The mechanism of upstream activation in the \textit{rrnB} operon of \textit{Mycobacterium smegmatis} is different from the \textit{Escherichia coli} paradigm

Kristine B. Arnvig, B. Gopal, K. G. Papavinasasundaram, Robert A. Cox and M. Joseph Colston

Division of Mycobacterial Research and Division of Protein Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Mycobacteria are slow-growing bacteria with a generation time of from 2–3 h up to several weeks. Consistent with the low growth rate, mycobacterial species have a maximum of two rRNA operons, \textit{rrna} and \textit{rrnB}. The \textit{rrna} operon is present in all mycobacteria and has between two and five promoters, depending on species, whereas the \textit{rrnB} operon, with a single promoter, is only found in some of the faster-growing species. The promoter region of the \textit{rrnB} operon of a typical fast grower, \textit{Mycobacterium smegmatis}, was investigated. By using \textit{lacZ} reporter gene fusions it was demonstrated that the \textit{rrnB} operon contains a highly activating region upstream of the core promoter, comparable to other bacterial \textit{rrn} operons. However, the results suggest that, unlike the situation in, for example, \textit{Escherichia coli}, the activating mechanism is solely factor dependent, and that no UP element is involved.

\section*{INTRODUCTION}

The genus \textit{Mycobacterium} consists of more than 80 species with large variations in growth rates (Wayne & Kubica, 1986). Mycobacteria have either one or two \textit{rrn} operons with different promoter architecture between operons as well as between species. The \textit{rrna} operon, located downstream of the \textit{murA} gene, is found in all species, including the human pathogens \textit{Mycobacterium tuberculosis} and \textit{Mycobacterium leprae}, but the number of promoters varies (Gonzalez-y-Merchand et al., 1997). An additional operon, \textit{rrnB}, which is found in some of the faster-growing species, is located downstream of the \textit{tyrS} gene and has a more conserved structure with a single promoter in all cases so far studied (Gonzalez-y-Merchand et al., 1999; Menendez et al., 2002). Both operons are differentially regulated and expressed depending on the growth stage of the bacterium (Gonzalez-y-Merchand et al., 1998; Verma et al., 1999). The −35 and −10 elements have been assigned by mapping of transcription start points, and additional putative elements have been identified by sequence comparison, but not by function (Gonzalez-y-Merchand et al., 1997, 1999; Menendez et al., 2002).

In \textit{Escherichia coli} the intrinsic strength of \textit{rrn} promoters is greatly increased by the presence of an upstream activating region (UAR) (Condon et al., 1995). The \textit{E. coli} UAR consists of two distinct components: the UP element which is factor independent (Newlands et al., 1991) and an element consisting of between three and five binding sites for the transcription factor Fis (Hirvonen et al., 2001). Both mechanisms require the C-terminal domain (CTD) of the RNA polymerase alpha subunit (RNAP\textsubscript{α}) for activation (Bokal et al., 1997; Ross et al., 1993). However, a major difference between the two lies in the positioning requirements relative to the core promoter. Factor-dependent activation, i.e. activation by Fis (or other class I transcription factors), is face-of-the-helix dependent; in other words the activator binding sites can be separated from the core promoter and still retain most of their activation potential, provided that they remain on the same face of the DNA helix (Bokal et al., 1997; Gaston et al., 1990; Ishihama, 1993; Newlands et al., 1992). In contrast, factor-independent activation involving the UP element requires a contact between the promoter proximal alpha subunit and the sigma factor: the element must be immediately upstream of, as well as in phase with, the −35 box (Hirvonen et al., 2001; Meng et al., 2001). Little is known about bacterial \textit{rrn} regulatory elements other than those of \textit{E. coli}, although a number of \textit{rrn} operons from other species have been sequenced and/or investigated. Most of these appear to harbour UP elements, as shown either by function or by the presence of characteristic AT-rich tracts immediately upstream of the −35 region (Aiyar et al., 2002; Amador et al., 1999; Dryden & Kaplan, 1993; Garnier et al., 1991;...
Helmann, 1995; La Fontaine & Rood, 1996; Zahn et al., 2001). Mycobacterial rrr operons do not contain phased AT-rich tracts characteristic of UP elements (Gonzalez-y-Merchand et al., 1997). However, in the thermophile Thermus thermophilus, sequence-specific interactions between RNAPz and a putative UP element with a different nucleotide composition have been reported, suggesting that a high AT content is not a universal requirement (Wada et al., 2000).

