Identification and characterization of *rel* promoter element of *Mycobacterium tuberculosis*

Vikas Jain, Subbanna Sujatha, Anil Kumar Ojha, Dipankar Chatterji*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Abstract

The *rel* gene is responsible for the maintenance of the level of (p)ppGpp in bacteria under nutrient starvation. This phenomenon known as stringent response plays an important role during survival of the microorganisms in stationary phase. We have cloned 1.6 kb upstream sequence of *rel* gene of *Mycobacterium tuberculosis* in a shuttle vector pSD5B containing promoterless *lacZ* gene and promoter activity was observed in *Mycobacterium smegmatis* cells by blue/white selection and was measured by β-galactosidase assay. In order to delineate the minimal promoter element of *rel* gene, a 200 bp fragment from this 1.6 kb upstream sequence was further cloned in promoterless *lacZ* shuttle vector pSD5B and promoter activity was observed in *M. smegmatis* cells in similar way. The 200 bp promoter fragment was found to be mycobacterium specific and did not respond when transformed in *Escherichia coli.* The +1 transcription start site was determined by primer extension method. The −10 promoter region was identified to be TATCCT. The three T bases when mutated, showed a remarkable decrease in the *lacZ* expression thus confirming the −10 region. The translation start site has also been identified by site directed frame shift mutagenesis. It appears that this *rel* promoter can be used for expression of proteins in mycobacteria.

Keywords: rel promoter; 200 bp; *lacZ* expression; Constitutive; Reporter assay

1. Introduction

Studies on the regulation of gene expression in any system are facilitated by simple and reliable assays, which can be quantitated and monitored both in vitro and in vivo. Reporter technology thus relies on fusing an assayable expression in both homologous and heterologus system, whose products are stable, with a promoter having sequence that can be regulated by different signals. Reporter genes have become convenient tools for studying mycobacteria and several such systems are known in the literature (Jain et al., 1997). Out of the many, few have become very popular and are widely used because of their control and inducibility (Parish et al., 1997; Stover et al., 1991).

By far the best candidate for reporter assay in *Escherichia coli* has been the *lacZ* expression system where the *E. coli lacZ* gene encoding β-galactosidase (Fowler and Zabin, 1983) has been extensively used with various substrates like lactose or its derivatives to catalyze the cleavage of β-1,4 linkage producing galactose and glucose as products. One of the common derivatives of lactose has been *o*-nitrophenyl-β-D-galactopyranoside (ONPG), which yields a yellow color product that can be monitored spectrophotometrically at 420 nm (Miller, 1972). In addition, the presence of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-D-galactopyranoside (X-Gal) in nutrient agar plates results in blue colored colonies because of the expression of *lacZ* thus

Abbreviations: Phsp60, promoter of hsp60 gene; Prelmt, promoter of *rel* gene of *M. tuberculosis*; OD, optical density; PCR, polymerase chain reaction; cDNA, DNA complementary to RNA.

* Corresponding author. Tel.: +91 80 2293 2836; fax: +91 80 2360 0535.
E-mail address: dipankar@mbu.iisc.ernet.in (D. Chatterji).

mark the presence of it on solid media as opposed to ONPG assay in an aqueous environment (Bannantine et al., 1997; Jain et al., 1997; Timm et al., 1994a,b). Varying degree of "blueness" in a colony, in principle, can tell the relative strength of a promoter.

Several attempts have been made in the past to fuse a mycobacterial promoter sequence with *lacZ* with varying degree of success (Dellagostin et al., 1995; Knipfer et al., 1998; Kumar et al., 1998). One of the problems was the instability of *lacZ* in *Mycobacterium smegmatis* due to transposition of an element IS 1096 and subsequent deletion of the vector (Chawla and Das Gupta, 1999; Cirillo et al., 1991).

We have been working on the carbon starvation induced stringent response pathway in *M. smegmatis* (Ojha et al., 2000, 2002; Chatterji and Ojha, 2001). The product of stringent response (p)ppGpp is maintained within the cell by two enzymes RelA and SpoT in Gram negative bacteria and in gram positive organisms like mycobacteria, both the enzymes are part of same gene known as *rel* (Mechold et al., 2002; Ojha et al., 2000). We report here the identification of the promoter of *rel* gene which has been cloned upstream to *lacZ* and is found to be specific for gene expression in mycobacteria where as β-galactosidase activity was not detectable in *E. coli* under the influence of same promoter. Thus this system will find wide range of application as a specific mycobacterial expression system. We have identified the +1 transcription start site by primer extension method and the −10 region by point mutations. We have also found the translation start site by frame shift mutagenesis. It was also observed that the plasmid bearing *lacZ* fused with 200 bp *rel* gene upstream fragment containing *rel* promoter is stable in *M. smegmatis*.

