Transcription analysis of the stretomyces-coelicolor A3(2) ribosomal-RNA operon

Wezel, G.P. van; Krab, I.M.; Stephen, D.; Bibb, M.J.; Vijgenboom, E.; Bosch, L.

Transcription analysis of the *Streptomyces coelicolor* A3(2) *rrnA* operon

Gilles P. van Wezel,† Ivo M. Krab,‡ Stephen Douthwaite,§
Mervyn J. Bibb,† Erik Vijgenboom and Leendert Bosch

Author for correspondence: Gilles P. van Wezel. Tel: +44 603 52571. Fax: +44 603 56844.
e-mail: vanwezel@bbsrc.ac.uk

Leiden University,
Department of Biochemistry,
PO Box 9502, 2300RA
Leiden, The Netherlands

Transcription start sites and processing sites of the *Streptomyces coelicolor* A3(2) *rrnA* operon have been investigated by a combination of *in vivo* and *in vitro* transcription analyses. The data from these approaches are consistent with the existence of four *in vivo* transcription sites, corresponding to the promoters P1–P4. The transcription start sites are located at −597, −416, −334 and −254 relative to the start of the 16S rRNA gene. Two putative processing sites were identified, one of which is similar to a sequence reported earlier in *S. coelicolor* and other eubacteria. The P1 promoter is likely to be recognized by the RNA polymerase holoenzyme containing σ^{hrdB}, the principal sigma factor in *S. coelicolor*. P2 also shares homology with the consensus for vegetative promoters, but has a sequence overlapping the consensus −35 region that is also present in the −35 regions of P3 and P4. The −35 sequence common to P2, P3 and P4 is not similar to any other known consensus promoter sequence. In fast-growing mycelium, P2 appears to be the most frequently used promoter. Transcription from all of the *rrnA* promoters decreased during the transition from exponential to stationary phase, although transcription from P1 and P2 ceased several hours before that from P3 and P4.

INTRODUCTION

Streptomycetes, like all members of the order Actinomycetales, have a complex life cycle that involves three stages of differentiation. These morphological changes are accompanied by a wide range of physiological events including the production of secondary metabolites, many of which have activity as antibiotics (Chater, 1989). The molecular processes regulating these events are presently only superficially understood.

Regulation of transcription of stable RNA operons is a pivotal process in all organisms. The only *Streptomyces*

rRNA operon for which the transcription has been studied (*Streptomyces coelicolor rrnD*; Baylis & Bibb, 1988) contains four promoters. Little is known about the regulation of rDNA transcription in *Streptomyces* and how this relates to differentiation processes. The most studied regulatory phenomenon is stringent control, elicited by accumulation of highly phosphorylated guanosine molecules as a response to starvation conditions. Although ppGpp has been implicated in secondary metabolism (Ochi, 1986, 1987), recent reports have failed to confirm this (Bascaran *et al.*, 1991; Strauch *et al.*, 1991). An intriguing observation was described by Granozzi *et al.* (1990), who reported a second round of stable RNA production on solid medium that coincided with the development of aerial hyphae.

In contrast to *Streptomyces*, transcription of *Escherichia coli* rRNA operons occurs from two promoters, P1 and P2. For *rrnB*, P1 is tightly regulated whereas P2 is used much less frequently than P1 during exponential growth (Gourse *et al.*, 1986). Recently, expression of all seven *E. coli rrn* operons was compared (Condon *et al.*, 1992). Although during fast growth hardly any difference could

be detected in transcription of these operons, slight but significant differences were observed in response to some stress conditions, including nutritional shift-down.

Recently, we described the cloning of three of the six rRNA operons of *S. coelicolor* (*rrnA*, *rrnC* and *rrnE*) and sequence analysis of the *rrnA* 16S rRNA gene and upstream region (van Wezel *et al.*, 1991). The nucleotide sequences of the upstream regions of the *S. coelicolor* ribosomal RNA operons *rrnA* and *rrnD* differ significantly, and this prompted us to investigate the locations and relative strengths of the *rrnA* promoters and compare the results with those obtained for the *rrnD* promoters. Here we analyse the relative contribution of each of the *rrnA* promoters to the total transcription of this operon under various growth conditions.

METHODS

Bacterial strains, plasmids and bacteriophages. *S. coelicolor* M145 was cultivated in Tryptone Soya Broth plus 10% (w/v) sucrose (TSBS), liquid minimal medium (Hopwood *et al.*, 1985), liquid minimal medium supplemented with Casamino acids [designated SMM (Takano *et al.*, 1993)], or on R2YE or minimal medium agar plates (Hopwood *et al.*, 1985). Liquid cultures were inoculated with spores at a density of 5×10^6 c.f.u. ml^{-1} and grown at 30 °C with vigorous shaking (300 r.p.m.), resulting in reproducibly dispersed growth. In SMM, nitrogen limitation results in a rapid transition into stationary phase approximately 18 h after inoculation of the cultures (Strauch *et al.*, 1991; Takano *et al.*, 1993). RNA used for nuclease S1 mapping was isolated from SMM inoculated with spores germinated for 6 h. The RNA we used for nuclease S1 mapping experiments was isolated by Eriko Takano (John Innes Centre, UK) as described earlier (Takano *et al.*, 1993).

RNA used for primer-extension experiments was isolated from liquid minimal medium cultures inoculated with ungerminated spores. In this case, stationary phase was reached after approximately 33 h. pIJ4083 (Clayton & Bibb, 1990a) was used in the promoter-probing experiments. *E. coli* JM109 (Yanisch-Perron *et al.*, 1985) was the host for pUC18 and pUC19, and for bacteriophages M13 mp18 and mp19 (Yanisch-Perron *et al.*, 1985), and was cultivated in LB medium (8 g Difco Bacto-tryptone, 5 g NaCl and 5 g Difco yeast extract per litre).

***In vitro* DNA manipulations.** Transformation of *S. coelicolor*, and isolation and cloning of *Streptomyces* DNA, were carried out according to Hopwood *et al.* (1985). Transformation of *E. coli*, and DNA cloning, isolation and gel electrophoresis were performed by standard procedures (Sambrook *et al.*, 1989).

DNA sequencing. This was performed using the T7 DNA polymerase sequencing kit (Pharmacia) or Sequenase version 2.0 (US Biochemical). The primers used for sequencing were the same as those used for primer extension. The location of the primers is shown in Fig. 1(a).

RNA isolation. RNA was isolated from *S. coelicolor* M145 according to Hopwood *et al.* (1985). To remove residual DNA, the RNA was salt-precipitated twice in 0·3 M sodium acetate (pH 6·0). No DNase treatment was applied (DNA was not detectable on agarose gels).

