete DNA sequence of this insert has now been determined and evidence suggests a contiguous 20-kb section defines the blasticidin S biosynthesis cluster. The predicted functions of several open reading frames are consistent with the expected biochemistry and include an arginine 2,3-aminomutase, a cytosylglucuronic acid synthase, and a guanidino* N*-methyltransferase. Insight into other steps in the assembly of blasticidin S was evident from sequence homology with proteins of known function and heterologous expression of fragments of the cluster. Additionally, the gene that directs the production of free cytosine,* blsM, *was subcloned and expressed in* Escherichia coli. *Characterization of BlsM revealed that cytidine monophosphate serves as the precursor to cytosine.*

KEYWORDS:

antifungal agents · biosynthesis · heterologous expression · peptidyl nucleoside

Introduction

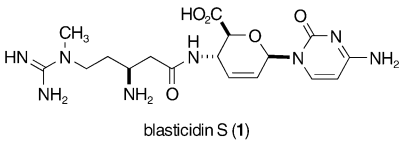
The peptidyl nucleoside family of antibiotics encompasses a structurally diverse group of compounds, many of which exhibit potent and varied biological activities. Most members of the family are comprised of three distinct structural elements: a heterocyclic base, an amino sugar, and an unusual amino acid or peptidyl moiety. Representative examples include blasticidin S (**1**), nikkomycin X (**2**), puromycin (**3**), and streptothricin F (**4**). Interest surrounding these compounds largely stems from their broad spectrum of biological activities, which include antitumor, antiviral, antibacterial, and antifungal activity. Formation of the individual structural components and assembly of the final product are expected to involve unusual biochemistry and there has thus been substantial effort devoted to studying their biosynthesis. The biosynthetic gene clusters for the nikkomycins,[1] puromycin,[2] and streptothricin F[3, 4] have been cloned and sequenced. Investigations of individual steps in nikkomycin formation have revealed unique routes to the pyridyl moiety[5] and the novel role of three proteins involved in transforming histidine into the 4-formyl-4-imidazolin-2-one base of nikkomycin X,[6] among other discoveries.[7–9] Similarly, the roles of several individual gene products of the puromycin pathway have been characterized.[10–13] More limited biochemical work has been directed at streptothricin F formation, about which a single study on the steps required for activation and incorporation of the β-lysine residues was recently reported.[14] Biochemical studies on blasticidin S have included purification and characterization of a unique glucuronosyltransferase that forms the cytosylglucuronic acid intermediate.[15] Work performed in crude cell-free systems detected activities for an aminosugar tautomerase and acetyl and methyltransferases and an unusual self-resistance mechanism.[16, 17] Previously, we identified and cloned the blasticidin S gene cluster from *Streptomyces griseochromogenes* and expressed the pathway in *S. lividans*.[18]

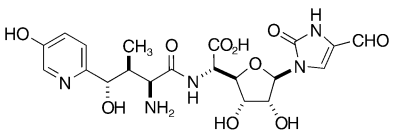
Blasticidin S (**1**) was first identified from an extract of *S. griseochromogenes* in 1958 in a screening effort to discover nonmercurial fungicides.[19] Specifically, **1** was effective at protecting rice plants from infection by the fungus that causes

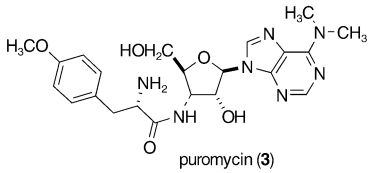
[a] *Prof. T. M. Zabriskie, Dr. M. C. Cone, Dr. X. Yin, L. L. Grochowski, M. R. Parker*
Department of Pharmaceutical Sciences
Oregon State University
Corvallis, Oregon 97331 (USA)
Fax: (+1) 541-737-3999
E-mail: mark.Zabriskie@oregonstate.edu

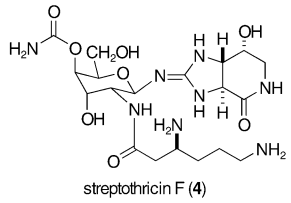
[b] *Prof. T. M. Zabriskie*
Program in Molecular and Cellular Biology
Oregon State University
Corvallis, Oregon 97331 (USA)

 DOI: 10.1002/cbic.200300583



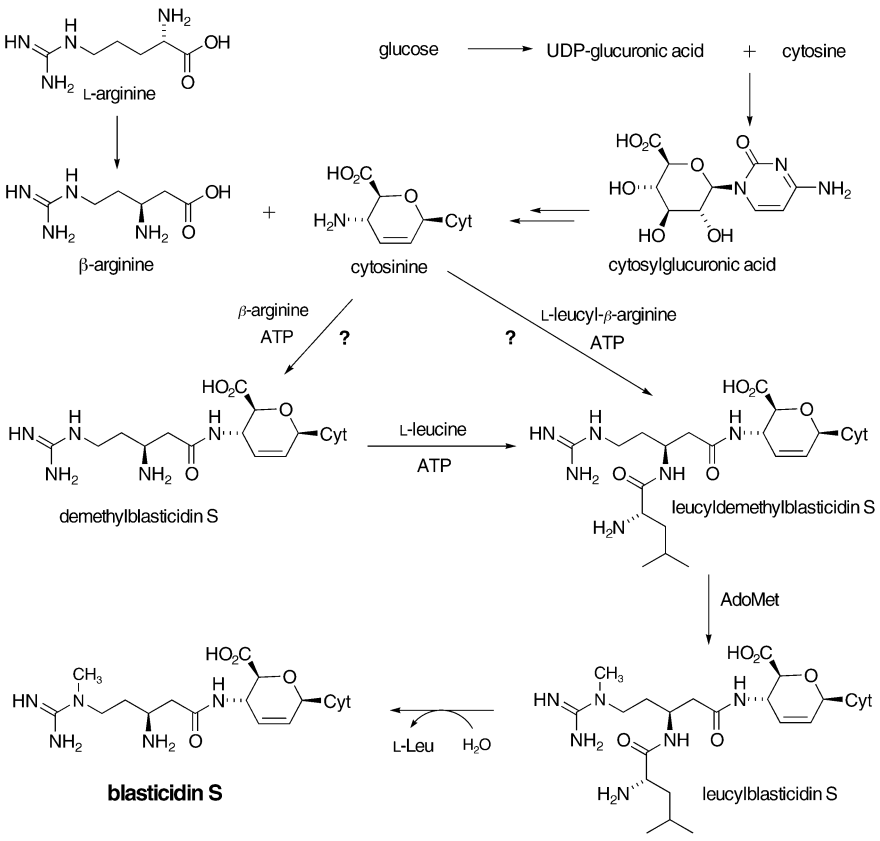






rice blast and was eventually produced on a large scale to support commercial agricultural use.[20] Blasticidin S exerts its cytotoxic action primarily through binding to the 50S ribosomal subunit, which results in inhibition of protein synthesis, and exhibits competitive binding with puromycin (**3**), a well-studied agent that affects the peptidyl transfer step on prokaryotic and eukaryotic ribosomes.[21, 22] Blasticidin S has recently found widespread use as a selectable marker in cell-culture applications when coupled with an acetyl transferase- or deaminase-resistance gene.[23–25]