2. Materials and methods

2.1. Bacterial strains, medium and growth condition

All the plasmids used in this study are enlisted in Table 1. *M. smegmatis*, mc^{2}155 (Snapper et al., 1990) was used in all experiments. The bacteria were grown in 7H9 medium supplemented with 2% glucose, 0.05% Tween-80 and 25 μg/ml kanamycin, unless mentioned otherwise. For plate culture, 1.5% agar was added to the liquid medium. For plate assay of *lacZ*, bacteria were grown on 7H9 plate containing 40 μg/ml of X-gal. The *E. coli* strains were maintained in LB or LB agar with either 50 μg/ml of kanamycin or 100 μg/ml of ampicillin. When reporter activity was assayed in *M. smegmatis* alone, MB7H9 medium was used. However, *E. coli* does not grow in this medium, so when a comparison of promoter activity was made between *M. smegmatis* and *E. coli*, LB medium was used and *M. smegmatis* grows well in this medium.

Table 1
Plasmids used in the present study

Plasmid	Size (kb)	Marker	Description
pGEMT Easy	3.0	Amp^R	pGEMTEasy vector (Promega)
pSD5B	9.5	Kan^R	Shuttle vector containing promoterless *lacZ* gene
pVJP16	11.1	Kan^R	pSD5B containing 1.6 kb *rel* upstream region
pVJP13	10.8	Kan^R	pSD5B containing 1.3 kb rel upstream region and lacking 200 bp promoter region
pSAK12	3.2	Amp^R	pGEMTEasy vector with 200 bp DNA fragment, upstream to start codon of *M. tuberculosis relA/spoT*
pAN12	9.7	Kan^R	pSD5B with 200 bp DNA fragment, upstream to start codon of *M. tuberculosis relA/spoT*, cloned upstream of *lacZ* gene
pSS12	9.7	Kan^R	1st 'T' of −10 region of promoter mutated to 'G' in pAN12
pSS22	9.7	Kan^R	2nd 'T' of −10 region of promoter mutated to 'G' in pAN12
pSS32	9.7	Kan^R	3rd 'T ' of −10 region of promoter mutated to 'C' in pAN12
pMV261	4.5	Kan^R	Shuttle vector containing hsp60 promoter
pHsplac	7.5	Kan^R	pMV261 in which *lacZ* cloned downstream to hsp60 promoter
pR300lac	7.8	Kan^R	hsp60 promoter in pMV261 replaced with 300 bp fragment of *rel* and *lacZ* cloned downstream to it to make a translational fusion construct

2.2. Transcriptional fusion of Mycobacterium tuberculosis relA/spoT upstream fragment to lacZ reporter and activity assay

The 1.6 kb fragment that contained upstream as well as some portion of the *rel* gene of *M. tuberculosis* was PCR amplified using a set of two primers Relprof (CGGGATCTAGAAGCTGATCTTCGCACC) and RelproR1 (ACGCGCGCATGCTGG TCTTAAGAGTCTCG) (Fig. 1) from cosmid MTCY227 (a gift from S.T. Cole, Cole et al., 1998), digested with XbaI and SphI and was then cloned in pSD5B (a mycobacteria-*E. coli* shuttle vector with promoterless *lacZ*, Jain et al., 1997) vector previously digested with same enzymes. The resulting recombinant plasmid, pVJP16, has the *lacZ* reporter gene transcriptionally fused to the 124th nucleotide of *relA/spoT* gene.