Promoter-probe experiments. *xylE* from *Pseudomonas putida* was used as a reporter gene (Zukowski *et al.*, 1983) to assay for *in vivo* promoter activity. Fragments containing different segments of the *rrnA* upstream region were cloned into the polycloning site upstream of the promoterless *xylE* gene of the

vector pIJ4083 and introduced into the appropriate *S. coelicolor* hosts by transformation of protoplasts.

S. coelicolor transformants containing the correct construct were grown in TSBS plus thiostrepton (5 µg ml^{-1}) until actinorhodin production was observed, typically after 40 h of growth. S30 extracts were prepared from these cultures as follows. Mycelium was spun down, washed twice and resuspended in 100 mM phosphate buffer (pH 7·2) followed by sixty 5 s pulses of ultrasound with intermittent pauses of 5 s to allow cooling of the samples. This was followed by a 15 min centrifugation at 30000 *g* to remove cell debris.

Expression of *xylE* was measured as the rate of catechol conversion by the *xylE* gene product, catechol 2,3-dioxygenase (EC 1.13.11.2). Catechol conversion rates were measured by the increase in absorbance at 375 nm (A_{375}) after addition of 0·2 mM catechol to diluted S30 extracts. From this, the specific catechol dioxygenase activity (SCDA) was determined as the increase in A_{375} per min corrected for plasmid copy number and protein concentration. The correction for plasmid copy number was performed as follows. DNA was isolated, a dilution series was prepared, and spotted onto Hybond-N filters (Amersham). These filters were hybridized with a probe recognizing the vector DNA [1 kb *Bcl*I fragment containing the *tsr* gene; hybridization conditions as described by van Wezel *et al.* (1991)], washed and counted in a Betascope radioactivity counter (Pharmacia). From these data, the relative plasmid copy numbers of the various cultures were determined and used to correct the data obtained from the catechol dioxygenase assay. Only slight differences in copy number were observed: the values fluctuated within a margin of approximately 25%. Protein concentrations were estimated by measuring the A_{280}. Using these values, all samples were diluted to yield approximately equal protein concentrations.

Primer extensions with reverse transcriptase. End-labelling of oligonucleotides and primer extensions with reverse transcriptase was performed as described earlier (Stern *et al.*, 1988). The experiments were carried out at least twice to establish reproducibility. The primers used in these experiments are deoxyoligonucleotides gvw-1 (-1 to $+18$), gvw-3 (-430 to -413) and gvw-5 (-214 to -198) (Fig. 1a). One microgram of RNA and 1 ng of ^{32}P end-labelled primer were used in each experiment.

Southern hybridization with oligonucleotides. DNA samples were electrophoresed through 0·7% agarose gels in TAE and transferred to Hybond-N, using 20 × SSC as the blotting buffer. Hybridization was carried out in 6 × SSC/0·1% SDS/0·1% pyrophosphate/5 × Denhardt's solution with 50 µg salmon sperm DNA ml^{-1}. Hybridization temperature was derived from the equation $T_H = T_m - 5$ °C, and was 47 °C for deoxynucleotide gvw-1, and 49 °C for deoxynucleotides gvw-3 and gvw-5. Filters were washed in 6 × SSC plus 0·1% SDS at a temperature 3 °C below T_m, with several changes of wash buffer until background was sufficiently reduced.

Nuclease S1 mapping. Hybridization of 5 µg RNA with the appropriate DNA probe was performed according to Murray (1986) in NaTCA buffer (Summerton *et al.*, 1983). All subsquent steps were carried out as descried by Gramajo *et al.* (1993), under conditions of probe excess.

***In vitro* transcription analysis.** Partially purified RNA polymerase holoenzyme preparations were isolated as described by Buttner & Brown (1985). *In vitro* run-off transcription experiments were performed as described earlier (Buttner & Brown, 1985). Products were analysed on denaturing 6% (w/v) polyacrylamide gels using ^{32}P end-labelled *Hpa*II fragments of pBR322 as size markers.

(a)

```
-751  GACGACCGCA CCGCGCTCGG TGATGTGCTG ACCCGCCTCC TGCGCAGGGT
                                                      FspI
      D  D  R  T   A  L  G   N  V  L   T  R  L  L   R  A  V

-701  GGAGCACGCA CAGCCGGACC GCGGTGGCCG GCTCAGCGAC CTCGATTAAA
                                      NaeI
      E  H  A   E  P  D  R   G  G  R   L  S  D   L  D  *

-651  GATCTTGACA GCGACGCTTG ACAGCCCTCT GCGGGATCCG TAAGGTTCTT
      BglII
         .P1
-601  CGGGTTGCCA CGGGACCGTA ACGGTTCTTC GGCAGCACCT CCCGCCGCAG

-551  CAGCGGCAAT CAAACCACCA GCACGACCCC CAAGTGGGAA TGACTTCGGC

-501  GCGCCCGAAT TCAATTCCGC TGGGTACTCG TCGGCCCGAT CGACCCCGAT
           EcoRI
                                              .P2
-451  TGGGAATCGC CGAGAGGGTC CGCTAAGGTT TGAGACGTCG GAACGGCCCA
                           gvw-3      AhaII

-401  ACAGCCGGGA CGACAACCCC TCCTGACGGG GAATCAGGGC CCGAAAGGAT
                                              ApaI
          .P3
-351  CTGATAGAGT CGGAACCGCC GGAAAGGGAA ACGCGGAACG GGGAACCTGG
                                      BstUI
                                                      .P4
-301  AAAGCACCGA GGAAATCGGA TCGGAAAGAT CTGATAGAGT CGGAAACGCA
                GGAAATCGGA TCGGAAAGAT           BglII

-251  AGACCGAAGG GAAGCGCCCG GAGGAAAGCC CACGAGAGTC TCCTGGGTGA
      AGACCGAAGG GAAGCGCCCG                        gvw-5

-201  GTACAAAGGA AGCGTCCGTT CCTTGAGAAC TCAACAGCGT GCCAAAAGTC
      RsaI                                         HincII

-151  AACGCCAGAT ATGTTGATAC CCCGACCTGA TCGGATCTCC GTTCGGGTTG

-101  AGGTTCCTTT GAAGTAACAC AACAGCGAGG ACGCTGTGA ACGGTCGGATT

 -51  ATTCCTCCGA CTGTTCCGCT CTCGTGGTGT CACCCGATTA CGGGTATACA
                                                      AccI

  -1  TTCACGGAGA GTTTGATCCT GGCTCAGGAC GAACGCTGGC GGCGTGCTTA
      gvw-1

 +51  ACACATGCAA .... 16S rRNA gene
```
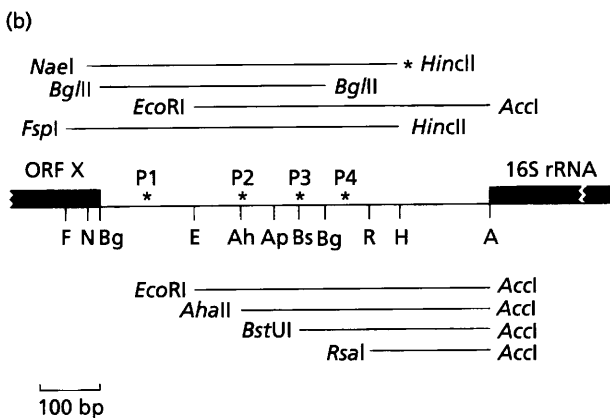
(b)

