Study of the bldG locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation

Dawn R. D. Bignell,1 Jason L. Warawa,1 Janice L. Strap,1 Keith F. Chater2 and Brenda K. Leskiw1

Author for correspondence: Brenda K. Leskiw. Tel: +1 780 492 1868. Fax: +1 780 492 9234. e-mail: brenda.leskiw@ualberta.ca

A cloned 2.5 kb DNA fragment that can restore antibiotic production and sporulation to a bldG mutant encodes a 113 aa protein showing similarity to a family of anti-anti-sigma factors from *Bacillus* and *Staphylococcus*; and the deduced product of a closely spaced downstream ORF, designated ORF3, shows similarity to cognate anti-sigma factors. The homologues in *Bacillus* regulate the activity of sporulation- and stress-response-specific sigma factors. However, there is no sigma factor gene near bldG and ORF3. bldG is transcribed both as a monocistronic and a polycistronic mRNA, the latter including the downstream ORF3 gene. The two transcripts were present at all time points during growth and both were upregulated when aerial mycelium and pigmented antibiotics were seen. At all time points, the monocistronic bldG transcript was two- to threefold more abundant than the polycistronic transcript. Mapping of the mRNA 5’ ends indicated that bldG transcription is initiated from two transcription start sites located 82 and 123 bp upstream of the bldG translation start. A constructed bldG null mutant had the same phenotype as previously isolated bldG point mutations, some of which were shown to have potentially significant base changes within bldG. When compared to the wild-type strain, the null mutant showed no differences in the levels of transcription from the two bldG promoters. These results suggest that bldG is not involved in autoregulation.

Keywords: *Streptomyces*, differentiation, antibiotic production, anti-sigma factor antagonist, anti-sigma factor

INTRODUCTION

Members of the genus *Streptomyces* are soil bacteria best known for their ability to produce a wide variety of secondary metabolites including many of the antibiotics used in human and veterinary medicine. In solid-phase laboratory culture, the production of these compounds coincides with the onset of morphological differentiation that leads to sporulation: aerial hyphae grow from the older parts of the substrate mycelium, then proceed through a series of changes that include nucleoid partitioning coupled to the formation of sporulation septa. The end result of this complex process is the formation of long chains of unigenomic spores (reviewed by Chater, 1998; Kelemen & Buttner, 1998). Differentiation is under the control of an elaborate regulatory system that coordinates gene-expression pathways to bring about the gross morphological changes in both the substrate and aerial hyphae. In addition, the onset of antibiotic production (physiological differentiation) is intimately associated with the onset of morphological differentiation, and these two processes are now known to share common elements in genetic regulation.

In *Streptomyces coelicolor* A3(2), the genetically most-studied streptomycete, many genes required for morphological and physiological differentiation have been identified. As a result of such studies we know that the regulatory mechanisms involve both pathway-specific and global regulatory genes (Champness, 2000). Of...
particular note are the bld (bald) genes, mutations in which often result in the failure to form both aerial hyphae and secondary metabolites. Thus the bld genes provide a key to understanding how the physiological and structural changes are initiated in the ageing substrate mycelium.

Several bld genes have been cloned, sequenced and characterized. This, together with functional analysis of several other genes, in which mutations give rise to a bald mutant phenotype but which had previously been given other gene designations (Chakraburty & Bibb, 1997; Ma & Kendall, 1994; Susstrunk et al., 1998), is beginning to paint a picture whereby the onset of physiological and morphological differentiation is governed by extracellular and metabolic signals, and is regulated at the level of transcription, translation and post-translational modification. When S. coelicolor is grown on rich media, Willey et al. (1991) showed that structural differentiation to form aerial mycelium depends on an extracellular protein, SapB. Although SapB production was impaired in all of the bld mutants tested, juxtaposition of pairs of bld mutants on the surface of agar plates resulted in the unidirectional restoration of both SapB and aerial-hyphae formation, suggesting that differentiation is governed by a hierarchical cascade of intercellular signals, with the bld genes themselves directly or indirectly responsible for the production of those signals (Willey et al., 1993). Recent studies by Nodwell et al. (1996) have provided biochemical evidence in support of this extracellular signalling model; the bldK locus encodes the subunits for an ATP-binding cassette family, oligopeptide permease that appears to be the importer of a bldJ (formerly bld261)-dependent, extracellular factor (Nodwell & Losick, 1998). However, acceptance of this model is complicated by the observation that not all of the bld genes fit into the proposed hierarchy (Nodwell et al., 1999; Willey et al., 1993). Of particular note is bldB, which encodes a putative DNA-binding regulatory protein and mutations in which, along with many of the other bld mutations (blda, -B, -C, -D, -G and -H), causes a deregulation of carbon utilization (Pope et al., 1996). That these mutants activate the gal operon promoter under conditions where it is normally repressed, is consistent with the idea that a change in metabolic state is associated with the onset of aerial mycelium formation (Karandikar et al., 1997; Nodwell et al., 1999). This idea is further supported by the finding by Susstrunk et al. (1998) that mutations in the adenylate cyclase gene, cya, which result in a classic bald phenotype, also cause a medium pH decrease, suggesting that the onset of differentiation involves a switch in metabolism to utilize organic acids released to the medium.

To formulate a model for the regulation of differentiation that takes into account all of these observations, it would be valuable to understand the function of all of the bld gene products. To this end we have cloned and sequenced the S. coelicolor bldG gene. bldG was first identified in a study in which mutantized colonies were screened for blocks in the formation of aerial hyphae and antibiotic biosynthesis (Champness, 1988). Three of the mutants identified, C103, C107 and C101J, mapped to the same region of the chromosome and defined the bldG locus. Here we show that bldG encodes a protein product showing similarity to anti-sigma factor antagonists from Bacillus and Staphylococcus. Transcriptional analyses have also revealed that bldG and the ORF immediately downstream are co-transcribed. The deduced product of the downstream ORF shows some similarity to anti-sigma factors, suggesting that it and BldG might function as a regulatory pair governing the activity of an unknown sigma factor(s).