M. smegmatis, mc^{2}155, transformed with pVJP16 was cultured till mid-log phase (OD_{600}=0.7) in 7H9 medium (supplemented with 2% glucose, 0.05% Tween-80 and 25 μg/ml kanamycin), harvested, washed once with PBS and transferred to 7H9 medium containing either 2% or 0.02% glucose and assayed for β-galactosidase activity in liquid culture using *o*-nitrophenyl-β-D-galactopyranoside (ONPG)

```
BamHI                                                                 Relprof →
GGATCCCTGG TAACCCCCGA CAGCTACCAG CGCACCGACT ACCCGTCGGC CGGGATCGAG
CAGCTGATCT TCGCACCACA AGGTTCACTC GCGCAAAGCC GCACCCGCCG CGCGCTCGCG
TTGTGTGTAC CCCGGGACGC GATCGCTCGG GATGCCGGGG TTCCGATTGC CAACTCGCGG
CTGTCCCCGG CGACCGACGA TGCCCTCACC GATGCCGACG GCGCCGCCGA AGCACGTCAG
TTCGGCCGGG TGGACCCCGC CGCCGCTCGC GACGCGCTGG GTGGTACGCC GCTGACCGTG
CGGATCGGCT ACGGCAGGCC CAACGCTCGG TTGGCGGCCA CCATCGGAAC CATTGCCGAC
GCCTGCGCCC CGGCCGGGAT CACCGTTTCG GATGTGACGG TGGACACACC CGGACCGCAA
GCGCTGCGGG ACGGAAAGAT TGACGTATTG TTGGCGAGCA CCGGTGGGGC CACCGGCAGC
GGATCGAGCG GATCGTGTGC GATGGATGCC TATGACTTGC ACAGCGGCAA CGGAAACAAT
CTATCGGGGT ACGCAAACGC TCAGATCGAC GGCATCATCA GCGCGCTCGC GGTGTCGGCC
GACCCCGCCG AGCGGGCCAG GTTGCTTGCC GAGGCCGCGC CGGTGCTCTG GGATGAGATG
CCAACCTTGC CGTTGTACCG GCAGCAGCGC ACGTTGTTGA TGTCGACGAA AATGTATGCG
GTGAGCAGGA ATCCGACGCG ATGGGGGGCA GGGTGGAACA TGGATCGCTG GGCGCTGGCG
CGGTGACGAT GGCCAGTGCC ATCTGCAGGT AATTGACAGA ATTCCACGAC GAGAAGCGGA
                                                        apt →
CTATCGGAGC GTAGTGTCGC AGGTGCTCCG GGCTGTCTGG GAGAGGATGT GTGCCATGGC
GGTACATGGG CTGGTGACTA CGTGTTGAAC GTGATCGCGA CGGGGCTCTC CTTAAAGGCA
CGGGGGAAGC GCCGCCGGCA GCGTTGGGTC GACGACGGGC GGGTATTGGC GCTCGGTGAG
TCCCGCCGGA GCTCAGCCAT ATCTGTGGCC GACGTGGTTG CGTCGCTGAC CCGGGATGTG
GCCGACTTTC CGGTTCCCGG CGTCGAGTTC AAGGACCTCA CCCCGCTATT CGCCGACCGA
AGAGGATTGG CCGCGGTAAC CGAAGCGCTG GCCGATCGGG CGTCCGGAGC TGACCTGGTG
GCCGGCGTCG ACGCCCGCGG GTTTCTGGTG GCAGCCGCGG TCGCCACCCG GCTCGAAGTG
GGTGTGCTGG CCGTTCGCAA GGGCGGCAAG CTGCCCCGGC CGGTGCTCAG CGAGGAGTAC
TACAGGGCGT ACGGCGCCGC CACTCTGGAG ATTCTCGCTG AGGGCATCGA GGTTGCGGGC

   ← RelproR2      Sak2 →
CGCCGTGTCG TGATCATTGA CGACGTGTTA GCAACCGGCG GCACCATCGG CGCGACGCGA
CGCCTGCTTG AGCGCGGTGG CGCCAACGTG GCCGGGGCGG CCGTAGTGGT GGAACTTGCG

                               Upstream ORF (apt) stop codon ←
GGGTTGAGCG GTCGCGCGGC GCTCGCACCG CTGCCGGTGC ACAGCCTGAG CCGCCTGTGA

              ← Sak1          relA/spoT start codon →
GGGATATCCT CTAGGTCGGA GGTGACGAAC GTGGCCGAGG ACCAGCTCAC GGCGCAAGCG
GTTGCACCGC CCACGGAGGC TTCTGCGGCT CTCGAGCCCG CTCTCGAGAC GCCCGAGTCG

                 ← RelproR1
CCGGTCGAGA CTCTTAAGAC CAGCATCAGC GCGTCGCGTC GGGTGCGGGC CCGATTGGCC
CGGCGGATGA CCGCCCAGCG CAGCACCACC AATCCGGTGC TCGAGCCGTT GGTGGCGGTG
CACCGGGAGA TCTATCCCAA GGCCGACCTG TCGATCTTGC AGCGAGCCTA CGAGGTCGCT
        BglII

GACCAAAGGC ATGC
        SphI
```

Fig. 1. The nucleotide sequence of 1.6 kb DNA fragment of *M. tuberculosis relA/spoT* gene locus. The primers used for amplifying 1.6 kb, 1.3 kb and 200 bp regions and the putative start codon for the Rel protein have been shown. The upstream ORF *apt* start and stop codons have also been mentioned (Cole et al., 1998).

exactly as described (Miller, 1972) at different time intervals. The activity is represented in terms of Miller units that is calculated using the formula,

$$\text{Activity (Miller Units)} = 1000 \times \frac{A_{420} - 1.75(\text{OD}_{550})}{\text{Time} \times \text{Vol}_{\text{culture}} \times \text{OD}_{600}}.$$

In liquid culture at least three readings from three different cultures were taken. *M. smegmatis* transformed with pSD5B was used as negative controls.