Fig. 1. Nucleotide sequence of the region upstream of the 16S rRNA of *rrnA* and location of restriction fragments and oligonucleotides. (a) Nucleotide sequence of the *rrnA* upstream region. The sequence of the 16S rRNA gene is shown in italics. The amino acid sequence of the translation product of the upstream open reading frame is shown. Oligonucleotides are indicated in bold face. Underlined sequences show two direct repeats, one of 22 nt and one of 57 nt. Locations of transcriptional start sites were determined by primer extension (Fig. 3), and are represented by dots. (b) Restriction fragments used for promoter-probing experiments (below), *in vitro* transcription (top, without asterisk) and nuclease S1 mapping

RESULTS

Promoter-probing of the *rrnA* upstream region

To determine the approximate location of the *rrnA* promoters and to obtain a rough estimate of their relative strengths, we transformed *S. coelicolor* M145 with derivatives of the multicopy vector pIJ4083 in which a variety of *rrnA* upstream fragments has been cloned in front of the promoterless *xylE* gene. The origin of these fragments is shown in Fig. 1(b). Transformants were grown in TSBS and SCDA levels determined (Table 1). Transcriptional activity increased with the length of the fragments. As can be deduced from the table, promoters are likely to be present between −318 and −199, between −416 and −318 and between −494 and −416.

Repeated attempts to subclone fragments containing the *FspI–EcoRI* (−708 to −494) part of the *rrnA* upstream region into pIJ4083 were unsuccessful. The occurrence of multiple copies of this fragment may be lethal to *S. coelicolor* M145. Perhaps it contains a binding site for a factor essential for vegetative growth. Interestingly, a protein with a binding site in the −708 to −494 region, which is centred around the *BglII* site at −651 (see Fig. 1a), has been detected and its identity is presently being investigated (van Wezel, 1994).

Specificity of oligonucleotides

Three deoxyoligonucleotides, gvw-1 (−1 to +18), gvw-5 (−214 to −198) and gvw-3 (−430 to −413), were used for primer extension experiments, the results of which are given in the next section. Although optimum conditions for Southern hybridization and DNA–RNA hybridization in primer-extension experiments differ, *Bam*HI- and *Sal*I-digested chromosomal DNA were probed in triplicate with ^{32}P end-labelled oligonucleotides (Fig. 2), to assess the ability of these primers to recognize transcripts originating from the other *rrn* operons. As expected, gvw-1 recognizes the 5′ end of the 16S rRNA gene in all six rRNA operons, and therefore signals resulting from primer extension using this oligonucleotide will be a mixture derived from all these six rRNA operons. However, since all rRNA transcripts are processed at the same positions (Baylis & Bibb, 1988), the bands arising from primer extension with gvw-1 should reveal the processing sites of *rrnA* in addition to those of the other *rrn* operons. Gvw-3 is operon-specific as it hybridized only with a *Bam*HI and a *Sal*I fragment corresponding to *rrnA* (Fig. 2). Gvw-5 hybridized with bands corresponding to *rrnA*, but also to an 18 kb *Bam*HI fragment and a 4·0 kb *Sal*I fragment. The latter signal does not seem to correspond to an rRNA operon, since all rRNA operons are located on *Sal*I fragments larger than 5·8 kb (van Wezel *et al.*, 1991). The origin of this band is unclear.

(top, asterisk shows 5′ ^{32}P end-label). Abbreviations of restriction enzymes: A, *Acc*I; Ah, *Aha*II; Ap, *Apa*I; Bg, *Bgl*II; Bs, *Bst*UI; E, *Eco*RI; F, *Fsp*I; H, *Hinc*II; N, *Nae*I; R, *Rsa*I. Not all restriction sites are shown. Putative transcription start sites are represented by an asterisk.

Table 1. Promoter-probing of the *rrnA* upstream region

Specific catechol dioxygenase activity (SCDA) values are the mean of three independent experiments. The *Bst*UI–*Acc*I construct resulted in a catechol conversion rate measured as a mean increase of 0·055 A_{375} units min^{-1}. After correction for plasmid copy number and protein concentration, the SCDA (given in arbitrary units) of the *Bst*UI–*Acc*I construct was defined as 1·0. ND, No catechol dioxygenase activity detected.

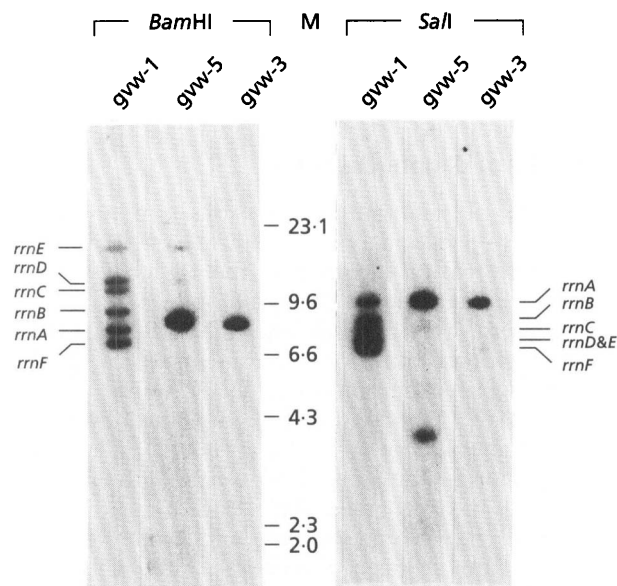| Fragment	Segment of *rrnA* upstream region	SCDA
No insert	–	ND
*Rsa*I–*Acc*I	−199 to −6	ND
*Bst*UI–*Acc*I	−318 to −6	1·0
*Aha*II–*Acc*I	−416 to −6	4·0
*Eco*RI–*Acc*I	−494 to −6	8·0

Fig. 2. Southern hybridization of *S. coelicolor* A3(2) chromosomal DNA with the oligonucleotides used for primer extensions. Digestions were carried out with *Bam*HI (left panel) and *Sal*I (right panel). Southern hybridizations were performed using the oligonucleotides indicated above the lanes. Operon classification is shown on the left and right side of the figure. Lane M shows a DNA size marker, *Hin*dIII-digested λ DNA, in kb.