The commercial importance and novel structural features of **1** prompted biosynthesis studies that are summarized in Scheme 1. Classical radiolabeled precursor studies established the primary precursors as glucose, cytosine, L-arginine, and methionine.[26] Through the use of both radioactive and stable isotope-labeled compounds, β-arginine, cytosylglucuronic acid (CGA), and cytosinine were identified as advanced intermediates in the pathway.[16, 27, 28] Efforts to unravel the biochemical steps in the assembly of **1** resulted in the purification and characterization of CGA synthase, the enzyme catalyzing the first committed step in the pathway, the condensation of cytosine and UDP-glucuronic acid (UDP, uridine diphosphate) to form CGA.[15] To further our investigations into the biosynthesis of **1**, we identified a 36-kb DNA fragment from an *S. griseochromogenes* genomic library in pOJ446 that contains the blasticidin S gene cluster and demonstrates heterologous production of CGA and other intermediates in the biosynthetic pathway in *S. lividans*.[18] We later revised the biosynthetic pathway to include leucylblasticidin S (LBS) as the penultimate compound in the pathway (Scheme 1).[17] Here we report the complete nucleotide sequence and organization of the blasticidin S biosynthesis gene cluster and describe the initial characterization of individual gene products.



Scheme 1. *Overview of blasticidin S biosynthesis.*

Results and Discussion

Sequence and organization of the biosynthesis gene cluster

Several segments of *S. griseochromogenes* chromosomal DNA cloned in the pOJ446 vector produced intermediates of blasticidin S biosynthesis when expressed in *S. lividans*.[17, 18] One cosmid, cos9 (Figure 1 A and B), produced late intermediates, including leucylblasticidin S (LBS), and was selected for DNA sequencing. Sequence analysis of the 36.7-kb cos9 insert with the Frameplot program[29] identified 28 open reading frames (ORFs) organized into three groups based on transcriptional direction (Figure 1 C). The first group consists of *blsA* to *blsN*. The second set of ORFs, *blsO* to *blsS*, is transcribed in the opposite direction. An additional nine ORFs between *blsS* and the 5′ end of the insert have mixed orientations and are not believed to participate in blasticidin S formation.

The boundaries of the *bls* gene cluster were defined by correlating metabolite expression profiles with restriction analysis of cosmids that directed blasticidin intermediate production in *S. lividans*. All cosmids lacking *blsS*, such as cos14, only produced early intermediates like cytosine and CGA (Figure 1 A). Cosmid clones con-

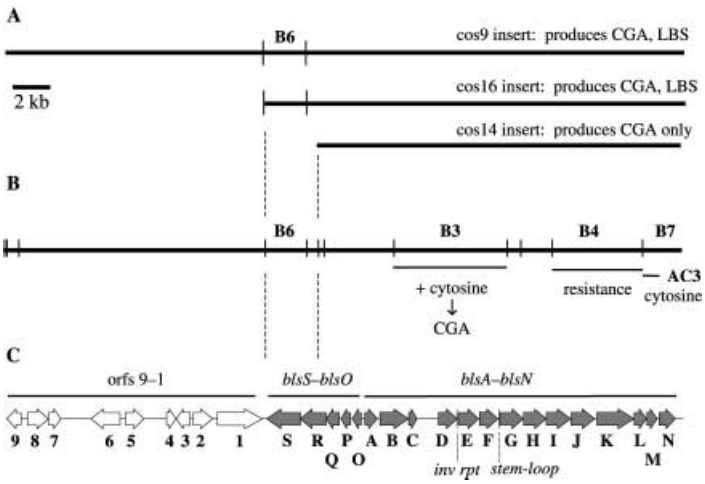


Figure 1. *A) Overlapping cosmids that produced blasticidin-S-related metabolites when expressed in* S. lividans*; B) partial restriction map of cos9 illustrating* Bam*HI cleavage sites and functions associated with particular fragments; C) organization of the fully sequenced cos9 insert containing the* bls *gene cluster (filled arrows).*

taining an intact B6 *Bam*HI fragment harboring *blsS* were able to produce later intermediates like LBS. This result suggests the entire blasticidin S cluster, with the possible exception of the final peptidase that cleaves the leucyl-β-arginine bond of LBS, is contained in an approximately 20-kb fragment of DNA.

Function of *bls* open reading frames deduced through sequence homology

Predicted functions for all 28 gene products identified on the cos9 insert (Figure 1 C) were assigned by BLAST analysis and exhibit varying degrees of sequence similarity to proteins in the public databases.[30] The proposed functions of each ORF and its closest homologues are presented in Table 1. Many of the ORFs encode products without significant sequence similarity to proteins of known function. Open reading frames with clear or demonstrated roles in the assembly of blasticidin S are detailed below.