METHODS

Strains and media. S. coelicolor strains used in this study include J1501 [hisA1 uraA1 strA1 pgI SCP1- SCP2- (Chater et al., 1982)] and its derivatives C103 [bldG103 (Champness, 1988)], C107 [bldG107 (Champness, 1988)] and C536 [bldG536; W. Champness, personal communication]; C101J [cysD18 mtbB2 NF, bldG101J (Champness, 1988)] and M145 [prototrophic, SCP1- SCP2- Pgl+ (Hopwood et al., 1985) provided by M. Bibb, John Innes Centre]. C103, C107, C536 and C101J were all provided by W. Champness, Michigan State University. Streptomyces lividans 66 (John Innes strain 1326) was the host for pC31 propagation and for the transfection of protoplasts. Media, culture conditions and protoplast transformation and transfection were as described by Hopwood et al. (1985).

Escherichia coli host strains were DH5α (Gibco-BRL) and ET12567 (MacNeil et al., 1992) (a gift from D. MacNeil, Merck Sharp & Dohme Research Laboratories). Media and culture conditions were as described by Sambrook et al. (1989).

Plasmid and bacteriophage vectors. Streptomyces vector KC304 is a derivative of bacteriophage oC31 and contains the tsr (thiostrepton resistance) gene for vector selection; the rph (viomycin resistance) gene, flanked by BamHI sites, as a ‘stuffer’ fragment for replacement by up to 6 kb of insert DNA; and the att–int region to allow efficient integration at single copy number into the chromosomal att site for oC31. KC304 derivatives were manipulated as described by Hopwood et al. (1987). Cosmid H5 (Redenbach et al., 1996) was obtained from Helen Kieser, John Innes Centre, Norwich, UK. Prior to their use to transform S. coelicolor, the EcoR–Streptomyces bifunctional vectors pSET152 (Bierman et al., 1992) (NRRL B-14792) (containing a multiple-cloning site and replication of pUC plasmids, the att–int region of oC31 and the apramycin-resistance gene for vector selection in either E. coli or Streptomyces) and pKC1218 (Bierman et al., 1992) (NRRL B-14790) (containing a multiple-cloning site and replication of pUC plasmids, the SCP2- replication and the apramycin-resistance gene for vector selection in either E. coli or Streptomyces) were replicated in the dam dcm host E. coli ET12567 using standard procedures (Sambrook et al., 1989). Streptomyces plasmids were maintained by selection for resistance to thiostrepton (50 µg ml⁻¹; a gift from S. Lucania, Bristol-Myers Squib, Princeton, NJ, USA) or Apralan [50 µg ml⁻¹ and containing 50%, w/w, apramycin; Provel (Division of Eli Lilly Canada)]. E. coli plasmids pAUS5 (Giebelhaus et al., 1996) and pBluescript II SK/KS (Stratagene) were manipulated as described by Sambrook et al. (1989).

Isolation of a bldG103-complementing clone from an S. coelicolor phage library. A library (generously provided by R. Passantino, Instituto di Biologia dello Sviluppo del Consiglio
Nazionale dello Ricerche, Palermo, Italy) of *S. coelicolor* M145 DNA fragments in the *Streptomyces* phage vector KC304 was screened as previously described by Elliot et al. (1998). In brief, the library screening involved spotting 5 x 20 µl aliquots of the phage-library suspension onto a lawn of *S. coelicolor* C103 mycelial fragments on the surface of each of two R2YE (Hopwood et al., 1985) agar plates. The lawns were allowed to grow for 2 weeks and then the mycelium in and around the spotted areas was scraped from the surface of the plates, pooled and resuspended in sterile Milli-Q water. The mycelia were then homogenized to break the hyphae into small fragments, diluted and plated to give about 100 single colonies per plate on minimal medium (MM; Hopwood et al., 1985) containing glucose as the carbon source and thiostrepton (50 µg ml⁻¹) to select for phage-containing lysogens. Colonies showing aerial mycelium and pigmentation typical of the antibiotic-producing wild-type strain were sought as evidence of *bldG* complementation. Recombinant phages, containing the cloned *bldG* gene, were recovered from sporulating lysogens after CHCl₃ fuming (to prevent transfer of viable spores) and replicated onto Difco nutrient agar plates with SNA (Soft Nutrient Agar) overlays containing *S. lividans* 1326 spores (Hopwood et al., 1985) where free phages released from *S. lividans* lawn resulted in plaque formation in the *S. lividans* lawn.

**Subcloning and sequencing.** The ~5-5 kb *bldG*-complementing fragment from KC741 (see Results) was removed from the øC31 vector as a ~65 kb EcoRV fragment containing 1 kb of flanking øC31 vector DNA. The 65 kb blunt-ended fragment was ligated into EcoRV-digested pSET152 and the ligation mixture was used to transform *E. coli* DH5α. Digestion of the recombinant plasmid, designated pAU61, with HindIII allowed the subcloning of a ~2.5 kb fragment that extended from the unique HindIII site in the cloned *bldG*-containing DNA rightwards to the HindIII site located in the vector polylinker (see Fig. 1). The fragment was ligated into the HindIII site of pKCI218 and the recombinant plasmid, designated pAU63, was isolated after transformation of *E. coli* DH5α and selection for Apralan resistance. pAU63 was then passaged through *E. coli* ET12567 and used to transform protoplasts of *S. coelicolor* C103. Apralan-resistant transformants were selected and visually scored for their phenotypes. The 2.5 kb complementing fragment was subcloned in both orientations into the HindIII site of pBluescript SK + and recombinant plasmids were designated pAU64 and pAU65. Each of the pBluescript plasmids was used to generate a series of exonuclease III deletion derivatives by the method of Henikoff (1984) using the Erase-a-Base system (Promega). Additional DNA sequence for the region downstream of the right-hand end of the clone was obtained by first subcloning an overlapping 2.3 kb BamHI fragment extending rightwards from the predicted *bldG* ribosome-binding site (see Fig. 1). The BamHI fragment from the *S. coelicolor* cosmid H5 was shotgun cloned in both orientations into pBluescript KS + and recombinant plasmids were identified by hybridization using an ORF3-specific oligonucleotide, JWA9 (5'—CTCTGACGGAGCTGAGG-3'), as the probe. As described above, the pBluescript plasmids, designated pAU66 and pAU67, were then used to generate a series of exonuclease III deletion derivatives. The DNA sequence for the entire insert was determined both by manual and automated (Applied Biosystems model 373A) sequence analysis. The Universal sequencing primer and all other specific primers were obtained from the Department of Biological Sciences Synthesis Service, University of Alberta.