Location of potential *rrnA* start and processing sites by primer-extension

To locate the potential transcription start and processing sites precisely, primer-extension experiments were performed on RNA isolated from cultures grown in liquid minimal medium (without addition of Casamino acids). *S. coelicolor* M145 mycelium was harvested at different times (12, 22, 29 and 42 h after inoculation with ungerminated spores) and RNA was isolated. Under these growth conditions, stationary phase was reached after approxi-

mately 33 h. The oligonucleotides gvw-1, gvw-3 and gvw-5 were used as primers to locate RNA 5′ ends resulting from either transcription initiation or processing in the region between the 16S rRNA gene and the upstream open reading frame which ends at nt −653 (van Wezel *et al.*, 1991). From the work published by Baylis & Bibb (1988), it follows that the region downstream of −200 is essential for correct processing of rRNA transcripts of all six rRNA operons. Secondly, as shown below, no *in vitro* transcription start sites have been observed for *rrnA* in this region, a result confirmed by the promoter-probe experiments. Therefore, we can assume that RNA end-points corresponding to positions upstream of −200 are very likely to have arisen from *in vivo* transcription initiation, and those corresponding to positions between −200 and −1 from processing of the various rRNA precursors.

Processing sites were identified by primer-extension using gvw-1 (−1 to +18, see Fig. 1a). Three RNA end-points were identified, corresponding to nt positions −168, −142 and −89, and referred to as bands I, II and III, respectively (Fig. 3a). The relative intensities of the three bands, as estimated by scanning the autoradiograms, were invariable in all four RNA preparations.

Upon extension of oligonucleotide gvw-5 (−214 to −198) with reverse transcriptase, six major RNA end-points were found (bands C–H in Fig. 3b), located at nt positions −416, −334, −308, −254, −248 and −245, respectively. Bands C, D and F correspond closely to sites of transcription initiation *in vitro* (P2, P3 and P4, respectively; see below). Since gvw-5 also hybridizes to a band that does not correspond to *rrnA* (Fig. 2), it is possible that at least one of the bands E, G and H observed in Fig. 3(b) reflects non-*rrnA*-derived transcripts.

To analyse whether transcription initiated upstream of nt position −400, oligonucleotide gvw-3 (−430 to −413) was used for primer extension. Two bands were observed, corresponding to RNA end-points at positions −597 (band B) and −640 (band A) (Fig. 3c). The size of band B is consistent with transcription initiation at P1 (see below). Band A is very strong in lane 1 (12 h), but less intense in lanes 2 (22 h) and 3 (29 h), and strong again in lane 4 (42 h), though weaker than in lane 1. This unexpected (but reproducible) growth-phase-dependence raises the question as to the origin of band A. If it corresponds to a real transcription start site, then the corresponding promoter appears to be utilized with almost equal efficiency in early exponential-phase and in stationary-phase cultures.

Location of transcription end-points by nuclease S1 mapping

RNA was isolated at different time-points, between 9 (OD$_{450}$ 0·3, corresponding to early exponential phase) and 33 h (OD$_{450}$ 2·1, stationary phase) after inoculation of SMM with pregerminated spores. The *Nae*I–*Hin*cII (−672 to −151) fragment, ^{32}P end-labelled at the *Hin*cII site (Fig. 1b), was used as a probe. As can be seen in Fig.
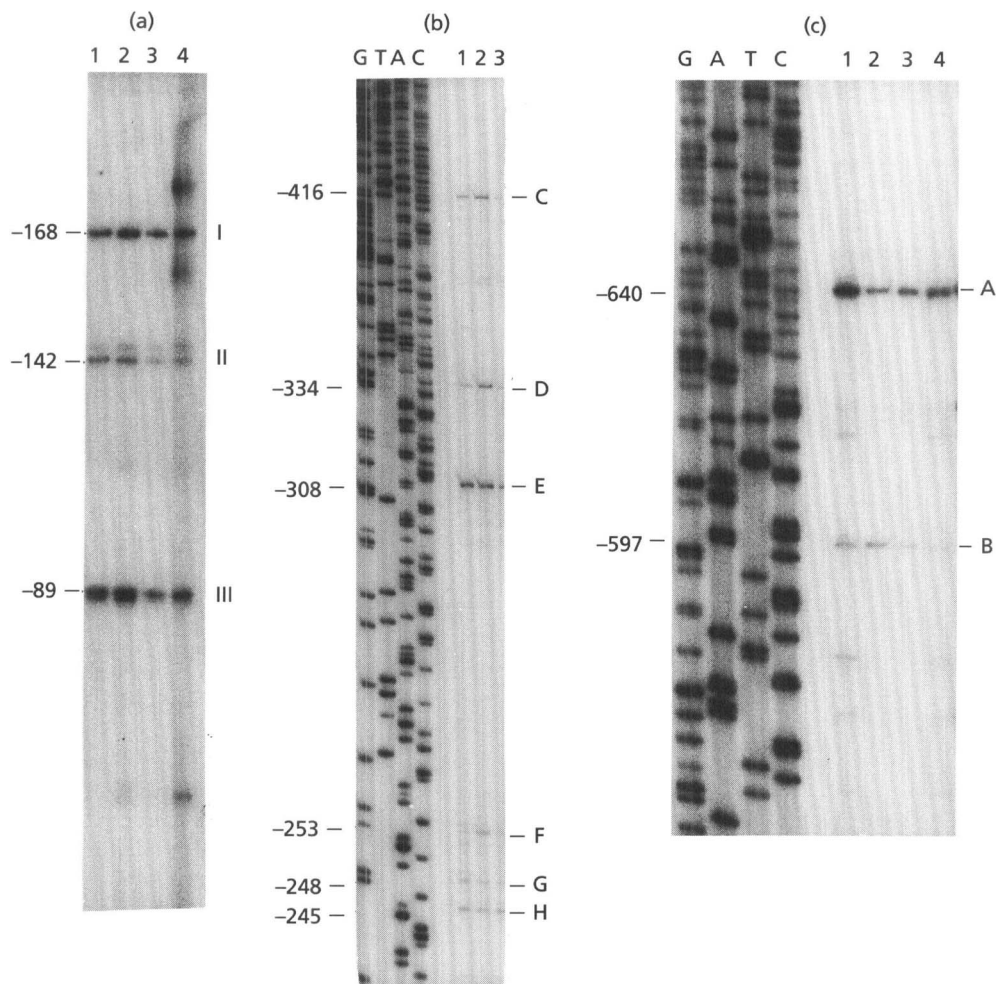
Fig. 3. Location of potential transcription start and processing sites by primer-extension using oligonucleotides gvw-1 (a), gvw-5 (b) and gvw-3 (c). Lanes 1–4: RNA isolated 12, 22, 29 and 42 h, respectively, after inoculation with ungerminated spores. Values to the left indicate the ends of the extension products with respect to the start of the mature 16S rRNA (Fig. 1a). Lanes G, A, T, C: nucleotide sequence determined with the same oligonucleotides as also used for primer extension.