The objective of the present study was to dissect the promoter region of the rrrB operon of Mycobacterium smegmatis in order to identify and characterize activating strength of the P1B promoter and shown that its upstream reporter gene fusions, we have characterized the intrinsic regulatory regions in this operon. Using primers (Stratagene) according to manufacturer’s instructions and with PCR amplified from E. coli were grown in liquid L broth with 50 μM sodium phosphate buffer, 250 mM NaCl, pH 8.0 and sonicated. Post-lysis, the cell extract was cleared by centrifugation, and the proteins in the supernatant were bound to the Talon resin pre-equilibrated in buffer A. The Talon column was washed with 10 column volumes of buffer B (50 mM sodium phosphate buffer, 250 mM NaCl, 5 mM imidazole, pH 7–5) and the protein was subsequently eluted from the column in buffer C (50 mM sodium phosphate buffer, 250 mM NaCl, 300 mM imidazole, pH 7–0). The eluted protein was about 85% pure and was subjected to size exclusion chromatography on a Pharmacia S-75 column in 50 mM Tris, 250 mM NaCl, pH 7–5.

Electrophoretic mobility shift assays (EMSA). The entire region of RNAPz known to interact with the UP element is 100% conserved between M. tuberculosis and M. smegmatis and so the M. tuberculosis RNAPz was used in all experiments. The binding reaction was performed at room temperature in 20 μl reaction buffer: 20 mM HEPES, pH 7–5, 0.2 mM EDTA, 10 mM (NH4)2SO4, 1 mM DTT, 15 mM MgCl2, 15 mM KCl, 50 μg BSA ml–1 and 2 μg poly dI : dC ml–1. The samples were electrophoresed on native 8% acrylamide gels in 0.5 x TBE, dried and exposed to a phosphor imager screen.

RESULTS

The rrrB minimal promoter

The structure and sequence of the M. smegmatis rrrB P1 promoter region has been outlined in Fig. 1. The proposed P1B promoter elements have a similar structure to E. coli α promoters, and constructs harbouring these elements are recognized in E. coli, as seen by blue colony colour on X-Gal plates (data not shown). To determine the strength of the P1B promoter fragments, we used the lacZ-based promoter probe vector pEJ414 (Papavinasasundaram et al., 2001). The region containing the attachment site and integrase gene of the mycobacteriophage L5, a kanamycin resistance cassette and a promoterless lacZ gene. Integration is thus site-specific and required for transformation of mycobacteria. The cloned P1B fragments including their exact boundaries are shown in Fig. 1. Reporter constructs were transformed into M. smegmatis mc2155 and the promoter activity was determined as the β-gal activity of the cell-free extracts.

Initially we wanted to determine the intrinsic strength of a P1B minimal promoter fragment as our point of reference. The region spanning the −10 and −35 hexamers (i.e. −36 to −8 relative to the transcription start point) was inserted into the promoter probe vector, and the resulting reporter construct gave rise to a β-gal activity of 24 units (mg protein)−1. For comparison, we made a construct with the M. tuberculosis rrrA PCL1 promoter fragment covering the same region, i.e. −36 to −8. This construct gave rise to a β-gal activity of almost 200 units (mg protein)−1, i.e. approximately eight times the activity of the corresponding