**Sequencing of *bldG* mutant alleles.** The mutant *bldG* gene was amplified from the chromosomal of the *bldG* mutant strains *S. coelicolor* C103, C101J, C107 and C536 using the oligonucleotide primers BKL65 (5’-CCGCTCAGGAGCA-TGC-3'; spanning nt −241 to −225) or JW A5 (5’-CACC- CAGTCTGCACAGC-3'; spanning nt −160 to −144), located in the 5' flanking region of *bldG*, and JW A6 (5’-GTTGACGGTGCCATG-3'; complementary to nt 451–467), located downstream of the *bldG* ORF (nucleotide positions are relative to the *bldG* translation start; see Fig. 3). PCR amplification was carried out with 1 µg *S. coelicolor* genomic DNA as the template and 40 pmol each primer in 100 µl reaction volumes. Expand Taq DNA polymerase (1.25 U; Roche) was used in each reaction. The reaction mixtures were denatured at 95 °C and then subjected to 30 cycles of 95 °C for 30 s, 52 °C for 30 s and 72 °C for 1 min. The major 708 bp (BKL65 and JW A6) or 628 bp (JWA5 and JW A6) amplification product was purified from a 5% polyacrylamide gel by crushing and soaking (Sambrook et al., 1989), and sequenced directly (DNA Sequencing Service, Department of Biological Sciences, University of Alberta) using as primers the oligonucleotides BKL65 (or JWA5) and JWA6 (described above); DBG4 (5’-GACCTGTCCCTGTCGAC-3'; nt 4–20); BKL81 (5’-GCGGAATTCGCTGAGTCGAGGTCGAC-3'; nt 102–118 and containing a 10 nt non-homologous extension [underlined]); BKL92 (5’-CGCCACCGCTCCTCAGA-3'; complementary to nt 314–330); and JWA1 (5’-TCAGTGCTTACACACCAGC-3'; complementary to nt 186–202) (nucleotide positions are relative to the *bldG* translation start; see Fig. 3) were also used.

**Creation of a *bldG* disruption mutant.** The *bldG* disruption construct pAU68 was made as follows. First, the 1.4 kb apramycin-resistance gene was cloned as an EcoRI–BamHI fragment from pBASpR (a pBluescript derivative containing the apramycin-resistance gene cassette; a gift from S. E. Jensen, Department of Biological Sciences, University of Alberta) into similarly digested pUC119 to give pUC119ApR. The 2.5 kb HindIII fragment containing *bldG* (see above) was purified from pAU64 and ligated into the HindIII site of pUC119ApR. Disruption of *bldG* was achieved by insertion of a blunted 1 kb *tsr* fragment (isolated from pAUS using SmaI and XbaI) into the blunted BglII site located in the *bldG*
coding region (see Fig. 3). The resulting recombinant plasmid was designated pAU68 and was passaged through *E. coli* ET12567 before being used to transform *S. coelicolor* J1501 and MI145 protoplasts. Thioesterase-resistant and Apralan-sensitive double-crossover mutants of each strain were selected and further tested by Southern blotting to confirm the double-crossover genotype. The *S. coelicolor* J1501 and MI145 bldG disruption strains were designated *S. coelicolor* C3b (*bldG3b*) and *S. coelicolor* C1a (*bldG1a*), respectively.

**Complementation of bldG mutants.** A 917 bp fragment containing the *bldG* ORF and upstream promoter was generated by PCR using pAU64 as template and the oligonucleotides DBG12 5'-GGCCGAATTCGTCGGATCGGGTCGG-3'; nt 451–467 and containing a non-homologous tail (underlined) with an EcoRI site and DBG3 5'-GGCCTCTAGAGTCGGATCGGGCATG-3'; complementary to nt 451–467 and containing a non-homologous tail (underlined) with an XbaI site as primers (nucleotide positions are relative to the *bldG* translation start). The amplified DNA was gel purified by the trough-purification method (Zhen & Swank, 1993), digested with EcoRI/XbaI and ligated into EcoRI/XbaI-digested PSET152. The resulting recombinant plasmid, designated pAU69, was isolated after transformation of *E. coli* DH5α and selected for Apralan resistance. After passage through *E. coli* ET12567, pAU69 was introduced into protoplasts of *S. coelicolor* bldG mutants CS36, C107, C101J and C3b by protoplast transformation. Apralan-resistant transformants were scored for their morphological and pigmentation phenotype.

**RNA isolation.** Streptomyces cultures were grown on cellophane discs (75 mm, 325P discs; Courtaulds Films) on the surface of R2YE agar as previously described (Leskiw *et al*., 1993). RNA was extracted essentially as described elsewhere (Hopwood *et al*., 1985) except that mycelia were scraped directly from the cellophane discs into modified Kirby mix. The RNA was isolated at various time points as described in Results.

**Northern blot analysis.** This was performed according to Williams & Mason (1985). RNA (40 μg) was denatured with glyoxal and DMSO, and size fractionated by electrophoresis at 4 V cm⁻¹ on a 1.25% agarose gel using a 10 mM Na₂HPO₄/NaH₂PO₄ (pH 7.0) recirculating buffer system. DNA molecular mass marker III (625 ng; Boehringer), treated probe was an [³²P]dCTP random-primer-labelled marker DNA to the filter at 44 °C. As a control for RNA loading levels, a probe for 16S rRNA was hybridized to the same blot. The 16S rRNA probe was an oligonucleotide, 5'-CCCGCTTTGCAACCGGT-3', corresponding to a conserved region in *Streptomyces* 16S rRNA sequences. Hybridization and washing were performed at 55 °C without formamide according to Procedure B described by Hopwood *et al.* (1985).