4, seven protected bands were observed. Band F corresponds to full-length protection of the probe and may be due to DNA–DNA hybridization, a phenomenon that occurs regularly in nuclease S1 experiments. Bands P1–P4 correspond to protected fragments of approximately 450, 265, 185 and 105 nt, respectively, and are consistent with the possible sites of transcription initiation identified by primer extension, namely those starting at nt positions −597, −416, −334 and −254. Since it was difficult to determine the exact position of the transcription start site of the putative P2 promoter, we performed a nuclease S1 mapping experiment combined with Maxam and Gilbert sequencing using the *Bam*HI–*Bgl*II (−609 to −274) fragment as a probe, which indicated that the position shown in Fig. 6 is the most likely start site for P2 (data not shown).

Additional bands were found at positions R1 (approximately 100 nt in length) and R2 (less than 60 nt). These bands probably reflect incomplete hybridization of RNA

derived from *rrnB–F* to the *rrnA* probe. The bands around position R2 displayed a similar growth-phase-dependence to P3 and P4. Surprisingly, band R1 was only visible late in growth: it became apparent approximately 14 h after inoculation (corresponding to transition phase) and was present throughout stationary phase. It is tempting to suggest that this is due to the expression of a particular *rrn* operon late in growth, although we cannot rule out the possibility that band R1 originated from an additional (fifth) *rrnA* promoter (see Discussion).

In vitro transcription analysis

In vitro run-off transcriptions were performed to confirm the location of the *rrnA* promoters, which were identified by the *in vivo* transcription analyses described above, and to get an idea of which RNA polymerase holoenzymes are involved in the recognition of these promoters. Run-off transcriptions were performed using fractions of *S. coelicolor* RNA polymerase that had been partially purified
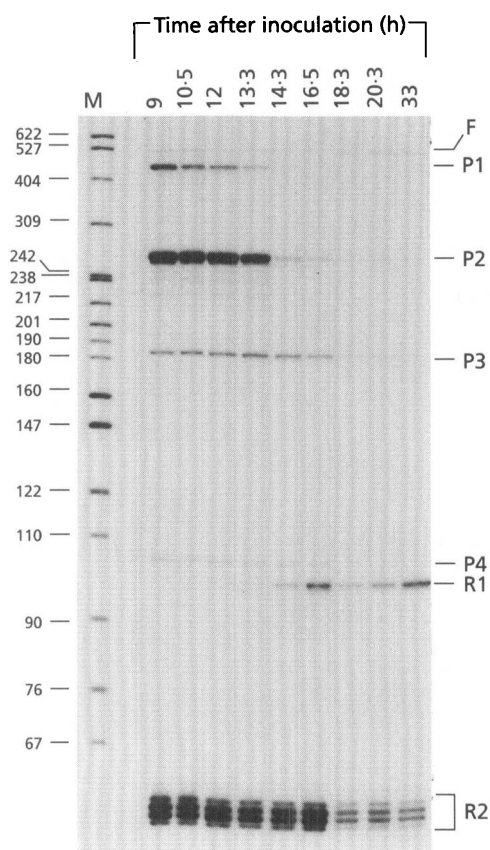
Fig. 4. Location of transcription end-points by nuclease S1 mapping. The probe used was the 557 nt *Fsp*I–*Hinc*II fragment (see Fig. 1). Bands P1–P4 represent promoter transcripts; band F indicates full length protection of the probe; R1 and R2 are bands that could not be ascribed to *rrnA* transcription starts. Lane M, DNA size markers (*Hpa*II-digested pBR322) in nt; the other lanes are labelled with time in hours after inoculation with germinated spores.

into different holoenyzme forms by Superose-6 FPLC chromatography (Buttner & Brown, 1985). The 557 bp *Fsp*I–*Hinc*II (−708 to −151), 377 bp *Bgl*II (−651 to −274) and 502 bp *Eco*RI–*Acc*I (−494 to −6, with an additional 14 bp of vector sequence at the *Acc*I site, since the fragment was subcloned into the polycloning site of pUC18 prior to digestion) fragments were used as templates (Fig. 1b). Five major run-off transcripts were observed with the 557 bp *Fsp*I–*Hinc*II fragment (Fig. 5a). Band F probably results from end-to-end transcription of the template. The four bands designated P1–P4 correspond to transcripts of approximately 445, 260, 165 and 100 nt and correspond to putative transcription start sites, identified by primer extension and nuclease S1 mapping, located at nt positions −597, −416, −334 and −254, respectively. An additional faint band was visible between the P1 and P2 transcripts. Since primer extension and nuclease S1 experiments failed to detect a corresponding transcription start site (see below) it may well be a result of artificial initiation by the RNA polymerase.

The locations and orientations of the putative promoters were verified by *in vitro* run-off transcription assays using

the 377 bp *Bgl*II fragment and the 502 bp *Eco*RI–*Acc*I fragment (including 14 bp of the pUC18 multiple cloning site sequence) with RNA polymerase fraction 28 (Fig. 5b). Run-off transcription using the *Bgl*II fragment generated three major transcripts of approximately 325, 145 and 65 nt. These correspond to the putative transcription start sites for P1, P2 and P3 already identified with the *Fsp*I–*Hinc*II fragment, located at nt positions −597, −416 and −334, respectively. The faint band F′ presumably represents end-to-end transcription.