METHODS

Plasmids and strains and media. Reporter constructs were cloned in pEJ414 (Papavinasasundaram et al., 2001) as XbaI–HindIII fragments with the cloned rrrB region indicated in Fig. 1. Promoter inserts of less than 100 bp were created by annealing two complementary oligonucleotides containing the promoter fragment, including the desired overhangs. All other promoter inserts were created by PCR amplification with primers containing the desired restriction sites. The inserts for the phased mutations were created by end-filling with Klenow (Promega) before ligation or by site-directed mutagenesis (Quickchange kit, Stratagene). Ligation was performed into E. coli DH5α; the resulting plasmids were isolated, sequenced and retransformed into M. smegmatis mc2155 (Novagen). The gene was PCR amplified from M. tuberculosis genomic DNA with PfuTurbo (Stratagene) according to manufacturer’s instructions and with primers 5′-TCGAATCCGATATGCTGATCTCACAG-3′ and 5′-ATGAGATCCCTTCGCGGTCGCTGTGG-3′. E. coli was grown in liquid L broth with 50 μg kanamycin ml–1. Overnight cultures of E. coli were diluted 500-fold into fresh medium and grown to OD600 0.5–0.8, measured by a UNICAM UV2 spectrophotometer, before harvesting. M. smegmatis was inoculated directly from 7H11 agar plates (Difco) into 50 ml Dubos broth (Difco) supplemented with 15 μg kanamycin ml–1 and grown overnight to OD600 0.5–0.8.

Bacterial cell extracts. Cell extracts were prepared as described by Papavinasasundaram et al. (2001). An aliquot of each sample was used to determine the total protein concentration using a BCA protein assay kit (Pierce). β-Galactosidase (β-gal) activity was determined as described by Miller (1972), normalized to total protein concentration and expressed as units (mg protein)−1. The promoter activities were assayed in at least three individual cultures for each strain.

Expression and purification of M. tuberculosis RNAPz. The plasmid encoding M. tuberculosis rrrA (pK2a) was transformed into E. coli BL21 (DE3) pLysS. The cells were grown in Terrific Broth (Sambrook et al., 1989) and induced with 1 mM IPTG when the OD600 reached 0.6. After a further 6 h growth at 30 °C, the culture was harvested by centrifugation for 30 min at 3000 r.p.m. RNAPz of M. tuberculosis was highly soluble. The His-tagged protein was purified by affinity chromatography on a cobalt resin column (Talon, Clontech). Briefly, the cells were resuspended in buffer A (50 mM phosphate buffer, 250 mM NaCl, pH 8.0) and sonicated.
P1B promoter fragment (data not shown). Next we compared our minimal promoter construct with constructs that had been extended downstream to either +1 or +10. Both of these 3′-end additions gave rise to a small increase in β-gal activity relative to the minimal construct (see Fig. 1). Thus, extending the promoter insert to include the native transcription start point increased activity 1.5-fold, while a further extension to +10 relative to the transcription start point increased activity another twofold (Fig. 1). These results demonstrate that during exponential growth the P1B minimal promoter gives rise to a relatively low β-gal activity, compared to that of the M. tuberculosis PCL1 minimal promoter fragment, and moreover that the region downstream of −8 supports only a limited increase in β-gal activity.

The rrnB upstream activating region

Our initial experiments indicated that the activity of the P1B promoter was significantly increased by including the region upstream of the −35 hexamer (K. B. Arnvig, unpublished data). Therefore a detailed investigation of the P1B UAR was carried out in order to identify putative activating elements. A series of P1B constructs was made in which the 3′-end was maintained at −8, as in the minimal promoter construct, while the UAR was gradually extended to +200 relative to the transcription start point. The results, shown in Fig. 1, demonstrate how the promoter activity of P1B increases significantly as the UAR is extended towards +200, reaching a maximum β-gal activity of 8758 units (mg protein)⁻¹, which corresponds to more than 350-fold activation of the minimal construct. The most significant increase in activity (18-fold) is observed as the UAR is extended from −39 to −52. Interestingly, this fragment contains a highly conserved sequence found in a number of mycobacterial rrn operons, indicated as UR1 in Fig. 1 (Gonzalez-y-Merchand et al., 1997). The region between −52 and −39, which contains the tyrS stop codon as well as a second UR element (UR2), increases β-gal activity another three- to fourfold, whereas no increase, but in fact