2.3. Cloning and characterization of 200 bp upstream sequence proximal to the start codon of relA/spoT

A set of two primers sak1 (CGGCCACGTTCGG-TACCTCCGACCTAGA) and sak2 (GCCGTGTCGTGA-GAATTCACGACGTGTTAG) were used to amplify the 200 bp immediately upstream to *relA/spoT* (see Fig. 1) from pVJP16. The 200 bp amplicon was subcloned into pGEMT Easy vector (Promega) according to manufacturer's instruction to form pSAK12. The clone with the correct orientation (the end proximal to the gene was towards SphI site) was picked and the 200 bp insert was released by SphI-SpeI and ligated to SphI-XbaI ends of pSD5B to form a recombinant plasmid pAN12. The promoter activity of the 200 bp fragment was analyzed by assaying the *lacZ* activity in *M. smegmatis* as well as in *E. coli* transformed with pAN12 in LB medium. The *lacZ* activity was assayed on plate as well as in liquid culture using *o*-nitrophenyl-β-D-galactopyranoside (ONPG) exactly as described (Miller, 1972). In liquid culture at least three readings from three different cultures were taken. *E. coli* and *M. smegmatis* transformed with pSD5B were used as negative controls.

The stability of pAN12 in the host strain, both *M. smegmatis* and *E. coli* was further assessed by repeated subculturing for 10 generations, expressing *lacZ* gene on X-gal containing plate.

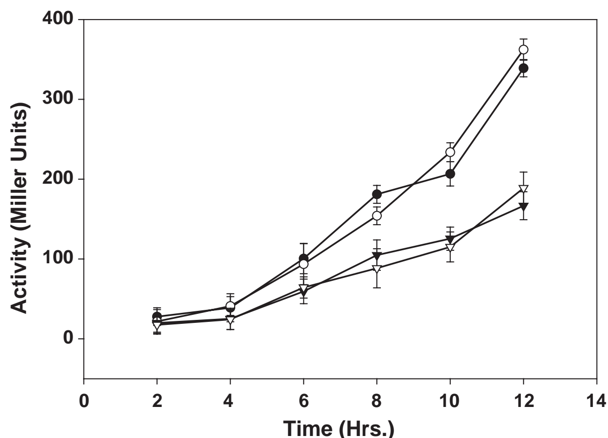


Fig. 2. Comparison of the activities of pVJP16 (represented by circles) and pVJP13 (represented by inverted triangles) vectors when assayed using *lacZ* reporter system. The time in hours represent the time after inoculating mid-log phase (OD=0.7) bacteria into the fresh medium containing 2.0% (-●-, -▼-) or 0.02% (-○-, -▽-) glucose.

2.4. Transcriptional fusion of M. tuberculosis relA/spoT upstream fragment lacking 200 bp region to lacZ reporter and activity assay

The 1.3 kb *rel* gene upstream fragment that lacks the 200 bp region was PCR amplified using a set of two primers Relprof (CGGGATCTAGAAGCTGATCTTCGCACC) and RelproR2 (TAACACGTCGTGCATGCTCACGACACGG) (Fig. 1) from cosmid MTCY227 (a gift from S.T. Cole, Cole et al., 1998), and was cloned in similar way as mentioned above in pSD5B shuttle vector with promoterless *lacZ* giving rise to pVJP13.

β-galactosidase activity in *M. smegmatis*, mc^2155, transformed with pVJP13 was assayed in the same way as mentioned above.

2.5. Transcription start site mapping

The +1 transcription start site was identified using primer extension method as described (Sambrook et al., 1989). Total RNA was isolated using RNeasy midi kit (Qiagen) from *M. smegmatis* transformed with pAN12 and grown till $OD_{600}=0.8$. A total of 10 µg RNA was used to make cDNA using map2 (GGAAGTGATTCCTCCGGA-TAT CG) primer end labeled with $\gamma P^{32}ATP$ (Perkin Elmer) and RevertAid M-MuLV reverse transcriptase (Fermentas) following manufacturer's instructions. The primer was designed approximately 83 nucleotides downstream to −10 region. Sequencing reaction was run using fmol DNA Cycle sequencing system (Promega) with end labeled map2 primer and the template pAN12 in accordance with the manufacturer's instructions except that the annealing temperature was 50 °C. The sequencing product was separated on 10% denaturing polyacrylamide gel containing 7 M urea. The gel was dried and phosphor imaged (Fujifilm FLA2000).

2.6. Mutation of the promoter element: identification of −10 region

Site-specific mutagenesis was carried out by the Quick-change protocol (Stratagene) in the −10 region of the promoter (TATCCT). The three highly conserved T bases in the −10 region of the promoter were mutated to either G or C bases (see Fig. 5a). The PCR conditions were 94 °C for 3 min, 65 °C for 30 s and 72 °C for 3 min (for 30 cycles), using pSAK12 as template. The mutations were confirmed by sequencing of DNA. Both wild type and mutants 200 bp inserts were released by SphI-SpeI digestion of pSAK12 and ligated to SphI-XbaI ends of pSD5B to form pSS12, pSS22, pSS32 (see Fig. 5a). Both plate as well as liquid culture assays were done to assess the activity of β-galactosidase.