In vitro transcription with the *Eco*RI–*Acc*I template also gave four bands. The transcripts of 425 nt, 345 nt and 265 nt correspond to the putative transcription start sites for P2, P3 and P4 located at nt positions −416, −334 and −254, respectively. Band X does not correspond to any of the transcripts produced by *in vitro* transcription of the *Fsp*I–*Hinc*II or *Bgl*II fragments, and may therefore be the result of artificial initiation of the RNA polymerase.

DISCUSSION

Putative processing sites

Primer extension with oligonucleotide gvw-1 revealed an RNA endpoint between −169 and −168 (band I in Fig. 3a). Transcripts from the *S. coelicolor rrnD* operon are processed at exactly the same position by RNaseIII (Baylis & Bibb, 1988) and, since the sequence of the *rrnA* operon between −200 and −1 is identical to that of *rrnD* (van Wezel *et al.*, 1991), it is extremely likely that this represents the major processing site for the *rrnA* precursors. This is also in accordance with the consensus found for the major processing site in other Gram-positive bacteria (Ogasawara *et al.*, 1983; Taschke & Hermann, 1986), including *S. ambofaciens* (Pernodet *et al.*, 1989). Furthermore, no transcription initiation was detected downstream of nt position −200 when the *Eco*RI–*Acc*I (−494 to −6) fragment was used as a template for *in vitro* run-off transcription with *S. coelicolor* RNA polymerases. Interestingly, an RNA end-point was also detected between the adenosines at positions −90 and −89 (band III in Fig. 3a) and occurs at the base of a computer-predicted stem–loop structure (Baylis & Bibb, 1988). The end-point is positioned exactly opposite the cleavage site that results in mature 16S rRNA. This suggests a mechanism with at least three maturation steps, resulting primarily in a precursor with a 168 nt leader, which is then further processed via an 89 nt leader into mature 16S rRNA. The physiological relevance of the RNA end-point between positions −143 and −142 (band II in Fig. 3a) remains unclear. Since all six rRNA operons are processed in the same way [no comparable band was found by nuclease S1 mapping (Baylis & Bibb, 1988)], it is likely that band II results from an artificial stop of the reverse transcriptase *in vitro*.

Identification of the transcription start sites of *rrnA*

The promoter-probing experiments predicted the presence of promoters in the regions −494 to −416, −416 to −318 and −318 to −199. These predictions appear to
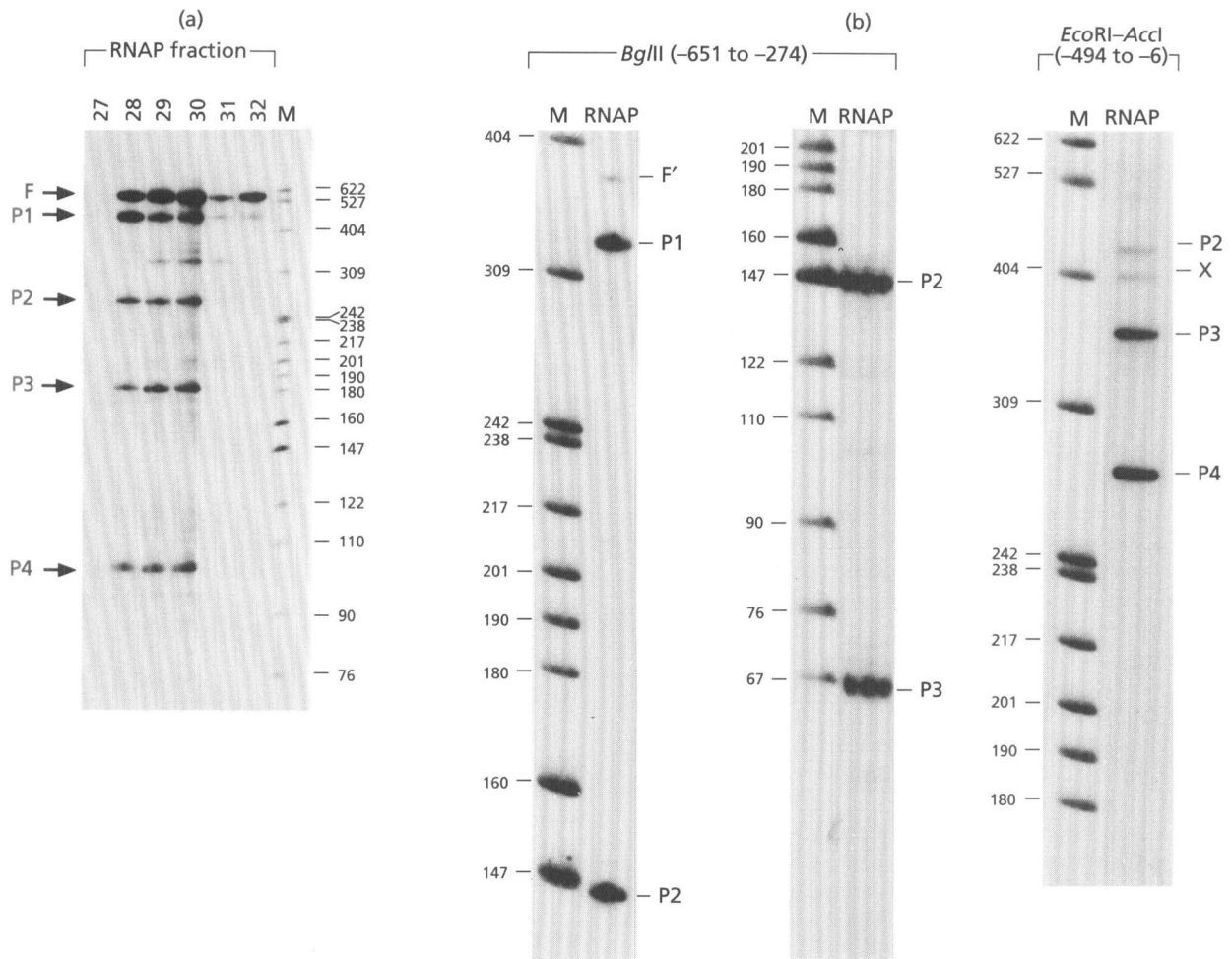
Fig. 5. *In vitro* transcription analysis of *rrnA*. Lane M, DNA size markers (*Hpa*II-digested pBR322) in bp. RNAP, different fractions of RNA polymerase from a Superose-6 FPLC column. Bands P1–P4 indicate *in vitro* run-off transcriptions originating at each of the promoters. (a) *In vitro* transcription using the *Fsp*I–*Hinc*II (−708 to −151) fragment as a template. F indicates end-to-end transcription of the RNA polymerase. (b) *In vitro* transcription using the *Bg*lII (−651 to −274) and the *Eco*RI–*Acc*I (−494 to −6) fragments as templates, using RNAP fraction 28. F′ indicates end-to-end transcription; band X was probably the result of artificial initiation of the RNA polymerase.

have been fulfilled by the identification, by *in vitro* transcription, primer extension and nuclease S1 mapping, of three promoters in the corresponding parts of the *rrnA* upstream region: P2 in the −494 to −416 region, P3 in the −416 to −318 region and P4 in the −318 to −199 region (see below).