Fig. 1. The M. smegmatis rrnB promoter region. The figure shows the sequence of the rrnB promoter region from −80 to +10. Reporter constructs are shown with rrnB sequences in upper case and vector sequences in lower case. The promoter elements (−35 and −10) and the tyrS stop codon (−56 to −54) are underlined. Numbers indicate the boundaries of each construct and are relative to the P1B transcription start point, indicated by arrows (the transcription start point was not determined for transcripts with 3′ end at −8). UR (upstream region) indicates a conserved sequence found in a number of mycobacterial rrn operons (see text). Boxed sequences represent repeated elements containing the UR. The β-gal activity obtained for each construct is shown in the column on the right and is expressed as units (mg protein)⁻¹. The empty vector gave a β-gal activity of 0.5–1.0 units (mg protein)⁻¹.
a slight decrease, is observed between −70 and −80. Finally, there is another activating region between −80 and −200, which increases β-gal activity approximately fivefold. However, the increase from −140 to −200 is minor (1.1-fold), indicating that most of the activating sequences of the rrnB operon are located within the first 100 nt of the UAR. As a control, the UAR was cloned into pEJ414, but showed no detectable promoter activity on its own in either orientation (data not shown).

The activity of the P1B UAR was subsequently determined in constructs with the 3′ end at +1 or +10 in order to establish if the region between −8 and +10 affected UAR activity or vice versa, i.e. if they act independently. The results, taken from the activities in Fig. 1, indicate that the UAR (−140 to −37) activates the P1B promoter 320-fold with the 3′ end at −8, 450-fold with the 3′ end at +1 and 360-fold with the 3′ end at +10. Conversely, the contribution from the region between −8 and +10 varies from 2.8-fold (with the 5′ end at −36) to 3.1-fold (with the 5′ end at −140). We regard these variations as minor and therefore conclude that the two regions act independently.

**Lack of sequence-specific interaction between the rrnB UAR and RNAPα**

A number of bacterial promoters, such as *E. coli rrm* promoters, contain a third element, in addition to the −10 and −35 boxes, located immediately upstream of the −35 box (Newlands et al., 1991; Rao et al., 1994). This so-called UP element consists of AT-rich tracts in phase with the −35 box, and it has been shown to interact with RNAPα in a sequence-specific manner (Estrem et al., 1998; Rao et al., 1994; Ross et al., 1993). The *M. smegmatis* P1B UAR does not contain an apparent UP element sequence, but it is possible that an alternative sequence may function as a UP element in mycobacteria, analogous to the situation in *T. thermophilus* (Wada et al., 2000). A series of EMSAs was performed to establish whether or not the mycobacterial RNAPα interacts in a sequence-specific manner with the P1B UAR. Results of EMSAs performed with 104 bp P1B UAR (from −140 to −37) and various concentrations of purified RNAPα are shown in Fig. 2(a). The results demonstrate that there is complex formation between the P1B UAR and RNAPα at protein concentrations starting at 20 μM RNAPα, and that 75 μM RNAPα is required to shift more than 50% of the probe. For comparison, the *E. coli* UP element DNA displays 100% complex formation at 4 μM RNAPα (Ross et al., 2001), and we therefore consider the P1B UAR interaction to be of low affinity. Fig. 2(b) illustrates how the RNAPα–UAR complex can be disrupted equally well by the addition of excess unlabelled UAR DNA and non-UAR DNA (compare lanes 3–5 with lanes 6–8). To ensure that the non-UAR competitor did not contain UAR-like determinants by coincidence, the experiment was repeated with another non-UAR DNA (a 147 bp *Pvu*–*FspI* pUC19 DNA fragment) but, as before, the non-UAR and UAR DNA fragments competed to the same extent (data not shown). These results strongly suggest that the interaction between the UAR and RNAPα is not sequence specific. The ability of RNAPα to interact with DNA independent of sequence was further supported by EMSAs in which non-UAR and UAR DNA formed complexes with RNAPα equally well (data not shown). The results demonstrate the absence of sequence-specific interaction between the rrnB UAR and the RNAPα subunit, which in turn suggests that there is no UP element in the mycobacterial rrnB UAR.