**S1 nuclease mapping of the bldG transcription-start site.** The probe for S1 nuclease mapping of *bldG* was generated by PCR amplification of a 264 bp fragment using pAU64 as the template. The primers were an 18-mer synthetic oligonucleotide, DBG14 5'-CGTCAATTTCGCCACCG-3', corresponding to a sequence internal to the *bldG* ORF and a 27-mer synthetic oligonucleotide, JWA20 5'-CGGCAAGCTTGGGATCGGGTCGG-3', corresponding to a region 194 nt upstream of the *bldG* start codon and containing a 10 nt non-homologous extension (underlined). The probes for S1 nuclease protection of the intergenic region between *bldG* and the downstream ORF were generated by PCR amplification of a 408 bp fragment using pAU64 as template. The primers were a 17-mer synthetic oligonucleotide, JWA17 5'-GGTGAGGCCGGTGATACG-3', corresponding to a sequence internal to the downstream ORF, and either a 27-mer synthetic oligonucleotide, JWA18 5'-GGCAGTGGGTCCGGCACTGAGG-3' or a 27-mer oligonucleotide, BK81 5'-GGGGAATTCGTCGGAGCCTG-3', corresponding to regions 121 nt and 350 nt upstream of the start codon for the downstream ORF and containing 10 nt non-homologous extensions (underlined). The amplified DNA was purified from a 2% agarose gel by the trough-purification method (Zhen & Swank, 1993). The 5' ends of the amplified DNA (about 2 pmol) were labelled with [³²P]ATP using T4 polynucleotide kinase. The probes, labelled at both ends, were used without treatment since the non-homologous extensions would be removed by the S1 nuclease treatment and would not result in the appearance of labelled, protected fragments (Leskiw *et al*., 1993). The sequencing ladders for the *bldG* and intergenic region S1 nuclease mapping were generated by the dyeoxy chain-termination method (Sanger *et al*., 1977) using DBG14 and JWA17 as primers, and pAU64 as the template. For each S1 nuclease protection reaction, 40 μg RNA was hybridized to 50000 Cerenkov c.p.m. probe in formamide buffer as described by Hopwood *et al.* (1985) except that glycyogen (Roche) replaced the carrier tRNA. To control for RNA loading levels, the RNA samples were first subjected to Northern blot analysis using the 16S rRNA probe (see above) and aliquots showing equivalent signals were subsequently used for S1 nuclease mapping. The samples were run under standard conditions on a 6% polyacrylamide sequencing gel.

**Primer-extension mapping of the bldG transcription-start site.** The primer was a 17-mer synthetic oligonucleotide, DBG15 5'-GTGAAAGAGCCAGCTGTTTTC-3', that was designed to hybridize immediately downstream from the *bldG* start codon and approximately 100 bp away from the proposed transcription-start point. Primer (50 pmol) was end-labelled with [³²P]ATP as described above. Approximately 5 pmol of [³²P]-labelled primer and 40 μg RNA were dissolved in 1 x SB buffer (60 mM NH₄Cl, 10 mM Tris–acetate, pH 7.4, 6 mM 2-mercaptoethanol) (Hartz *et al.*, 1988), denatured by heating to 90 °C for 5 min and annealed by transferring to 75 °C, slow cooling to 55 °C and then incubating at 55 °C for 1 h. The primer-annealed RNA was ethanol precipitated, washed with 80% ethanol and air-dried for 10 min. Extension of the primer was performed at 45 °C for 1 h in a solution of [³²P]dCTP random-primer-labelled marker DNA to the filter at 44 °C. As a control for RNA loading levels, a probe for 16S rRNA was hybridized to the same blot. The 16S rRNA probe was an oligonucleotide, 5'-CCCGCTTTGCAACCGGT-3', corresponding to a conserved region in *Streptomyces* 16S rRNA sequences. Hybridization and washing were performed at 55 °C without formamide according to Procedure B described by Hopwood *et al.* (1985).
1 x SB buffer, 15 mM magnesium acetate, 3 mM dNTPs, 17.5 U RNAGuard (Amersham) and 12.5 U AMV reverse transcriptase (Roche). Loading dye (98% deionized formamide, 10 mM EDTA, pH 8.0, 0.025% xylene cyanol, 0.025% bromophenol blue) was added and the reaction was evaporated at 80 °C for 20 min. The entire reaction mixture was loaded onto a 6% polyacrylamide sequencing gel. A sequencing ladder was generated as described above using the same oligonucleotide as for the primer-extension reactions.

**Computer-assisted sequence analysis.** General raw sequence handling was done using the GeneTool 1.0 program (BioTools). A version of the frame (Bibb et al., 1984) program modified to run on an Apple Macintosh (obtained from S. E. Jensen, University of Alberta, Canada) was used to detect putative ORFs. Similarities between deduced protein products and known proteins in the databases were detected using BLAST at the internet site http://ncbi.nlm.nih.gov. Multiple sequence alignments were generated using the PILEUP program of the Genetics Computer Group. Potential RNA secondary structures, together with AG values, were determined using the Mfold 2.3 program with the folding temperature set to 30 °C, the standard growth temperature used for these studies. The Mfold programs are available at the internet site http://www.ubic.wustl.edu/~zuker/rna/. Analysis of the ORF3 protein sequence for conserved bacterial histidine kinase domains was done using the Prosite ProfileScan server located at the internet site http://www.isrec.isb-sib.ch/software/PFSCAN_form.html.