Eight RNA end-points were observed by primer extension with reverse transcriptase in the region upstream of nt position −200. Since oligonucleotide gvw-3 appears to be specific for *rrnA*, and since gvw-5 hybridized most strongly to DNA fragments corresponding to *rrnA*, most of the transcripts detected by primer extension will probably have arisen from *rrnA*. It is not clear to what extent the RNA 5′ ends detected with gvw-5 could have originated from the second DNA segment to which gvw-5 anneals (Fig. 2).

RNA end-points detected by primer extension at positions −597, −416, −334 and −254 were also detected by

nuclease S1 mapping. Furthermore, *in vitro* run-off transcription experiments with *S. coelicolor* RNA polymerase using three different templates revealed that these positions are transcription initiation sites *in vitro*. Together, these data strongly suggest that the RNA end-points at −597, −416, −334 and −254 correspond to sites of transcription initiation from four promoters, P1–P4, respectively. The other bands observed after primer extension (at positions −640, −308, −248 and −245) may reflect non-specific stops by reverse transcriptase (perhaps due to RNA secondary structure) or transcription start sites of the *rrnE* operon.

Sequence comparison of the *rrnA* promoters

The promoters P1 and P2 are similar to the consensus sequence for the major class of eubacterial promoters (Hawley & McClure, 1983), with P1 showing a perfect fit to the canonical −35 region. P3 and P4 lack a consensus

```
                  -35                    -10        .
P1    -642  AGCGACGCTTGACAGCCCTCTGCGGGATCCGTAAGGTTCTTCGGGT

                  -35                   -10      .
P2    -460  CGACCCGATTGGGAATCGCCGAGAGGGTCCGCTAAGGTTTGAGA

                -35                    -10          .
P3    -380  CCTGACGGGGAATCAGGGCCCGAAAGGATCTGATAGAGTCGGAACCG

                -35                   -10        .
P4    -298  GCACCGAGGAAATCGGATCGGAAAGATCTGATAGAGTCGGAAACG
```

Fig. 6. Alignment of the *S. coelicolor rrnA* promoters. Putative −35 and −10 sequences are underlined and transcription start sites (determined by primer extension) are indicated by dots. The sequence GG(G/A)ATC that occurs in the P2, P3 and P4 −35 regions is given in bold type. Numbers refer to the sequence shown in Fig. 1(a).

−35 region (Fig. 6). Sequence alignment of P1, P2, P3 and P4 (Fig. 6) shows a striking similarity between the P3 and P4 promoter regions, which are almost identical between positions −1 and −25 with respect to the transcription start site, and have similar −35 regions. These promoters are almost identical to the P3 and P4 promoter of the *S. coelicolor rrnD* operon (Baylis & Bibb, 1988). Interestingly, the sequence of *rrnA* P2 around nt position −35 is comparable to that of *rrnA* P3 and P4, namely G(G/A)AATC. Boxes with high homology to the sequence G(G/A)AATC can also be seen in the promoter regions of *hrdD* P2 (GAATTC, ending at −26; the gene encodes a sigma factor homologous to σ^{hrdB}), *aph*P2 (GGATTC, ending at −31; the gene encodes aminoglycoside phosphotransferase) and pA1 (GAACTC, ending at −31), a promoter which overlaps *aph* P1, but is oppositely oriented [see Strohl (1992) and references therein].

In addition to the 57 nt direct repeat that encompasses the −35 and −10 regions of P3 and P4, an additional direct repeat of 22 nucleotides is located at nt positions −620 to −599 and −591 to −570 in the region upstream of the 16S rRNA gene (see Fig. 1a), and flanks the P1 transcription start site.

Some clues about recognition of the *rrnA* promoters are given by the results of the *in vitro* transcription assay with the *FspI–HincII* (−708 to −151) fragment as the template (Fig. 5a). We used various RNA polymerase holoenzyme fractions, as used previously for *in vitro* run-off transcription of the four markedly heterogeneous *dagA* promoters (Buttner *et al.*, 1987). The *rrnA* P1 promoter, which conforms to the consensus sequence for vegetative promoters, is most efficiently transcribed by RNA polymerase present in fraction 28. This is the fraction which is enriched for the RNA polymerase holoenzyme containing σ^{hrdB}, the vegetative sigma factor, which also transcribes the *dagA* P4 promoter (M. J. Bibb, unpublished data). This result, in combination with the high homology of P1 with the vegetative promoter consensus sequence, makes it likely that the P1 promoter is recognized by σ^{hrdB}. For P2–P4, it is less clear which sigma factor is responsible for recognition. Clearly the P3 and P4 promoters are quite similar and the apparent lack of a sequence around the −35 position comparable to the consensus for vegetative promoters suggests that a sigma factor other than σ^{hrdB}

might be responsible for the recognition of these promoters. The P2 promoter is similar to the consensus sequence for the major class of eubacterial promoters, but shares a sequence around its −35 position with P3 and P4.

Relative contribution of the promoters to transcription *in vivo*

The half-life of rRNA precursors in *S. coelicolor* is approximately 40 s (Clayton & Bibb, 1990b). In our experiments, mycelial samples were taken at hourly intervals and therefore we can assume that the observed steady-state levels of the primary transcripts represent *de novo* RNA synthesis. P2 appears to be the most frequently used promoter in fast growing cultures. P1 and P3 are used to similar extents, and are considerably weaker, whereas P4 is the weakest promoter. This contrasts with the results obtained with *rrnD*, where P3 and P4 together contributed at least 85 % of *rrnD* transcription in rapidly growing cultures (Strauch *et al.*, 1991). The discrepancy is surprising since the sequences around P3 and P4 are almost identical for *rrnA* and *rrnD*.

We believe that nuclease S1 mapping is more reliable for comparing transcript levels than primer extension. In conditions of DNA probe excess, all RNA molecules will be hybridized to the probe and every RNA molecule should contribute to signal strength. This is not always the case for primer-extension experiments, since signal intensity will depend on the ability of the reverse transcriptase to transcribe to the end of the transcript, an event which may occur much less frequently for long RNA molecules than for short ones. In addition, factors such as stem–loop structures in the RNA molecules influence the amount of full-length cDNA made.