Assembly and modification of the amino dideoxynucleoside core

Cytosylglucuronic acid (CGA) is the earliest committed intermediate in blasticidin S formation (Scheme 1). CGA synthase catalyzes the specific coupling of UDP-glucuronic acid and cytosine, and has been purified and characterized from wild-type *S. griseochromogenes*.[15] The sequence of *blsD* exhibits some similarity to glycosyltransferase genes and is predicted to encode CGA synthase. During efforts to express the entire blasticidin S pathway in *S. lividans*, we also expressed fragments of cos9 to identify regions coding for the production of separate precursor components. Cos9 *Bam*HI fragments greater than 1 kb were cloned into the *Streptomyces* expression vector pIJ702 and the resulting plasmids were used to transform *S. lividans*. When cultures of the *S. lividans* transformant expressing the 6.5-kb *Bam*HI fragment (B3) harboring *blsD* (Figure 1 B) were supplemented with cytosine, CGA was produced, which supports the assigned function of BlsD.[18]

An intriguing feature of blasticidin S biosynthesis is that free cytosine is incorporated at exceptionally high levels (around 95 %)[26] and the addition of exogenous cytosine can greatly increase CGA levels (up to 70-fold) while having a modest effect on blasticidin S production (1.6-fold).[44] Hence, it appears that the availability of free cytosine may be a limiting factor in blasticidin S biosynthesis. The intracellular pyrimidine bases exist almost exclusively at the nucleotide level and free cytosine is not a precursor to cytidine or the cytidine nucleotides cytidine monophosphate (CMP), diphosphate (CDP), and triphosphate (CTP). Rather, CTP is synthesized at the triphosphate stage directly from uridine triphosphate (UTP) by the action of CTP synthetase.[45] Therefore, we anticipated an activity associated with the *bls* cluster specifically for producing cytosine from cytidine nucleotides. Such a system is found in *Streptoverticillium rimofaciens*, the producer of mildiomycin, which contains hydroxymethylcytosine (HMC). A purified nucleotide hydrolase from this organism is specific for HMC monophosphate and does not hydrolyze CMP.[46] A nucleosidase acting on pyrimidine nucleosides and 2′-deoxynucleoside monophosphates from *S. virginiae* has been reported.[47]

When the 2.1-kb cos9 *Bam*HI fragment B7 (Figure 1 B) was expressed from pIJ702 in *S. lividans* a 7-fold increase in cytosine production was observed relative to the wild-type species. Restricting B7 with *Apa*I afforded an 826-bp fragment, AC3, containing only the intact *blsM* gene. Both B7 and AC3 were cloned into the constitutive expression site of a new shuttle expression vector pXY200 derived from elements of the *Streptomyces* expression vector pIJ4123[48] and the *Escherichia coli* vector pT7 – 7[49]. The resulting plasmids, pXY270 and pXY280, respectively, and a control lacking an insert were introduced into *S. lividans* by protoplast transformation and after six days the cytosine levels in the broths were determined by HPLC. Quantities of cytosine in the pXY270 and pXY280 transformants were 20 – 27-fold greater than in wild-type *S. lividans* (Table 2).

The *blsM* gene encodes a 174 amino acid protein similar to nucleoside 2′-deoxyribosyltransferase (Ndt), a nucleoside recycling enzyme first found in lactobacilli, and its orthologues and paralogues (COG 3613, NCBI). These enzymes catalyze cleavage of the glycosidic bond of 2′-deoxyribonucleosides by way of a covalent deoxyribosyl-enzyme intermediate. Most members of COG 3613 are between 150 and 190 amino acids in length and show little sequence homology except at key functional residues. Studies of the crystal structure of the *Lactobacillus leichmannii* enzyme and mutagenesis experiments implicate Glu98 as the active site residue that undergoes deoxyribosylation and suggest that Asp92 and/or Asp72 function as general acid catalysts, depending on the substrate.[50] Genes for two related enzymes, DRTaseI and DRTaseII, were recently cloned from *L. helveticus* CNRZ32. DRTaseI is specific for purine nucleosides as donors, while DRTaseII shows preference for pyrimidine nucleosides as donors and purine bases as acceptors.[51] In the absence of an acceptor, DRTaseII exhibits hydrolase activity and

Table 1. *Deduced functions of* bls *open reading frames.*

Protein	Amino acids	Proposed function	Sequence similarity (protein, origin)	Similarity, identity	Accession number	Ref.
Orf9	273 partial	ATP-binding protein	SCF43A.14, *Streptomyces coelicolor*	60%, 42%	CAB48901	[31]
Orf8	392	2-component system sensor kinase	SCH10.32c, *S. coelicolor*	46%, 34%	CAB42041	[31]
Orf7	222	2-component system response regulator	SCI41.37, *S. coelicolor*	71%, 48%	CAB59507	[31]
Orf6	522	secreted peptidase	SlpD, *S. coelicolor*	55%, 44%	CAB38476	[32]
Orf5	328	hydrolase	IpbD, *Psuedomona putida*	49%, 30%	AAC03446	[33]
Orf4	155	oxidoreductase	MitO, *S. lavendulae*	58%, 46%	AAD28457	[34]
Orf3	244	transcriptional regulator	SC5A7.19c *S. coelicolor*	43%, 33%	CAA19948	[31]
Orf2	343	unknown; pqqE/moaA family	NirJ *Methanosarcina mazei*	50%, 30%	NP633750	[35]
Orf1	909	serine/threonine protein kinase	SC7A12.07, *S. coelicolor*	39%, 29%	CAB94054	[31]
BlsS	610	oxidoreductase	Rv0492c, *M. tuberculosis* H37Rv	57%, 46%	Q11157	[36]
BlsR	479	unknown	Rv0493c, *M. tuberculosis* H37Rv	50%, 40%	Q11158	[36]
BlsQ	239	transcriptional regulator; GntR family	Rv0494, *M. tuberculosis* H37Rv	46%, 32%	Q11159	[36]
BlsP	159	3-helix membrane protein	Putative plasmid transfer protein *S. coelicolor*	60%, 40%	CAC36626	[31]
BlsO	156	2-helix membrane protein	SCF56.03, *S. coelicolor*	52%, 32%	CAB62748	[31]
BlsA	246	methyltransferase	Hypothetical protein, *Leuconostoc mesenteroides*	53%, 36%	ZP00063806	[37]
BlsB	513	carboxylesterase	Putative carboxylesterase, *S. avermitilis*	69%, 59%	BAB69209	[38]
BlsC	141	regulatory protein; yjgF family	SC5F1.25, *S. coelicolor*	57%, 42%	CAC16451	[31]
BlsD	328	cytosylglucuronic acid synthase	SCC75A.28c, *S. coelicolor*	42%, 28%	CAB61728	[31]
BlsE	344	unknown	MoaA, *S. coelicolor*	42%, 31%	CAB59437	[31]
BlsF	317	unknown	SCJ11.24c, *S. coelicolor*	43%, 31%	CAB52909	[31]
BlsG	410	arginine 2,3-aminomutase	KAM, *Clostridium subterminale*	66%, 48%	AAD43134	[39]
BlsH	392	aminotransferase	RifK, *Amycolatopsis mediterranei*	48%, 32%	AAC01720	[40]
BlsI	398	ligase	NikS, *Streptomyces tendae*	40%, 28%	CAC11141	[41]
BlsJ	414	self-resistance; transporter	SC8F4.05, *S. coelicolor*	37%, 26%	CAB70631	[31]
BlsK	579	Unknown	lysyl-tRNA synthetase, *S. coelicolor*	42%, 30%	CAD55312	[31]
BlsL	213	guanidino methyltransferase	Guanidinoacetate methyltransferase, *Rattus norvegicus*	45%, 27%	NP036925	[42]
BlsM	160	CMP hydrolase	2′-deoxyucleoside transferase, *Leuconostoc mesenteroides*	63%, 42%	ZP00063804	[43]
BlsN	264	unknown	SCG11A.10c, *S. coelicolor*	77%, 71%	CAB61591	[31]

releases 2′-deoxyribose. DRTaseI and DRTaseII share approximately 80% identity with NdtI and NdtII from *L. helveticus* ATCC 8018.[52, 53]