![Fig. 2](image-url)

**Fig. 2.** Interactions between RNAPα and the rrnB UAR. (a) EMSA using *M. tuberculosis* RNAPα and the UAR (−140 to −37) region of rrnB. Lane 1, free UAR probe; lanes 2–8, UAR probe and 20, 30, 40, 50, 75, 100 and 150 μM RNAPα, respectively. (b) Competition between UAR probe and unlabelled UAR (Spec) DNA or non-UAR (Non spec) DNA (a 120 bp *EcoRI–PvuI* fragment of pUC19). Lane 1, free UAR probe; lane 2, UAR probe and 50 μM RNAPα; lanes 3–5, as lane 2 with 1-, 10-, 50-fold excess unlabelled UAR DNA; lanes 6–8, as lane 2 with 1-, 10-, 50-fold excess unlabelled pUC19 DNA.
Face-of-the-helix-dependent activation of the *rrnB* promoter

The apparent absence of a UP element in the *rrnB* UAR was subsequently tested with a phasing experiment. Due to the highly flexible interdomain linker of the bacterial RNAPz, the zCTD can move more than 4 nm away from the rest of the RNAP molecule (Blatter et al., 1994; Zhou et al., 1994). This implies that promoters activated by class I transcription factors (Ishihama, 1993) can be separated from the upstream region by an integral turn of the DNA helix and still retain most of their activity. If the activating element and the core promoter are separated by non-integral helical turns, such as half or three-quarter turns, the activity will still retain most of their activity. If the activating element is likely to be recognized in *E. coli* due to the conservation of RNAPz. If, on the other hand, activation of the P1B promoter were purely factor mediated, then there would be a high probability that recognition of the UAR was restricted to mycobacteria and related species. Therefore, in order to determine the ability of the P1B UAR to activate the P1B promoter in a different organism, we measured the β-gal activities of the P1B promoter with and without UAR in *E. coli*. Curiously, the minimal construct (−36 to −8) gave a β-gal activity of 614 ± 28 units (mg protein)^−1, whereas the construct with the UAR (−140 to −8) gave a β-gal activity of 145 ± 16 units (mg protein)^−1. In other words, the *rrnB* UAR confers a fourfold decrease in promoter activity when expressed in *E. coli*. Therefore the *M. smegmatis rrnB* UAR is not recognized as an activating element in *E. coli*. Combined with the known sequence conservation of the RNAPz subunit, we believe this supports the notion of factor-dependent activation in the *M. smegmatis rrnB* operon.

**DISCUSSION**

In the current study we have dissected the promoter region of the *M. smegmatis rrnB* operon and found that sequences upstream as well as downstream of the basic promoter elements (−35 and −10 boxes) appear to contribute to the overall activity of the P1B promoter. However, the contribution from the region downstream of −8 is marginal. Moreover, we cannot exclude the possibility that constructs with different transcription start sites have different lacZ activity.

**Species-specific activation of the P1B promoter**

This finding is consistent with a factor-dependent mechanism and thus supports the notion that the *rrnB* UAR does not contain a UP element.