**RESULTS**

**Cloning of the bldG gene**

Before the *S. coelicolor* genome project (see http://www.sanger.ac.uk/Projects/S_coelicolor/) began, we had cloned bldG by complementation of the *S. coelicolor* bldG103 mutation. bldG-containing recombinant KC304 phage derivatives were identified from a library of *S. coelicolor* M145 DNA fragments by their ability to restore antibiotic production and aerial hyphae formation to the *S. coelicolor* C103 mutant strain. The phage library was introduced into the bldG mutant, and two thiostrepton-resistant lysogens that were able to sporulate and produce the pigmented antibiotics actino-rhodin and undecylprodigiosin were selected for further study out of ~4000 colonies screened. The phages released from the lysogens were resorted on a *S. coelicolor* C103 lawn to verify that the cloned DNA was responsible for restoring antibiotic production and sporulation. It was not determined whether the production of calcium-dependent antibiotic, the other antibiotic produced by *S. coelicolor* J1501, was restored. Both lysogens appeared to have identical inserts and therefore only one, designated KC741, was chosen for further study.

**Sequencing and analysis of ORFs**

The ~5.5 kb DNA fragment containing bldG was subcloned from KC741 as a 6.5 kb EcoRV fragment containing 1 kb of flanking oC31 vector DNA. The fragment was introduced into pSET152, which integrates site specifically into the *S. coelicolor* chromosome at the oC31 attB site, generating pAU61. Further subcloning of a 2.5 kb HindIII fragment from pAU61 into pKC1218, a low-copy-number *E. coli–Streptomyces* shuttle vector, generated pAU63. Introduction of pAU63 into the *S. coelicolor* C103 mutant restored apparently wild-type levels of both pigmented-antibiotic production and aerial-mycelium formation. Nucleotide sequencing of the 2.5 kb HindIII fragment revealed two partial (ORF1 and ORF3) and one complete ORF (ORF2) (Fig. 1). BLAST analysis indicated that the predicted product of ORF2 resembles a group of anti-sigma factor proteins that include RsbV from *Staphylococcus aureus* (Wu et al., 1996) and SpoIIA and RsbV of *Bacillus subtilis* (Kalman et al., 1990; Dufour & Haldenwang, 1994; Duncan et al., 1996). Since ORF2 was the only complete ORF present on the subcloned 2.5 kb DNA fragment that was able to complement the bldG103 mutation, this ORF was designated bldG (GenBank accession no. AF134889). Since genes encoding anti-anti-sigma factors are typically encoded as the first gene in an operon also encoding a cognate anti-sigma factor and sigma factor (Kalman et al., 1990), the remainder of the ORF3 sequence, as well as an additional ~1 kb of downstream sequence was determined. Comparison with sequences in the databases revealed that the ORF3 protein product resembles anti-sigma factor proteins of *B. subtilis*, including SpoIIAB (Duncan & Losick, 1993; Min et al., 1993) and RsbW (Benson & Haldenwang, 1993a); anti-sigma F and anti-sigma B proteins, respectively. Surprisingly, a sigma factor was not encoded in the sequence downstream of the putative anti-sigma factor. Instead, a partial ORF showing similarity to pyrophosphate synthases was located 339 nt downstream of the ORF3 stop codon. The long intergenic region, together with the existence downstream of the ORF3 stop codon of a sequence capable of forming, in RNA, a stable stem–loop structure [AG = −21.3 kcal mol⁻¹ (−89 kJ mol⁻¹)] that might serve as a transcription-termination signal, suggested that the putative pyrophosphate synthase gene is not part of a bldG operon. The partial sequence of the divergently expressed ORF1 suggests that the protein product belongs to a group of bacterial RNA helicases. The sequences for bldG and the surrounding ORFs on cosmID H5 have been now been determined as part of the *S. coelicolor* genome sequencing project (http://www.sanger.ac.uk/Projects/S_coelicolor) and have been deposited under accession no. AL033636. It is clear from the sequence analysis that the 2.5 kb HindIII fragment present in our clone does not arise from contiguous chromosomal sequences and has come from the ligation of two non-contiguous *Sau3AI* fragments during library preparation. The non-contiguous DNA is shown as a hatched box in Fig. 1.

**Alignment of BldG and the ORF3 protein products with known proteins from the databases**

The bldG gene encodes a 113 aa protein with end-to-end similarity to a number of anti-anti-sigma factors that regulate the activity of sigma factors responsible for
stress-induced or growth-stage-specific transcription. BldG is most closely related to the *Sta. aureus* and *B. subtilis* RsbV proteins (38% and 40% identity, and 61% and 60% similarity, respectively), and to *B. subtilis* SpoIIAA (26% identity, 56% similarity). Alignment of BldG with RsbV and SpoIIAA proteins (Fig. 2a) reveals a highly conserved region of sequence surrounding the serine residue known to be phosphorylated on the *B. subtilis* SpoIIAA (Najafi et al., 1995). Likewise, alignment of the ORF3 product with other anti-sigma factors revealed about the same degree of similarity to RsbW and SpoIIAA proteins (26% identity, 56% similarity) and to *B. subtilis* SpoIIAA (Najafi et al., 1995). Likewise, alignment of the ORF3 product with other anti-sigma factors revealed about the same degree of similarity to RsbW and SpoIIAA proteins (26% identity, 56% similarity). However, the ORF3 protein product contains only two of the five conserved amino acid residues that are thought to be important for ATP and magnesium binding in *B. subtilis* SpoIIAA and RsbW. These findings suggest that the ORF3 protein product lacks the kinase activity and SpoIIAB in *B. subtilis* have been shown to have kinase activity that regulates their interaction with either their cognate sigma factor or anti-anti-sigma factor by phosphorylation of the anti-anti-sigma factor antagonist (Dufour & Haldenwang, 1994; Min et al., 1993). The ORF3 amino acid sequence was aligned with RsbW and SpoIIAB sequences to look for conserved kinase domains. As shown in Fig. 2b, the ORF3 protein product contains only two of the five conserved amino acid residues that are thought to be important for ATP and magnesium binding in RsbW (Kang et al., 1996) and SpoIIAB (Min et al., 1993), and in bacterial histidine kinases in general (Stock et al., 1995). Also, analysis of the ORF3 protein sequence using the Prosite ProfileScan Server did not reveal any conserved histidine kinase domains within the sequence. These findings suggest that the ORF3 protein product lacks the kinase activity
Fig. 3. Nucleotide sequence of the bldG promoter and coding region. The deduced amino acid sequence of bldG and the partial amino acid sequences of ORF1 and ORF3 are shown below in single letter code. Nucleotide positions are indicated relative to the first nucleotide of the bldG start codon (designated +1). The proposed transcription start sites (P1, P2; ○→) and putative −10 and −35 sequences for bldG are shown, as well as potential ribosome-binding sites (RBS) for bldG and ORF3. The two inverted repeats at the end of the bldG coding region are indicated by the arrows. Point mutations identified in the bldG mutant strains are shown in bold, with the base changes indicated above. Oligonucleotide primers used for sequencing of the bldG mutations are indicated by (®) above the sequence. The unique BglII site used to construct the bldG disruption vector is underlined.