Figure 2 illustrates the alignment of nucleoside 2′-deoxyribosyltransferase active sites with the corresponding region of BlsM. A glutamate residue corresponding to position 98 of *L. leichmannii* Ndt is strictly conserved, as is Asp72. Interestingly, a serine residue replaces Asp92 in BlsM and Lmes1293, a hypothetical protein from *Leuconostoc mesenteroides* and the closest homologue to BlsM; a histidine residue occupies the same position in NdtII. Additional biochemical evidence will be required to confirm if these changes affect substrate specificity

Table 2. *Levels of cytosine production in* S. lividans *transformants expressing* blsM.

Plasmid	Cytosine [mg L^{-1}]
pIJ702/B7	48
pXY220 (control)	7
pXY270	145 – 152
pXY280	177 – 190

amount when BlsM was incubated with CMP (Figure 4), while dCMP and, to a lesser extent CDP, led to lower levels of cytosine. Unlike the nucleoside 2′-deoxyribosyltransferases to which BlsM shows active site similarity, BlsM does not accept cytidine as a substrate and the observed function is most similar to the hydroxymethylcytosine nucleotidase activity reported in *S. rimofaciens*.[46]

Expression of cosmids missing all or part of the B6 fragment (for example, cos14, Figure 1 A) did not yield detectable

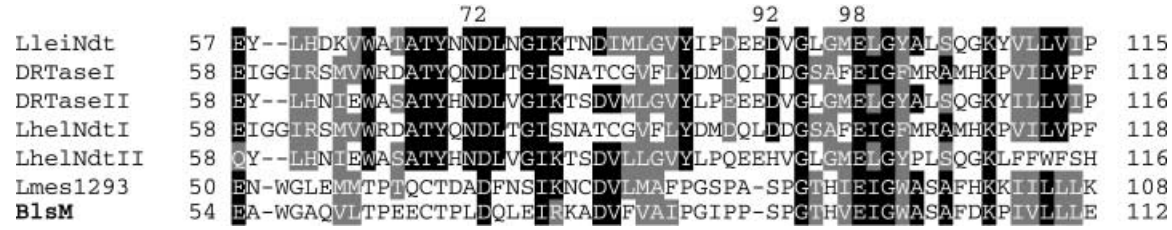


Figure 2. *Alignment of BlsM with known nucleoside 2′-deoxyribosyltransferase active sites. LleiNdt is from* L. leichmannii; *DRTaseI and DRTaseII are from* L. helveticus *CNRZ32; LhelNdtI and II are from* L. helveticus *ATCC 8018, and Lmes1293 is from* L. mesenteroides.

and hydrolase versus transferase activity. A final notable feature of the *blsM* gene is the presence of a TTA codon, which makes the translation dependent on cellular levels of the *bldA* gene product, a specific tRNA^{Leu} that coordinates the events of antibiotic biosynthesis with the development cycle in many *Streptomyces* species.[54] A similar situation is found in the puromycin pathway.[55]

To confirm the function of BlsM in vitro and explore the scope of the reaction, *blsM* was amplified from cos9 and cloned into the pET41a + vector to yield a construct encoding a glutathione-*S*-transferase (GST)/BlsM fusion protein that possesses a C-terminal His_6 tag. The GST/BlsM fusion protein was expressed in *E. coli*, purified, GST was removed by thrombin cleavage, and soluble BlsM was isolated by Co^{2+} affinity chromatography (Figure 3).

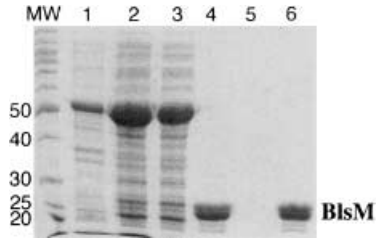


Figure 3. *SDS-PAGE analysis of GST/BlsM fusion protein overexpressed in* E. coli, *and purification of BlsM. Lane 1, insoluble fraction; Lane 2, cleared cell lysate; Lane 3, unretained fraction from GST-Bind (Novagen) affinity column; Lane 4, fraction released from GST-Bind column after thrombin treatment; Lane 5, unretained fraction from Co^{2+} affinity column; Lane 6, BlsM eluted from the Co^{2+} affinity column.*

To identify the natural substrate(s) of BlsM, various cytosine nucleosides or nucleotides were incubated with BlsM in 50 mM sodium phosphate buffer at pH 7.0 and 37 °C for 1 hr. The product and substrate were separated and quantified by reverse-phase HPLC. Free cytosine was generated in the greatest

intermediates beyond CGA. Partial and complete genes contained on B6, *blsR* and *blsS*, are predicted to encode a protein of unknown function and a glucose-methanol-choline (GMC) oxidoreductase homologue, respectively, and are likely involved in the elaboration of CGA to an amino deoxynucleoside such as cytosinine (Scheme 1). Along with *blsQ*, these genes are highly similar to, and occur in the same order as three uncharacterized genes found in the *Mycobacterium tuberculosis* H37Rv genome (Table 1).[36] Both Rv0494 and BlsQ are predicted to contain a conserved domain that places them in the FadR subgroup of the gntR family of DNA-binding transcriptional regulatory proteins.[56]

The product of *blsH* is also predicted to function in amino deoxynucleoside formation. BlsH exhibits closest similarity to aminotransferases such as the pyridoxyl phosphate-dependent perosamine synthase[57] and RifK from *Amycolatopsis mediterranei* S699.[40] The latter similarity is interesting because RifK was shown to function as the 3-amino-5-hydroxybenzoic acid (AHBA) synthase in rifamycin biosynthesis and more recently was found to also operate in tandem with the NAD^+-dependent dehydrogenase RifL in the conversion of UDP-glucose to UDP-kanosamine.[58, 59]