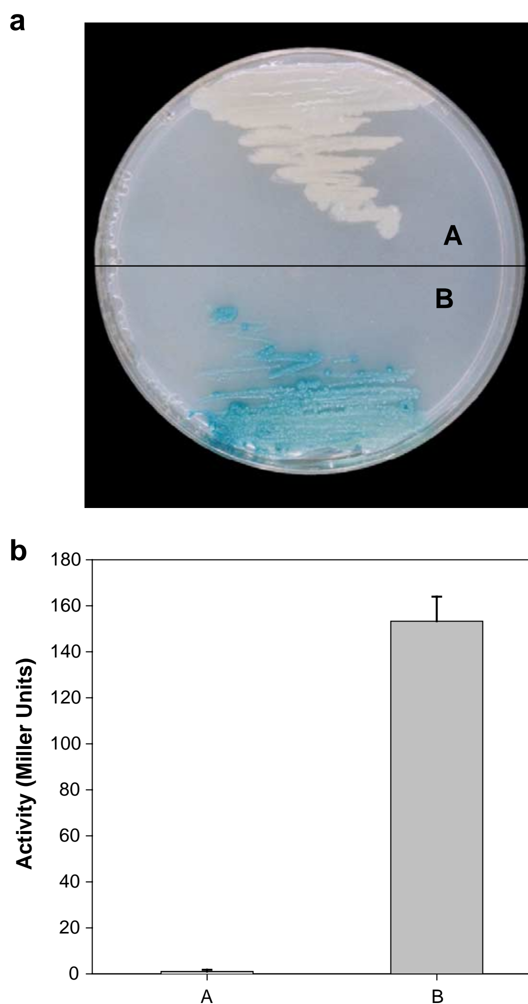


Fig. 3. (a) Promoter activity of 200 bp fragment immediately upstream to the start codon of *M. tuberculosis relA/spoT*. pSD5B (promoterless *lacZ*) (A) and pAN12 (200 bp DNA fragment+*lacZ*) (B) were transformed in *M. smegmatis* and the colonies were streaked on 7H9 agar containing X-gal. (b) Quantitative analysis of the promoter activity in A, pSD5B and B, pAN12. The cells were grown in 2.0% glucose concentration in 7H9 broth till mid-log phase before harvesting. The assays were done in triplicates and data represents the average.

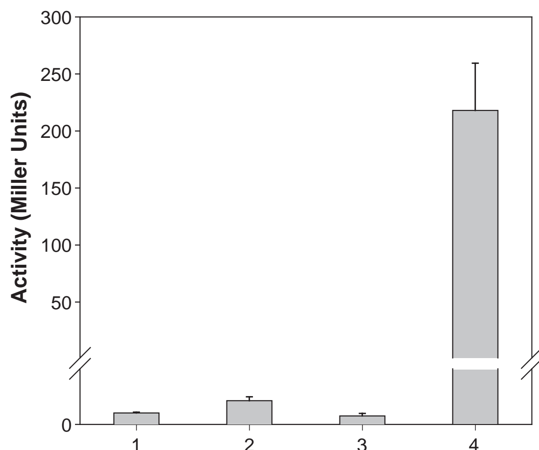


Fig. 4. Estimation of the *lacZ* expression from pSD5B in *E. coli* (1) and *M. smegmatis* (2). Similarly expression from pAN12 in *E. coli* (3) and *M. smegmatis* (4) shows that 200 bp promoter region is functional in *M. smegmatis*, but not in *E. coli*. The *Y*-axis is broken to show the level of activity from 1, 2 and 3. Both *M. smegmatis* and *E. coli* were grown in LB broth containing 25 μg/ml kanamycin as mentioned in Material and methods. The assays were done in triplicates and the data represents the average of the three.

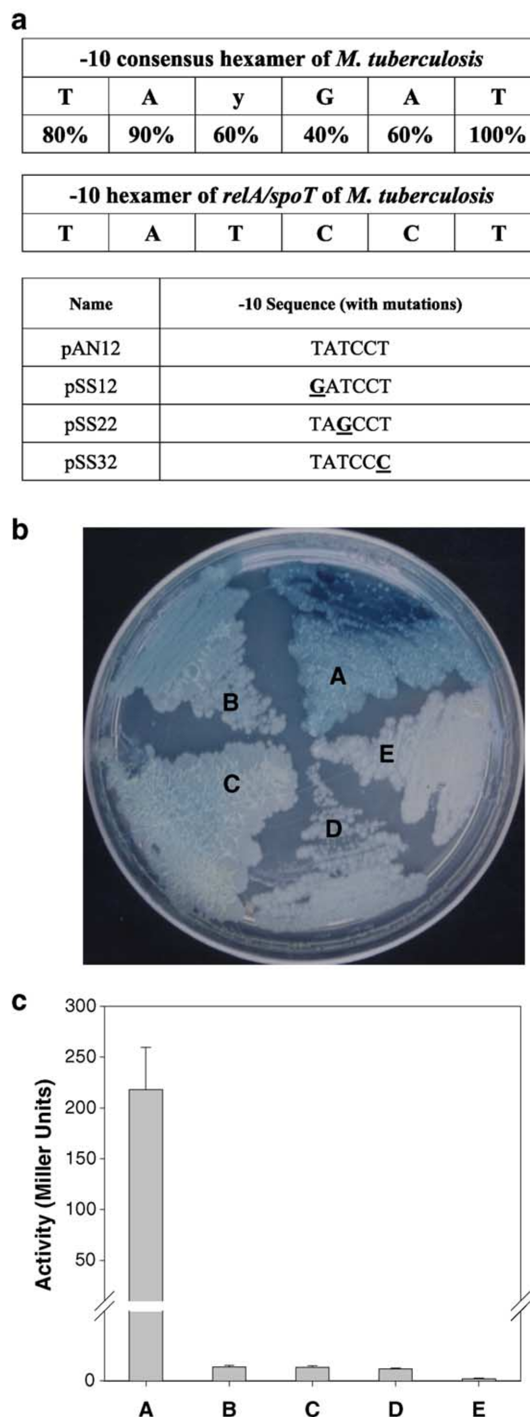
2.7. Identification of translation start site