A constructed bldG null mutant has the same phenotype as the strains with bldG point mutations

A bldG null mutant allele was constructed by inserting the thiostrepton-resistance gene (tsr) at a unique BglII restriction site internal to bldG (Fig. 3). The mutant allele was used to replace the wild-type allele in the chromosomes of S. coelicolor J1501 and M145 as described in Methods. Disruption of bldG in the chromosome of representative J1501 and M145 bldG mutants was confirmed by Southern blotting (data not shown) and the strains were designated S. coelicolor C3b (bldG3b) and S. coelicolor C1a (bldG1a), respectively. Both phenotypically resembled the bldG mutants of Champness (1988) in being devoid of blue or red pigments and failing to make detectable aerial mycelium on R2YE or MM+glucose (but sporulating aerial mycelium was formed, as with other bldG mutants, on MM+mannitol). The M145 bldG mutant was not studied further.

To confirm that the constructed J1501 bldG null mutant did not contain any other defects, pAU69, a pSET152-derivative containing a copy of bldG together with its promoter region, was introduced into bldG3b. The same complementation plasmid was also introduced into the bldG point mutants used in this study. pAU69 restored the wild-type phenotype to bldG3b as well as the C103, C107 and C536 strains; however, introduction of the plasmid had no effect on the phenotype of C101J.

Sequencing of bldG mutations

To map the location of the bldG mutation in C103, C107 and C536, as well as to confirm the absence of a mutation in the bldG coding region of C101J, the bldG ORF from the four bldG mutants was amplified by PCR and the resulting products were sequenced. As expected on the basis of the complementation studies, bldG101J did not have a mutation in the bldG coding or promoter region. Since the mutation in this strain maps to the same region of the chromosome as the mutations in C103, C107 and C536, it may reside in one of the nearby ORFs. Alternatively, the C101J strain may be a double mutant, making the genetic mapping data difficult to interpret and possibly misleading. As shown in Fig. 3, the bldG103 and bldG536 mutations both involve an A→T substitution at the same base, introducing a stop codon that would generate a truncated 87 aa BldG protein. The bldG107 mutation, the substitution of an adjacent CT for AG, leads to an amino acid change of aspartate to glutamate at position 56, and of a serine to alanine at position 57 of BldG. The latter residue corresponds to one known to be phosphorylated in SpoIIAA by the SpoIIB protein kinase. These data suggest that the C-terminal 26 aa of BldG, as well as the conserved region corresponding to the phosphorylation that has been shown in B. subtilis to be important for the regulation of anti-anti-sigma factor and anti-sigma factor interactions.
site on the *B. subtilis* SpoIIA, are important for BldG function.

**Although some bldG transcripts are monocistronic, some also include ORF3**

Northern blotting was performed using RNA isolated from *S. coelicolor* J1501 to determine the size of the *bldG* transcript as well as the timing of expression. The experiments were performed at least twice using RNA from three different time courses, and representative results are shown in Fig. 4. The probe, a 207 bp fragment internal to *bldG*, hybridized to two mRNA species with sizes of 600–700 nt and 1100–1200 nt. The size of the smaller transcript is consistent with a monocistronic *bldG* transcript terminating in the intergenic region between *bldG* and the start of ORF3, whereas the larger transcript has a size comparable to that expected for a polycistronic transcript including both *bldG* and the downstream ORF3 gene. When a second blot, prepared at the same time and in the same manner as the first, was probed with a PCR-amplified probe internal to the putative anti-sigma factor-encoding ORF3 (Fig. 4), only a band corresponding to the larger transcript was observed, and the same results were observed when a single blot was first hybridized with the *bldG*-specific probe, stripped and reprobed with the ORF3-specific probe (data not shown). The two transcripts were present at low levels during vegetative growth and at higher levels from 24 h post-inoculation, the time point corresponding to the appearance of aerial hyphae and pigmented antibiotics. Quantitative analysis of the two bands indicated that the smaller transcript is present at a two- to three-fold higher concentration than the larger transcript.

**The putative anti-sigma factor encoded downstream of bldG does not have its own promoter**

To further investigate whether the putative anti-sigma factor gene is co-transcribed with *bldG* rather than being transcribed from its own promoter, S1 nuclease protection studies were performed using a probe corresponding to the intergenic region between the two ORFs. RNA samples isolated from three independent time courses were tested and representative results are shown in Fig. 5. The major S1 product observed at all the time points tested has a size of 172 nt, which is equivalent to the size of the full-length probe excluding the 10 nt non-homologous extension. Full-length protection was also observed with a 401 bp probe generated using the same anti-sigma factor gene-specific internal primer (JWA17) and the oligonucleotide BKL81 (corresponding to nt 102–118 relative to the *bldG* start codon), which anneals to a region within the *bldG* ORF (data not shown). These data confirm that an anti-sigma factor gene-specific promoter is not located either in the intergenic region or within 3′ sequence of the BldG promoter.
transcription from either the possible that ORF3 could be expressed by readthrough 4).

probe, confirmed the existence of such transcripts (Fig. 5).
bldG3b promoter. Both the ability to complement the strain with a plasmid containing only the coding region. In the light of these data, the question was raised as to why disruption of bldG did not lead to polar effects on expression of the downstream gene. Analysis of the sequence of the disruption plasmid revealed that tsr was transcribed in the same direction as bldG and ORF3. Since the tsr terminator was not present on the cloned tsr-containing fragment, it is possible that ORF3 could be expressed by readthrough transcription from either the bldG promoter (located upstream of the inserted tsr gene) or from the tsr gene promoter. Both the ability to complement the bldG3b strain with a plasmid containing only the bldG ORF, as well as a Northern blot analysis using an ORF3-specific probe, confirmed the existence of such transcripts (Fig. 4).