Formation, attachment, and methylation of β-arginine

The β-arginine moiety found in blasticidin S was shown by Seto's group to originate from L-arginine.[26] Stable isotope labeling studies by Gould et al. established an intramolecular migration of the α-nitrogen, as previously observed for β-lysine formation catalyzed by lysine-2,3-aminomutase from *Clostridium subterminale* SB4.[60] Extensive work by Frey's group revealed that native lysine 2,3-aminomutase (KAM) is a hexamer of 48-kDa subunits possessing three [4Fe – 4S] centers and six molecules of pyridoxal phosphate (PLP), requires AdoMet and is O_2-sensitive.[61a] BLAST analysis reveals the predicted product of *blsG* is 48% identical and 65% similar to *C. subterminale* KAM, and includes

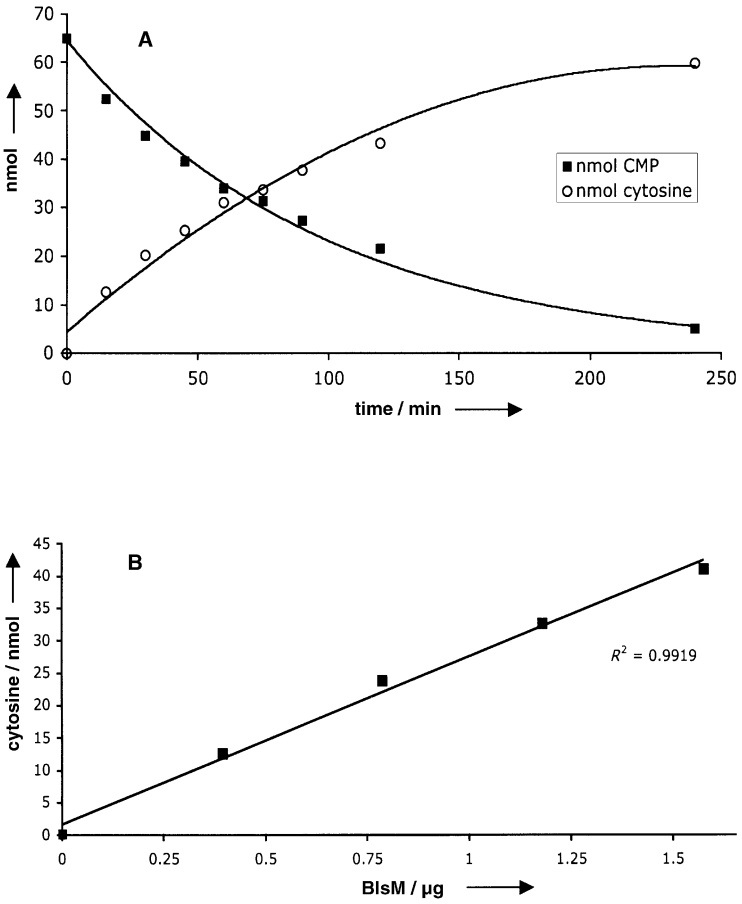


Figure 4. *Conversion of CMP to cytosine catalyzed by BlsM. A) Time-dependent formation of cytosine and concomitant loss of CMP. B) Increased production of CMP as a function of BlsM concentration.*

the conserved lysine for PLP attachment. We have expressed and purified BlsG from *E. coli* and *S. lividans* but have not detected aminomutase activity[61b].

The process for attaching β-arginine to the nucleoside core of blasticidin S has been of considerable interest to us. Prior to obtaining the DNA sequence of cos9, attempts to amplify nonribosomal peptide synthetase (NRPS) gene fragments from cos9 or detect β-arginine carboxy activation in cell-free extracts failed. Similar efforts were fruitless when we attempted to discern how leucine was incorporated into leucylblasticidin S.[17] Mechanisms for amino acid incorporation by other peptidyl nucleosides vary. Genes similar to NRPSs are present in the streptothricin F cluster from *S. rochei*.[3, 4] More detailed investigation of how β-lysine is incorporated into streptothricins (also known as nourseothricins) in *S. noursei* identified a stand-alone adenylating enzyme (NpsA) that specifically activates β-lysine for transfer to an N-terminal peptidyl carrier protein (PCP) domain on NpsB.[14] In the case of puromycin, attachment of the tyrosine moiety to the nucleoside core is predicted to be mediated by Pur6, a protein that shares little resemblance to NRPSs and lacks an obvious AMP binding domain.[10] NikS is a ligase in the nikkomycin pathway that activates 4-pyridyl-2-oxo-4-hydroxyisovaleric acid (POHIV) as the acylphosphate and belongs to the ATP-grasp-fold superfamily of enzymes.[8] Members of this superfamily are characterized by a unique ATP binding structure and include D-Ala-D-Ala ligase, glutathione synthetase, and several CoA ligases, all of which activate their substrate carboxy groups as acylphosphates, rather than acyladenylates. The product of *blsI* shares the ATP-grasp fold and is predicted to catalyze the coupling of peptidyl and nucleoside moieties in blasticidin S. Possible roles for BlsI include activating β-arginine or leucyl-β-arginine for coupling with an amino deoxynucleoside, or activating leucine for attachment to β-arginine or demethylblasticidin S to give the dipeptide or demethylleucylblasticidin S, respectively (Scheme 1). Precedent for these ligases activating a dipeptide carboxy group is found in the cases of the D-Ala-D-Ala adding enzyme (MurF) in peptidoglycan biosynthesis[62] and glutathione synthetase.[63] The intermediacy of leucylblasticidin S in the pathway requires the involvement of two amino acid activating enzymes; one to activate leucine and one to activate β-arginine or leucyl-β-arginine. The best candidate for the second sort of protein is BlsK, which exhibits modest similarity to a putative lysyl-tRNA synthetase from *S. coelicolor*.

The penultimate step in the biosynthesis of **1** is methylation of the β-arginine guanidine residue.[17] BlsL is similar to a number of guanidoacetate methyltransferases and is the best candidate for carrying out the final constructive step in the pathway.