A translational fusion construct of promoter region and *lacZ* gene was made and translation start site was identified by frame shift mutagenesis as described previously except that instead of mutating start codon (Feltens et al., 2003), we have done frame shift mutations. Approximately 300 bp region that includes 103 bp upstream and 180 bp downstream to −10 region was PCR amplified using primers relup1f (CGGCCGTAGTG GTACCACTTGCGGG) and relup1r (GGTGGTG-CTGCAGTGGGC GGTCATCC) and cloned in pMV261 vector (Stover et al., 1991) in place of Phsp60 promoter at KpnI and PstI sites to give rise to pR300. *lacZ* was taken out from pSD5B using PstI and was cloned in pR300 at PstI site thus resulting in pR300lac. The clone was checked for correct orientation and transformed in *M. smegmatis* and β-galactosidase assays were carried out. At least three readings from three separate cultures were taken in each case.

Three frame shift mutations were made using 3 sets of primers and pR300lac template. PCR conditions were 96 °C for 1 min, 50 °C for 1 min and 68 °C for 14 min (for 15 cycles). All the mutant plasmids were transformed in *M. smegmatis* and the *lacZ* assays were done to measure the activity of β-galactosidase.

2.8. Miscellaneous

The strategy of molecular cloning was followed as described (Sambrook et al., 1989). The electroporation of *M. smegmatis* was carried out in cell electroporator (BTX) with 2 mm-gap cuvette at 1.25 kV/mm.


a

-10 consensus hexamer of *M. tuberculosis*					
T	A	y	G	A	T
80%	90%	60%	40%	60%	100%

-10 hexamer of *relA/spoT* of *M. tuberculosis*					
T	A	T	C	C	T

Name	-10 Sequence (with mutations)
pAN12	TATCCT
pSS12	**G**ATCCT
pSS22	TA**G**CCT
pSS32	TATCC**C**

Fig. 5. (a) −10 hexamer identified in *M. tuberculosis relA/spoT* gene and mutations that were made in the three conserved T bases to study promoter activity. (b) Effect of three mutations in the −10 region on *lacZ* expression in *M. smegmatis* transformed with (A) pAN12, (B) pSS12, (C) pSS22, (D) pSS32 and (E) pSD5B and grown on 7H9 agar containing X-gal containing 25 μg/ml kanamycin. (c) Measurement of β-galactosidase activity in *M. smegmatis* transformed with (A) pAN12, (B) pSS12, (C) pSS22, (D) pSS32 and (E) pSD5B. The cells were grown till mid-log phase before harvesting. The assays were done in triplicates and data here represents the average. The *Y*-axis is broken to show the level of activity in B, C, D and E.

3. Results

3.1. The 1.6 kb DNA fragment has a constitutive promoter activity

The 1.6 kb DNA fragment (Fig. 1) showed promoter activity when cloned in front of *lacZ* gene in pSD5B vector (Jain et al., 1997). An ONPG assay was done using this construct to obtain a quantitative data as a function of time of growth. As the conversion of ONPG to ONP (*o*-nitrophenol) by β-galactosidase can be monitored spectrophotometrically, it gives an estimation of the amount of expression of *lacZ* and hence the promoter activity (Miller, 1972). It was observed that the expression from the 1.6 kb fragment was constitutive in nature and the activity changed negligibly when the cells were shifted to carbon starved medium (0.02% glucose) (Fig. 2). All the experiments were carried out in triplicates. As 1.6 kb fragment was long for promoter analysis, a search for promoter element nearest to *relA/spoT* gene was carried out using nested PCR.