S1 nuclease mapping and primer-extension analysis reveal two bldG mRNA 5’ ends
To determine the transcription-initiation site for the bldG transcripts, high-resolution S1 nuclease mapping studies were performed using a 264 bp PCR-amplified probe labelled at the 5’ ends. As above, RNA samples isolated from three independent time courses were tested and representative results are shown in Fig. 6.

Interestingly, two major S1 nuclease-protected products of 152 nt and 193 nt were detected when the hybridization was carried out at 5 °C and at 12 °C (not shown) above the predicted probe melting temperature. If these products correspond to transcription-start points, the +1 positions would be situated 82 and 123 nt upstream of the bldG translation-start codon. The same mRNA 5’ ends were identified by primer-extension analysis of the same RNA samples using an end-labelled 17 nt oligonucleotide primer (Fig. 6), showing that the shorter protected species (see Fig. 3) was not an artefact caused by an S1 nuclease hypersensitive site. The transcription-start site proximal to the translation-start codon is located just downstream of a putative −10 sequence similar to those found in E. coli-like Streptomyces promoters (Strohl, 1992). At an atypically short spacing of 16 bp, a −35-like sequence was also present in this region. Similarly, a putative −10 sequence was found upstream of the distal transcription-start point. In this case, it was not possible to identify appropriately spaced −35 sequences. As was seen in the Northern analysis, both transcripts are expressed at low levels during vegetative growth and are upregulated when aerial mycelium and pigmented antibiotics become visible. The relative intensities of the signals for transcripts initiating at both promoters correlate well with the relative intensities of the long and short transcripts seen by Northern analysis. These results may suggest that initiation from the distal promoter gives rise to the polycistronic transcript and that transcription initiation at the proximal promoter gives rise to the monocistronic transcript. However, the low resolution of the Northern blots did not permit us to address this directly.

The bldG gene product does not appear to be involved in an autoregulatory circuit
The well-studied bldG homologue spoIIA plays an indirect part in its own transcriptional regulation. Like bldG, spoIIA is transcribed from two promoters. One of these is dependent on the σ7 factor encoded by spoIAC, the third gene in the spoIIA operon. Inactivation of spoIIA leaves σ7 locked in a complex with SpoIAB, preventing use of the σ7-dependent promoter (Schuch & Piggot, 1994). To find out whether bldG is likewise important for transcription of one or both of its promoters, S1 nuclease protection studies using RNA samples isolated from the constructed bldG null mutant, bldG3b, were performed (Fig. 6). The same two mRNA 5’ ends that were identified in bldG+ strain were detected at comparable relative abundance in the mutant, indicating that the loss of bldG activity has no effect on bldG transcript initiation.

DISCUSSION
The simplest interpretation of the extracellular-signalling model for regulation of differentiation in S. coelicolor is that all of the bld genes might directly
encode the intercellular signals, or the components for sensing or uptake of those signals. On the basis of what we know about the bld gene products characterized to date, this simple interpretation is not supported. So far, only bldK and bldJ appear to encode direct components of the hierarchical signalling cascade (Nodwell et al., 1996) and all of the other bld genes appear to encode products that must have indirect roles in signal synthesis, sensing or uptake. bldA encodes a leucyl tRNA that is the only means for efficient translation of rare UUA sensing or uptake. ORF3 displays similarity or uptake.

products that must have indirect roles in signal synthesis, developmentally regulated transcription by regulating the activity of a specific sigma factor(s).

Despite the sequence similarities between BldG and the deduced protein product of its co-transcribed, downstream ORF3 with proteins known to regulate the activity of certain sigma factors in Bacillus, we see differences between the way that the genes are organized and regulated in the two organisms. In contrast to both the spoIIA and rsb operons from Bacillus, a sigma factor is not encoded at the bldG locus. So far, there are no reports in the literature of genes for an anti-sigma factor antagonist and its anti-sigma factor lying at a distant location from the gene for the cognate sigma factor. If bldG and ORF3 do in fact encode an anti-anti-sigma/anti-sigma factor pair, this raises the possibility that the regulatory pair might regulate the activity of more than one sigma factor, or that they may regulate a particular sigma factor only in a subset of the conditions in which that sigma factor directs gene expression. The antibiotic- and acylmecylceum deficient phenotype of bldG mutants then presents us with two possibilities: the regulatory pair could control one or more globally acting sigma factors that serve to activate transcription of both antibiotic and sporulation-specific genes; or alternatively, the pair could regulate two sigma factors, one of which activates transcription of antibiotic biosynthetic genes and one which activates sporulation-specific gene expression. Interestingly, of two sporulation-specific sigma factors so far identified in Streptomyces (Kelemen et al., 1996; Méndez & Chater, 1987), one (σ^F) is of the subfamily regulated by homologues of BldG/ORF3 in Bacillus. Sigma factors recognizing antibiotic-biosynthetic-gene promoters have not been definitively identified. It is attractive to suggest that immobilization of the putative anti-sigma factor on an affinity column might lead us to one such sigma factor. A search of the S. coelicolor database reveals that there is no shortage of potential sigma factor targets. The Bacillus sigma factors that are regulated by anti-anti-sigma/anti-sigma factor pairs make up a subfamily of σ^F-like sigma factors (Lonetto et al., 1992), and so far at least 9 sigma factors that would fall into this subfamily are found on the S. coelicolor chromosome; several of them (in addition to sigF) are not located next to bldG/ORF3-like gene pairs (Gabriella Kelemen, personal communication).