Self-resistance

Cosmids harboring the blasticidin S gene cluster were originally identified by hybridization with a 4.8-kb *Bam*HI fragment of *S. griseochromogenes* chromosomal DNA that conferred resistance to the antibiotic on *S. lividans*.[18] This fragment is represented as B4 in Figure 1 and contains *blsJ*. BlsJ is predicted to contain 11 membrane-spanning domains and is similar to a number of *S. coelicolor* proteins thought to be involved in metabolite efflux and transport. *S. lividans* resistance to blasticidin S was increased when a plasmid carrying a gene for an ATP-binding protein was coexpressed with the B4 fragment, which indicates that BlsJ is probably a component of an ABC transporter.[64]

Experimental Section

General methods and materials: The cloning of the blasticidin S gene cluster from *S. griseochromogenes*, conditions for heterologous expression in *S. lividans* TK24, and conditions for HPLC analysis have been previously reported.[18] Vectors used in these studies include: *Streptomyces* vectors pIJ702 (obtained from Prof. D. A. Hopwood, John Innes Institute) and pIJ2925 (a gift from Dr. M. J. Virolle, Université Paris-Sud), pBluescript II KS(+) (Stratagene), and pGEM11zf (Promega). *E. coli* strains DH5α (Life Technologies, Inc.) and JM109 (Promega) were routinely used as hosts for *E. coli* plasmids and *E. coli – Streptomyces* shuttle vectors. The new *E. coli – Streptomyces* shuttle expression vector pXY200 was constructed by replacing the kanamycin resistance marker present in pIJ4123[48] with the apramycin selection marker *accC4*[65] and introducing an intact ColE1 origin and ampicillin resistance marker (*bla*) derived from the *E. coli*

expression vector pT7 – 7.[49] All *E. coli* manipulations were performed according to standard protocols.[66] Standard media and methods of culture for *Streptomyces* are described in Kieser et al.[67] Restriction endonucleases, DNA ligase, DNA polymerase, and alkaline phosphatase were purchased from various sources and used according to the manufacturers' recommendations.

DNA sequence analysis: A cosmid clone (cos9) carrying a 36.7-kb insert of *S. griseochromogenes* genomic DNA in pOJ446 and able to direct LBS production in *S. lividans* was selected for sequencing.[18] Digestion with *Bam*HI yielded 13 fragments, each of which was subcloned in pBluescript II KS (+). Smaller inserts, ranging from 0.15 – 4.88 kb, were sequenced by a combination of delta-subcloning and primer walking at the Center for Gene Research and Biotechnology at Oregon State University by using the Amplitaq dye-terminator sequencing system (Perkin Elmer) and Applied Biosystems automated DNA sequencers. The two largest *Bam*HI fragments (B1, 13.3 kb and B3, 6.2 kb) were sequenced by MWG Biotech Inc. (High Point, NC). Nucleotide sequences were determined for both strands. Sequence analysis was carried out with MacVector (Oxford Molecular Group), VectorNTI (Informax), and FramePlot[29] software packages. Nucleotide and amino acid sequence similarity comparisons were carried out in public databases by using the BLAST (basic local alignment search tool) program.[30] The DNA sequence of cos9 has been deposited in GenBank; accession number: AY196214.

Heterologous expression of cos9 subfragments: *Bam*HI fragments of cos9 cloned in pBluescript II KS (+) that have more than 1 kb were excised with *Bam*HI, gel purified, and ligated with *Bgl*II restricted pIJ702. The resulting plasmids, and an empty vector control, were individually introduced into *S. lividans* by protoplast transformation. Transformants were cultured and analyzed for blasticidin S precursors as previously described.[18]

The 2.1-kb *Bam*HI fragment from cos9 (B7) was also cloned in pIJ2925 linearized with *Bam*HI. The B7 insert was removed by restriction with *Hin*dIII and *Eco*RI, gel purified, and ligated with *Hin*dIII – *Eco*RI-digested pXY200 to yield pXY270. The B7 fragment was also restricted with *Apa*I and the 0.83-kb fragment (AC3) was cloned into pGEM11zf. The AC3 fragment was then excised by treatment with *Hin*dIII – *Eco*RI and cloned into similarly restricted pXY200 to yield pXY280. Plasmids pXY270, pXY280, and the empty vector control were introduced into *S. lividans* by protoplast transformation. Transformants were cultured and analyzed for cytosine as previously described.[18]

Preparation of a *blsM* expression construct: The *blsM* gene was amplified from cos9 by PCR with the forward primer 5′-ctcgaggg<u>aattc</u>ggtgcgcagcgtctttctcgccggt-3′ (*Eco*RI site underlined) and the reverse primer, 5′-acgatgcggtgcacggttcgg<u>ctcgag</u>cg-3′ (*Xho*I site underlined). The PCR was carried out in a volume of 50 μL containing approximately 10 ng cos9 as template, 1X Promega Thermophilic Buffer, $MgCl_2$ (1.25 mM), dNTP Mix (0.4 mM, MBI) and dimethylsulfoxide (5%) and Taq DNA polymerase (5 units, MBI). PCR products were purified, digested with *Eco*RI and *Xho*I, and inserted into a similarly restricted pET41a+ vector. The resulting plasmid, pET41/BlsM, was used to transform *E. coli* JM109 cells and then transferred to Rosetta (DE3)pLysS competent *E. coli* cells (Novagen) for expression.

Overproduction and purification of BlsM: Transformants carrying pET41/BlsM were grown overnight in Terrific broth supplemented with chloramphenicol and kanamycin (50 $\mu g\,mL^{-1}$ each). This seed culture (5 mL) was used to inoculate Terrific broth (1 L) supplemented to a final concentration of 50 $\mu g\,mL^{-1}$ chloramphenicol and 50 $\mu g\,mL^{-1}$ kanamycin. The cells were grown at 37 °C to an optical density at 600 nm (OD_{600}) of 0.55 and then induced with isopropyl-β-D-thiogalactopyranoside (final concentration, 0.4 mM). The culture was then grown for an additional 8 h at 37 °C. After centrifugation, the cells were resuspended in phosphate buffered saline (PBS; 50 mL, containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4), pH 7.3. Cells were lysed by sonication (7 × 60 sec at 10 watts, with 1 min cooling on ice between bursts) and cellular debris was pelleted by centrifugation (18000 × g, 40 min, 4 °C). The soluble GST – BlsM fusion protein was applied to a glutathione affinity column (GST-Bind, Novagen) and the resin-bound complex was treated with thrombin (10 U) for 18 h at room temperature. Free BlsM was eluted from the column and further purified by Co^{2+} affinity chromatography (Talon resin, BD Biosciences). The BlsM was eluted in imidazole (8 mL, 150 mM) in PBS at pH 7.3, followed by dialysis against PBS (4 L, pH 7.3). The purified BlsM was diluted to 400 $\mu g\,mL^{-1}$ with PBS at pH 7.3 and glycerol (20%). The BlsM was shown to retain equal activity when stored at either −80 °C or 4 °C for 15 days.