3.2. The promoter activity was contained in a 200 bp sequence immediately upstream to relA/spoT gene

With a set of two primers, sak1 and sak2, a 200 bp DNA fragment (Fig. 1) upstream to *relA/spoT* was amplified and cloned ahead of *lacZ* in pSD5B vector (Jain et al., 1997) to form promoter–reporter construct, pAN12 which was transformed in *M. smegmatis* for β-galactosidase activity assays. Fig. 3a shows that 200 bp fragment was sufficient to show promoter activity. Quantitative analysis of the promoter-*lacZ* system in liquid culture (Fig. 3b) corroborated the data obtained with plate culture. All the experiments were carried out in triplicates. Consistent with the promoter activity of 1.6 kb with *lacZ* reporter, the activity of 200 bp was observed to be constitutive with a high level of expression even under nutrient enriched condition. In carbon starved condition, no additional increase in β-galactosidase activity was noticed (data not shown). Although a set of nested PCR products with increment of 200 bp were also amplified, they were not analyzed further since the entire promoter activity was observed in the proximal 200 bp fragment. To further confirm this, one construct pVJP13 was prepared by cloning the 1.3 kb upstream sequence of *rel* gene lacking the 200 bp region (which is having promoter activity) (Fig. 1) and transformed in *M. smegmatis* cells and assayed for β-galactosidase activity. It was observed that *M. smegmatis* transformed with pVJP13 construct showed lesser activity as compared with the *M. smegmatis* transformed with pVJP16 (Fig. 2) which clearly demonstrates the contribution from 200 bp region. However, there remained some residual activity from the pVJP13 construct alone which was probably because of the promoter present upstream to

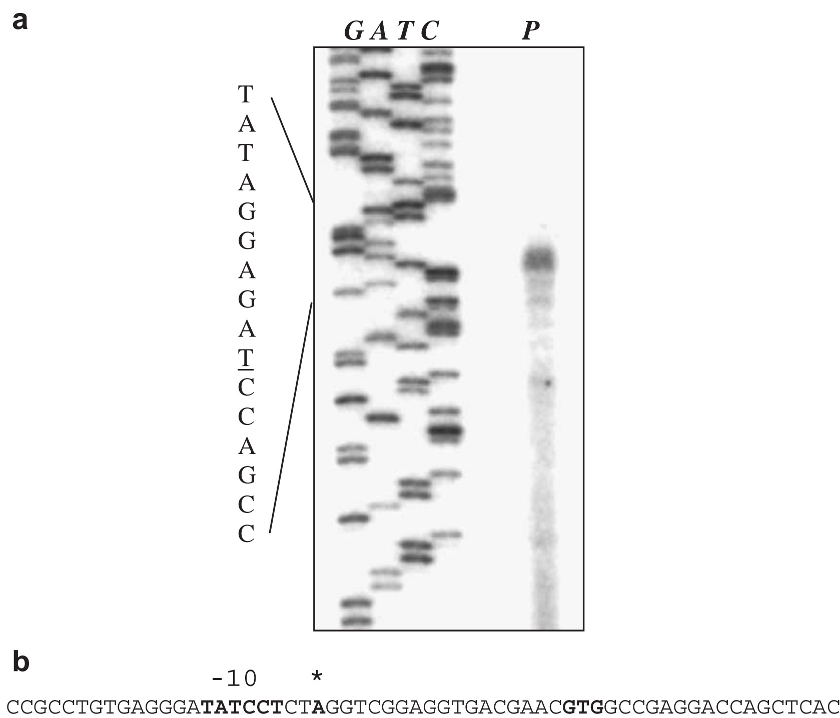


Fig. 6. Identification of the transcription initiation site. (a) Primer extension analysis. The transcription start site was determined by primer extension method using total RNA from *M. smegmatis* cells transformed with pAN12 and grown till OD=0.7. The reaction product was run in lane P alongside the sequencing reaction represented by G, A, T and C. The sequence on the template strand has been shown on the left and the identified +1 transcription start site has been underlined. (b) DNA sequence of the *relA/spoT* promoter region showing the +1 base (bold and underlined marked with an asterisk). −10 region and the translation start codon have also been underlined.

the *apt* gene as shown in Fig. 1. For further work we have referred the pSD5B vector containing 200 bp fragment as Prelmt. Though we do not observe a regulated expression from this promoter in *M. smegmatis*, we cannot rule out the possibility of *rel* promoter being regulated in *M. tuberculosis* which is probably due to the differences in the transcriptional machinery (Triccas et al., 2001).

Interestingly, the promoter activity of Prelmt was specific to mycobacteria and was completely absent in *E. coli* (Fig. 4). When *E. coli* and *M. smegmatis* were transformed with pAN12, specific β-galactosidase activity was noticed in the case of *M. smegmatis* only. A quantitative analysis by an ONPG assay of the same (Fig. 4) shows a remarkable difference between the two thus confirming that Prelmt is functional only in mycobacteria. This observation was consistent with the general property of most of the *M. tuberculosis* promoters that they are not active in *E. coli* (Jain et al., 1997; Mulder et al., 1997). Here both *E. coli* and *M. smegmatis* transformed with pSD5B acted as negative control.