---

**Table 1. Face-of-the-helix-dependent activation**

The Sequence column shows the sequence immediately upstream of and including the −35 box; the latter is shown in bold italic type. β-gal activity [in units (mg protein)^−1] is given as a percentage of wild-type (KAM87) activity.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Sequence</th>
<th>No. base pairs inserted</th>
<th>Helical turns</th>
<th>β-gal activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAM87</td>
<td>ATTTGACT</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>KAM92</td>
<td>ATATTTGACT</td>
<td>2</td>
<td>1/5</td>
<td>16</td>
</tr>
<tr>
<td>KAM74</td>
<td>ATCTAGATTGACT</td>
<td>6</td>
<td>1/2</td>
<td>1-2</td>
</tr>
<tr>
<td>KAM70</td>
<td>ATCTAGACTTTGACT</td>
<td>8</td>
<td>3/4</td>
<td>0-4</td>
</tr>
<tr>
<td>KAM76</td>
<td>ATCTAGCTAGAGTTGACT</td>
<td>11</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>KAM75</td>
<td>ATCTAGCTAGAGATTGACT</td>
<td>13</td>
<td>5/4</td>
<td>74</td>
</tr>
</tbody>
</table>
mRNA stability, which in turn could affect the resulting β-gal activity. However, the rrrB UAR is highly activating during exponential growth and comparable in strength to the E. coli rrr UARs (Hirvonen et al., 2001). Unlike upstream activation of E. coli rrr promoters, the activation of the P1B promoter was found to be face-of-the-helix dependent, a strong indication of factor-dependent activation (Hirvonen et al., 2001; Meng et al., 2001). Moreover, we have demonstrated the absence of sequence-specific interactions between the M. smegmatis rrrB UAR and RNAP, another indication that the rrrB UAR contains no UP element. The lack of activation by the rrrB UAR in E. coli also suggests that there is no UP element involved. It is possible that a putative mycobacterial UP element would not be recognized in E. coli but, based on the fact that the amino acids known to interact with the E. coli UP element are 100% conserved between mycobacteria and E. coli, we believe that the lack of activation by the rrrB UAR in E. coli points towards factor-dependent activation. The somewhat unexpected higher activity of the core promoter construct compared to the full-length construct in E. coli may be due to vector sequences, which in the absence of UAR are located immediately upstream of the P1B promoter. Since this region consists of transcriptional terminators in the form of T-tracts, it is possible that these could act as weak UP elements.

The region that conferred the highest increase in promoter activity is located between −41 and −52. This segment contains an element, TCTGACC3/4GGG, which is repeated exactly two helical turns further upstream (see Fig. 1). A similar repeat is present in the rrrB operon of Mycobacterium fortuitum, although the sequence has been changed slightly from TCTGACCAGGG-N5-TCTGACCCTGGG in M. smegmatis to TCTGACCTGGC-N5-AATGACCTGGC in M. fortuitum. Moreover, the core sequence, PuACCNG, is also found in a single copy immediately upstream of most, if not all rrrA promoters (Gonzalez-y-Merchand et al., 1997). We therefore propose this to be a likely recognition site for a transcriptional activator common to both rrrA and rrrB and thus to all mycobacteria. We did not identify a similar sequence in the region from −8 to −140. Due to the different relative positions of tyrS and P1B in M. smegmatis and M. fortuitum, the location of this element differs so that it overlaps with the tyrS coding sequence in M. smegmatis but not in M. fortuitum. Whether this has implications for the function of the element has not been determined, although it is clear that such an overlap puts some constraints on the sequence. This type of arrangement, a promoter region within the coding region of an upstream gene, is not unknown and has been described before in both E. coli and cyanobacteria (Bogner et al., 1989; Grundstrom & Jaurin, 1982; Plansangkate et al., 2004).

The overall conclusion from these experiments is that one or more transcriptional activators are likely to be responsible for the more than 300-fold activation seen in the rrrB operon. As a comparison, factor-dependent activation is less than tenfold in E. coli, whereas factor-independent activation accounts for the majority of the total transcriptional activation in the seven rrr operons (Hirvonen et al., 2001). In a number of bacterial species, the rrr operons contain either functional UP elements (Aiyar et al., 2002; Ross et al., 1993; Wada et al., 2000) or phased AT-rich tracts, i.e. putative UP elements upstream of their rrr promoters (Amador et al., 1999; Dryden & Kaplan, 1993; Garnier et al., 1991; Helmann, 1995; La Fontaine & Rood, 1996; Zahn et al., 2001). Mycobacteria appear to rely exclusively on factor-dependent activation to achieve maximal promoter activity, and this is the first characterization of a UAR that differs significantly from the existing paradigm.

ACKNOWLEDGEMENTS

We thank Elaine Davis for pJ414 and Roger Buxton for critical reading of the manuscript. This work was supported by the Medical Research Council and by The European Community for research, technological and demonstration activities, Fifth Framework Programme (Contract EU-Cluster QLK2-2000-01761).

REFERENCES