Identification of the BlsM substrate: Enzyme assays were carried out in a total volume of 100 μL containing sodium phosphate buffer (50 mM, pH 7.0) and BlsM (1.2 ng) at 37 °C for 1 hr. Various cytosine nucleosides or nucleotides (100 mM) were evaluated as substrates. The reactions were quenched by boiling for 5 min. After centrifugation (3 min at 18000 × g), the product and substrate were separated and quantified by reverse-phase HPLC (Beckman Ultrasphere C_{18} 4.6 mm × 25 cm column; isocratic mobile phase containing 200 mM sodium phosphate buffer (pH 5.3)/acetonitrile, 95:5) with UV detection at 260 nm. For each assay, 50 μL was loaded onto a 100-μL injection coil for each run.

Assays to confirm the time-dependent formation of cytosine were carried out in 650 μL total volume containing $NaPO_4$ (50 mM, pH 7.0), BlsM (7.8 ng) and CMP (100 μM). An aliquot (65 μL) was removed at each time point and quenched and analyzed as described above.

We thank Prof. David Hopwood for plasmid pIJ702, Prof. Jean-Luc Pernodet for pHP45Ωacc(4), Dr. Marie-Joëlle Virolle for pIJ2925, and Dr. Yukio Miyazaki of Kaken Chemical Company, Ltd. for the culture of S. griseochromogenes *and generous samples of blasticidin S. Dr. Steven J. Gould is thanked for his help and interest in this work. Partial support came from the National Institutes of Health (Grant nº GM32110) and the Northwest Health Foundation.*

[1] C. Bormann, V. Mohrle, C. Bruntner, *J. Bacteriol.* **1996**, *178*, 1216 – 1218.
[2] R. A. Lacalle, J. A. Tercero, A. Jimenez, *EMBO J.* **1992**, *11*, 785 – 792.
[3] M. A. Fernández-Moreno, C. Vallín, F. Malpartida, *J. Bacteriol.* **1997**, *179*, 6929 – 6936.
[4] J. Propin, M. Fernandez-Moreno, F. Malpartida, **2000**, direct submission to GenBank, Accession no. AJ271405.
[5] C. Bruntner, B. Lauer, W. Schwarz, V. Mohrle, C. Bormann, *Mol. Gen. Genet.* **1999**, *262*, 102 – 114.
[6] H. Chen, B. K. Hubbard, S. E. O'Connor, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 103 – 112.
[7] C. Bruntner, C. Bormann, *Eur. J. Biochem.* **1998**, *254*, 347 – 355.
[8] B. Lauer, R. Russwurm, W. Schwarz, A. Kalmanczhelyi, C. Bruntner, A. Rosemeier, C. Bormann, *Mol. Gen. Genet.* **2001**, *264*, 662 – 673.
[9] B. Lauer, R. Süssmuth, D. Kaiser, G. Jung, C. Bormann, *J. Antibiot.* **2000**, *53*, 385 – 392.
[10] J. A. Tercero, J. C. Espinosa, R. A. Lacalle, A. Jiménez, *J. Biol. Chem.* **1996**, *271*, 1579 – 1590.
[11] M. A. Rubio, J. C. Espinosa, J. A. Tercero, A. Jimenez, *FEBS Lett.* **1998**, *437*, 197 – 200.
[12] J. C. Espinosa, J. A. Tercero, M. A. Rubio, A. Jimenez, *J. Bacteriol.* **1999**, *181*, 4914 – 4918.