Growth-phase-dependent expression of the *rrn* operons

Transcription from P1 and P2 in SMM fell markedly between 13 and 17 h after inoculation, which corresponds to the transition between exponential growth and stationary phase. However, transcription from P3 and P4 persists until 17 h after inoculation, and consequently P3 appears to replace P2 as the most frequently used promoter in the transition phase. This observation underlines the need to study relative promoter activities throughout a growth curve.

While the R2-protected fragments are apparent in the early stages of growth, the R1 band did not appear until 14 h after inoculation. This suggests that one of the six *rrn* operons may be expressed only late in growth. Such growth-phase-dependent expression might reflect a role for this operon in either morphological differentiation or antibiotic production, and is consistent with the observation by Granozzi *et al.* (1990) of a second phase of stable RNA synthesis on solid medium during the transition from vegetative to aerial mycelium. The intriguing possibility that one of the *rrn* operons may be expressed late in growth will be investigated by com-

paring the transcription of all six *rrn* operons. This should yield information about the relative contribution of the various *rrn* operons to *de novo* rRNA synthesis throughout the *S. coelicolor* life-cycle.

ACKNOWLEDGEMENTS

REFERENCES

Bascaran, V., Sanchez, L., Hardisson, C. & Brana, A. F. (1991). Stringent response and initiation of secondary metabolism in *Streptomyces clavuligerus. J Gen Microbiol* **137**, 1625–1634.

Baylis, H. A. & Bibb, M. J. (1988). Transcriptional analysis of the 16S rRNA gene of the *rrnD* gene set of *Streptomyces coelicolor* A3(2). *Mol Microbiol* **2**, 569–579.

Buttner, M. J. & Brown, N. L. (1985). RNA polymerase–DNA interactions in *Streptomyces. In vitro* studies of a *S. lividans* plasmid promoter with *S. coelicolor* RNA polymerase. *J Mol Biol* **185**, 177–188.

Buttner, M. J., Fearnly, I. M. & Bibb, M. J. (1987). The agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2): nucleotide sequence and transcriptional analysis. *Mol & Gen Genet* **209**, 101–109.

Chater, K. F. (1989). Aspects of multicellular differentiation in *Streptomyces coelicolor* A3(2). In *Genetics and Molecular Biology of Industrial Microorganisms*, pp. 99–107. Edited by C. L. Hershberger, S. W. Queener & G. Hegeman. Washington: American Society for Microbiology.

Clayton, T. M. & Bibb, M. J. (1990a). *Streptomyces* promoter-probe plasmids that utilise the *xylE* gene of *Pseudomonas putida*. *Nucleic Acids Res* **18**, 1077.

Clayton, T. M. & Bibb, M. J. (1990b). Induction of a φC31 prophage inhibits rRNA transcription in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **4**, 2179–2185.

Condon, C., Philips, J., Fu, Z.-Y., Squires, C. & Squires, C. L. (1992). Comparison of the seven ribosomal RNA operons in *Escherichia coli. EMBO J* **11**, 4175–4185.

Gourse, R. L., de Boer, H. A. & Nomura, M. (1986). DNA determinants of rRNA synthesis in *E. coli*, growth rate dependent regulation, feedback inhibition, upstream activation, antitermination. *Cell* **44**, 197–205.

Gramajo, H. C., Takano, E. & Bibb, M. J. (1993). Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Mol Microbiol* **7**, 837–845.

Granozzi, C., Billetta, R., Passantino, R., Sollazzo, M. & Puglia, A. M. (1990). A breakdown in macromolecular synthesis preceding differentiation in *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **136**, 713–716.

Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**, 2237–2255.

Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M. & Schrempf, H. (1985). *Genetic Manipulation of Streptomyces: a Laboratory Manual.* Norwich: John Innes Foundation.

Murray, M. G. (1986). Use of sodium trichloroacetate and mung bean nuclease to increase sensitivity and precision during transcript mapping. *Anal Biochem* **158**, 165–170.

Ochi, K. (1986). Occurrence of the stringent response in *Streptomyces sp.* and its significance for the initiation of morphological and physiological differentiation. *J Gen Microbiol* **132**, 2621–2631.

Ochi, K. (1987). Metabolic initiation of differentiation and secondary metabolism by *Streptomyces griseus*: significance of the stringent response (ppGpp) and GTP content in relation to A factor. *J Bacteriol* **169**, 3608–3626.

Ogasawara, N., Moriya, S. & Yoshikawa, H. (1983). Structure and organization of rRNA operons in the region of the replication origin of the *Bacillus subtilis* chromosome. *Nucleic Acids Res* **11**, 6301–6318.

Pernodet, J.-L., Boccard, F., Alegre, M.-T., Gagnat, J. & Guérineau, M. (1989). Organization and nucleotide sequence analysis of a ribosomal RNA gene cluster from *Streptomyces ambofaciens. Gene* **79**, 33–46.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Stern, S., Moazed, D. & Noller, H. F. (1988). Analysis of RNA structure using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol* **164**, 481–489.

Strauch, E., Takano, E., Baylis, H. A. & Bibb, M. J. (1991). The stringent response in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **5**, 289–298.

Strohl, W. R. (1992). Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res* **20**, 961–974.

Summerton, J., Atkins, T. & Bestwick, R. (1983). A rapid method for preparation of bacterial plasmids. *Anal Biochem* **133**, 79–84.

Takano, E., Gramajo, H. C., Strauch, E., Andres, N., White, J. & Bibb, M. J. (1993). Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **6**, 2797–2804.

Taschke, C. & Herrmann, R. (1986). Analysis of transcription and processing signals of the 16S–23S rRNA operon of *Mycoplasma hyopneumoniae. Mol & Gen Genet* **205**, 434–441.

van Wezel, G. P. (1994). *Transcriptional regulation of translational genes in Streptomyces coelicolor A3(2).* PhD thesis, Leiden University.

van Wezel, G. P., Vijgenboom, E. & Bosch, L. (1991). A comparative study of the ribosomal RNA operons of *Streptomyces coelicolor* A3(2) and sequence analysis of *rrnA. Nucleic Acid Res* **19**, 4399–4403.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Zukowski, M. M., Gaffney, D. F., Speck, D., Kaufmann, M., Findeli, A., Wisecup, A. & Lecocq, J.-P. (1983). Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene. *Proc Natl Acad Sci USA* **80**, 1101–1105.