The C-terminal region of sigma factors in this subfamily shares homology with the DNA-binding region of a group of bacterial transcriptional-activator proteins (Kahn & Ditta, 1991; Lonetto et al., 1992). Our analysis of the σ^F sequence using the Prowtie ProfileScan server revealed similarity between region 4.2 near the σ^F C terminus and the helix-turn-helix domain of LuxR-related DNA-binding proteins. This raises the alternative possibility that the BldG and ORF3 regulatory pair could target a DNA-binding protein rather than a sigma factor. However, this possibility seems unlikely since the candidate σ^F contact residues for SpoIIB are not located in region 4.2 of σ^F. SpoIIB contacts σ^F in three areas corresponding to conserved regions 2.1, 3.1 and 4.1 of σ^F-like sigma factors (Decatur & Losick, 1996). So far, only two putative contact residues on SpoIIB have been identified (Garsin et al., 1998); and one of these two residues, R20, is conserved in the ORF3 gene product.

In addition to the absence of an operon-encoded sigma factor, the bldG operon differs from the Bacillus operons in the way that the expression of the genes is controlled. In the case of both the spoIIA and rsb operons, the genes for the anti-anti-sigma factor, the anti-sigma factor and the sigma factor are expressed as a single transcript, from either a σ^H-dependent promoter in the case of spoIIA (Wu et al., 1991) or a σ^B-dependent promoter in the case of rsb (Wise & Price, 1995). Expression of the genes is also upregulated from a second promoter that is recognized by the operon-encoded sigma factor (Schuch & Piggot, 1994; Wise & Price, 1995). For the spoIIA operon, upregulation is dependent on the activity of a prespore-compartment-specific phosphatase, SpoILE, that activates the SpoIIB anti-anti-sigma factor by dephosphorylation, allowing the formation of SpoIIB/SpoIAB complexes and the liberation of σ^F (Duncan et al., 1995; Wu et al., 1998). Upregulation of σ^B occurs either in response to energy stress, where reduced ATP levels in the cell influence the phosphorylation state of the anti-anti-sigma factor, RsBV, or in response to environmental stress signals that activate a phosphatase that dephosphorylates RsBV and results in the liberation of σ^B (Voelker et al., 1995).

For the bldG operon we also see a complex transcription pattern involving not only two different promoters, but also the generation of two differently sized bldG-containing transcripts. Based on the Northern analyses, the longer and less abundant of the two transcripts extends through bldG to include the downstream putative anti-sigma ORF3. The shorter, more strongly expressed transcript appears to terminate in the inter-
genic region, and includes only the bldG coding sequence, a situation not seen with either the spoIIA or rsb operons. Analysis of the intergenic sequence for potential RNA secondary structures revealed two inverted repeats that could give rise to stem–loop structures with ΔG values of −22.2 and −31.3 kcal mol⁻¹ (−93 and −131 kJ mol⁻¹) (see Fig. 3), respectively, and either of which might play a role in mRNA stabilization or as a transcription-termination signal. Although a string of U residues is not present downstream of either inverted repeat, it is well documented that inverted repeats without poly(U) tails can act as terminators in Streptomyces (Deng et al., 1987; Ingham et al., 1993). The existence of the shorter transcript that hybridizes only to the bldG probe might suggest that a terminator is functional; however, an alternative mechanism for the generation of the shorter transcript could be endo- or exo-nucleolytic removal of the 3′ ORF3-containing end of the longer transcript. Similarly, although a comparison of transcript abundance as seen by Northern analysis, primer-extension and S1 nuclease protection studies may suggest that the longer transcript arises from initiation at the more upstream promoter, and that transcripts initiating from the proximal bldG promoter terminate in the intergenic region, our analyses do not allow a definitive conclusion.

For both the Bacillus spoIIA and rsb operons, the genes appear to be translationally coupled such that they are expressed in equimolar concentrations (Magnin et al., 1997; Benson & Haldenwang, 1993b; Kalman et al., 1990). The lack of a promoter in the intergenic region between bldG and the downstream ORF3 encoding the putative anti-sigma factor suggests that these two genes might also be translationally coupled. However, the very long intergenic region means that there cannot be a straightforward mechanism of coupling that involves either closely spaced or overlapping start and stop codons to ensure a 1:1 stoichiometry. This, together with the added complication of a bldG monocistronic transcript that might lead to an excess of BldG, highlights the need to address the levels of the BldG and ORF3 proteins throughout the life cycle. By analysing the protein levels we should be able to answer questions about how the equilibrium between BldG and the ORF3 product is shifted at different stages of growth. Certainly, appearance of a transcript does not mean that the RNA is being translated to protein, a fact that warrants further investigation because of the expression of a putative RNA helicase from a divergently expressed promoter that overlaps the −10 region of the most upstream bldG promoter (J. Stoehr & B. K. Leskiw, unpublished).

The availability of the purified proteins and antibodies to those proteins will also help us to explore the role, if any, of the conserved phosphorylation site on BldG. Although mutation at this site does abolish BldG activity in the same way that Ser58 mutation in SpoIIAA does, the existence of a second mutation in this region in bldG107, together with the lack in the putative anti-sigma factor of conserved residues found in bacterial histidine protein kinases, makes it difficult to draw any conclusions.

ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada and by the Alberta Heritage Foundation for Medical Research. We are grateful to Rosa Passantino for the use of her M145 library and to Wendy Champness for supplying all of the bldG mutant strains. We thank William Klimke for help with the computer alignments of protein sequences and Karen Willis for her help with the digital figures. We also thank Kim Bilida and Michelle Tougas for technical assistance, and Ken Roy for critical reading of the manuscript.

REFERENCES


Duncan, L. & Losick, R. (1993). SpolIB is an anti-sigma factor that binds to and inhibits transcription by regulatory protein
sigma F from Bacillus subtilis. Proc Natl Acad Sci USA 90, 2325–2329.


S. coelicolor bldG locus

and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome. Mol Microbiol 21, 77–96.