[13] J. A. Tercero, R. A. Lacalle, A. Jimenez, *Eur. J. Biochem.* **1993**, *218*, 963 – 971.
[14] N. Grammel, K. Pankevych, J. Demydchuk, K. Lambrecht, H. P. Saluz, H. Krugel, *Eur. J. Biochem.* **2002**, *269*, 347 – 357.
[15] S. J. Gould, J. Guo, *J. Bacteriol.* **1994**, *176*, 1282 – 1286.
[16] S. J. Gould, Q. Zhang, *J. Antibiot.* **1995**, *48*, 652 – 656.
[17] Q. Zhang, M. C. Cone, S. J. Gould, T. M. Zabriskie, *Tetrahedron* **2000**, *56*, 693 – 701.
[18] M. C. Cone, A. K. Petrich, S. J. Gould, T. M. Zabriskie, *J. Antibiot.* **1998**, *51*, 409 – 417.
[19] S. Takeuchi, K. Hirayama, K. Ueda, H. Sakai, H. Yonehara, *J. Antibiot.* **1958**, *11*, 1 – 5.
[20] T. Misato, I. Ishii, M. Asakawa, Y. Okimoto, K. Fukunaga, *Ann. Phytopathol. Soc. Jpn.* **1959**, *24*, 302 – 306.
[21] H. Yamaguchi, C. Yamamoto, N. Tanaka, *J. Biochem.* **1965**, *57*, 667 – 677.
[22] H. Yamaguchi, N. Tanaka, *J. Biochem.* **1965**, *60*, 632 – 642.
[23] M. Izumi, H. Miyazawa, T. Kamakura, I. Yamaguchi, T. Endo, F. Hanaoka, *Exp. Cell Res.* **1991**, *197*, 229 – 233.
[24] C. Karreman, *Nucleic Acids Res.* **1998**, *26*, 2508 – 2510.
[25] C. B. Mamoun, I. Y. Gluzman, S. Goyard, S. M. Beverley, D. E. Goldberg, *Proc. Natl. Acad .Sci. USA* **1999**, *96*, 8716 – 8720.
[26] H. Seto, I. Yamaguchi, N. Otake, H. Yonehara, *Agric. Biol. Chem.* **1968**, *32*, 1292 – 1298.
[27] P. C. Prabhakaran, N. T. Woo, P. S. Yorgey, S. J. Gould, *J. Am. Chem. Soc.* **1988**, *110*, 5785 – 5791.
[28] J. Guo, S. J. Gould, *J. Am. Chem. Soc.* **1991**, *113*, 5898 – 5899.
[29] J. Ishikawa, K. Hotta, *FEMS Microbiol. Lett.* **1999**, *174*, 251 – 253.
[30] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403 – 410.
[31] S. D. Bentley, K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, *Nature* **2002**, *417*, 141 – 147.
[32] M. Redenbach, H. M. Kieser, D. Denapaite, A. Eichner, J. Cullum, H. Kinashi, D. A. Hopwood, *Mol. Microbiol.* **1996**, *21*, 77 – 96.
[33] R. W. Eaton, K. N. Timmis, *J. Bacteriol.* **1986**, *168*, 123 – 131.
[34] Y. Mao, M. Varoglu, D. H. Sherman, *Chem. Biol.* **1999**, *6*, 251 – 263.
[35] U. Deppenmeier, A. Johann, T. Hartsch, R. Merkl, R. A. Schmitz, R. Martinez-Arias, A. Henne, A. Wiezer, S. Baumer, C. Jacobi, H. Bruggemann, T. Lienard, A. Christmann, M. Bomeke, S. Steckel, A. Bhattacharyya, A. Lykidis, R. Overbeek, H. P. Klenk, R. P. Gunsalus, H. J. Fritz, G. Gottschalk, *J. Mol. Microbiol. Biotechnol.* **2002**, *4*, 453 – 461.
[36] S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, L. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M.-A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, B. G. Barrell, *Nature* **1998**, *393*, 537 – 544.
[37] Direct submission to GenBank, NCBI Microbial Genomes Annotation Project, **2002**.
[38] S. Omura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12215 – 12220.
[39] F. J. Ruzicka, K. W. Lieder, P. A. Frey, *J. Bacteriol.* **2000**, *182*, 469 – 476.
[40] C. G. Kim, T. W. Yu, C. B. Fryhle, S. Handa, H. G. Floss, *J. Biol. Chem.* **1998**, *273*, 6030 – 6040.
[41] B. Lauer, R. Russwurm, C. Bormann, *Eur. J. Biochem.* **2000**, *267*, 1698 – 1706.
[42] H. Ogawa, T. Date, T. Gomi, K. Konishi, H. C. Pitot, G. L. Cantoni, M. Fujioka, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 694 – 698.
[43] Direct submission to GenBank, NCBI Microbial Genomes Annotation Project, **2002**.
[44] S. J. Gould, J. Guo, K. Dejesus, A. Geitmann, *Can. J. Chem.* **1994**, *72*, 6 – 11.
[45] G. A. O'Donovan, J. Neuhard, *Bacteriol. Rev.* **1970**, *34*, 279 – 343.
[46] H. Sawada, T. Suzuki, S.-I. Akiyama, Y. Nakao, *J. Ferment. Technol.* **1985**, *63*, 23 – 27.
[47] A. Imada, S. Igarasi, *J. Bacteriol.* **1967**, *94*, 1551 – 1559.
[48] E. Takano, J. White, C. J. Thompson, M. J. Bibb, *Gene* **1995**, *166*, 133 – 137.
[49] a) S. Tabor, C. C. Richardson, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1074 – 1078; b) Yin et al., unpublished work.
[50] S. A. Short, S. R. Armstrong, S. E. Ealick, D. J. Porter, *J. Biol. Chem.* **1996**, *271*, 4978 – 4987.
[51] P. A. Kaminski, *J. Biol. Chem.* **2002**, *277*, 14400 – 14407.
[52] K. Okuyama, T. Noguchi, direct submission to GenBank, Accession No. AB039914.
[53] K. Okuyama, T. Noguchi, *Biosci. Biotechnol. Biochem.* **2000**, *64*, 2243 – 2245.
[54] E. J. Lawlor, H. A. Baylis, K. F. Chater, *Genes Dev.* **1987**, *1*, 1305 – 1310.
[55] J. A. Tercero, J. C. Espinosa, A. Jimenez, *FEBS Lett.* **1998**, *421*, 221 – 223.
[56] S. Rigali, A. Derouaux, F. Giannotta, J. Dusart, *J. Biol. Chem.* **2002**, *277*, 12507 – 12515.
[57] W. C. Nierman, T. V. Feldblyum, M. T. Laub, I. T. Paulsen, K. E. Nelson, J. A. Eisen, J. F. Heidelberg, M. R. Alley, N. Ohta, J. R. Maddock, I. Potocka, W. C. Nelson, A. Newton, C. Stephens, N. D. Phadke, B. Ely, R. T. DeBoy, R. J. Dodson, A. S. Durkin, M. L. Gwinn, D. H. Haft, J. F. Kolonay, J. Smit, M. B. Craven, H. Khouri, J. Shetty, K. Berry, T. Utterback, K. Tran, A. Wolf, J. Vamathevan, M. Ermolaeva, O. White, S. L. Salzberg, J. C. Venter, L. Shapiro, C. M. Fraser, J. Eisen, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4136 – 4141.
[58] K. Arakawa, R. Müller, T. Mahmud, H. G. Floss, *J. Am. Chem. Soc.* **2002**, *124*, 10644 – 10645.
[59] J. Guo, J. W. Frost, *J. Am. Chem. Soc.* **2002**, *124*, 10642 – 10643.
[60] D. J. Aberhart, S. J. Gould, H.-J. Lin, T. K. Thiruvengadam, B. H. Weiller, *J. Am. Chem. Soc.* **1983**, *105*, 5461 – 5470.
[61] a) P. A. Frey in *Comprehensive Natural Products Chemistry, Volume 5*, (Eds.: K. Nakanishi, D. Barton), Elsevier, New York, 205 – 233; b) Grochowski and Zabriskie, unpublished results.
[62] T. D. Bugg, C. T. Walsh, *Nat. Prod. Rep.* **1992**, *9*, 199 – 215.
[63] S. C. Lu, *FASEB J.* **1999**, *13*, 1169 – 1183.
[64] C. Méndez, J. A. Salas, *FEMS Microbiol. Lett.* **1998**, *158*, 1 – 8.
[65] M. H. Blondelet-Rouault, J. Weiser, A. Lebrihi, P. Branny, J. L. Pernodet, *Gene* **1997**, *190*, 315 – 317.
[66] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning. A Laboratory Manual* 2nd ed., Cold Spring Harbor Laboratory Press, New York, **1989**.
[67] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, The John Innes Foundation, Norwich, **2000**.

Received: February 20, 2003
Revised version: June 10, 2003 [F 583]