3.3. Detection of promoter element

Putative mycobacterial promoter sequences, as published by Mulder et al., 1997 show that *M. tuberculosis* promoter consists of a −10 consensus sequence TAyGAT (y-pyrimidine) (Fig. 5a). Putative −10 region with the sequence TATCCT was identified in the 200 bp promoter region of *rel*. This −10 hexamer of *M. tuberculosis* is highly conserved at four positions. Fig. 5a shows varying degree of conserved T base in the −10 region sequence. Thus after comparing the −10 consensus of *M. tuberculosis* with *rel* gene −10 hexamer, we mutated the 1st position T to G, 3rd position T to G and 6th position T to C and studied their effect on *lacZ* expression. It was found that all these mutations severely affected the expression of *lacZ* as observed on X-gal plate (Fig. 5b). The ONPG assay data on these mutants also represent the same (Fig. 5c) thus confirming that the TATCCT is the actual −10 region which is located 3 bases upstream to +1 transcription start site.

Fig. 7. Identification of the translation initiation site. (a) DNA sequence of the *relA/spoT* N-terminal region along with the amino acid sequence of the wild type. Three putative start codons have been boldfaced and underlined. The three frame shift mutations have also been marked with the arrow and the inserted base has been mentioned. (b) Quantitative analysis of the wild type (1) and its comparison with mutants (2, 3 and 4) clearly shows that the first GTG is the translation initiation site. The assays were done in triplicates and data presented here is the average of the three. Bar 3 was found to be undetectable.

3.4. Identification of transcription initiation and translation initiation site

Primer extension method was used to identify +1 transcription start site in *rel* promoter. In general, +1 transcription start site is approximately 10 bases downstream to −10 hexamer but here we observed that the third base to −10 hexamer is the transcription initiation site (Fig. 6). It can also be seen that there are few other bands below the actual product. Since these products are diffused and of low intensity, we do not consider them to be the actual products as they can appear because of the tendency of reverse transcriptase to stop or pause in regions of high secondary structure in the template RNA (Harrison et al., 1998).

The translation initiation site was identified by making a translational fusion construct and doing frame shift mutations within it. As there were at least 3 codons which are in frame (Fig. 7a) and therefore could act as potential sites for translation initiation, approximately 300 bp fragment that includes all three putative initiation codons in the *rel* gene was taken and a translation fusion construct with promoterless *lacZ* was made. It was earlier thought that the translation will start from some downstream start codon as the number of bases between −10 hexamer and the first putative initiation codon was very less (20 nucleotides) (Fig. 7a). Therefore a hunt for the translation start site was initiated by making use of frame shift mutagenesis. Three mutations were made and the corresponding β-galactosidase activity was measured in *M. smegmatis* (Fig. 7a). It was found that shifting the frame after the first putative start codon itself (assuming that translation starts with this codon), the β-galactosidase activity was almost completely lost (Fig. 7b). Since such frame shift mutations will affect the translation, we conclude from the present data that the first codon is the actual initiation codon.

4. Discussion

The development of molecular genetic tools is needed to understand the mechanisms regulating gene expression in mycobacterial species. The lesser occurrence of strong promoters in mycobacterial genome can be one of the reasons why a sufficiently strong expression system has not yet been established for mycobacteria. Such an expression system can be achieved by providing a strong mycobacterial promoter upstream to the desired gene. With this vector, the gene of interest, from a slow growing pathogen, can be successfully expressed in the heterologus faster growing mycobacterial species such as *M. smegmatis*, which can act as a surrogate host.

Here, we show that the 200 bp *rel* promoter region obtained from 1.6 kb upstream fragment of *rel* gene of *M. tuberculosis* is sufficient for promoter activity and is constitutive in nature at least in *M. smegmatis*. The proximity between the promoter and the initiation site (2 bp) is noteworthy, although similar unusual observations were made earlier too in different systems. For eg., *purC* gene of *M. tuberculosis* (Jackson et al., 1996) and *bla* gene of *Mycobacterium fortuitum* (Timm et al., 1994b) have one single base representing both transcription start site as well as first base of the translation start codon; the distance of 4 bases between promoter and transcription start site has also been noticed before (Bashyam et al., 1996).

Such simple blue/white selection and a β-galactosidase assay would go a long way for both quantitative and qualitative assessment of the mycobacterial promoter strength. In addition, any gene cloned downstream of *rel* promoter in correct orientation would show good expression, expectedly. The stability of the plasmid for a considerable length of time is an added advantage. Although we expected a regulatory, starvation controlled promoter element of the *rel* gene, even the 1.6 kb upstream fragment showed constitutive expression with *lacZ* gene (Fig. 2). This can probably be attributed to the fact that the expression from this promoter is being monitored in *M. smegmatis* which is a heterologous system and may have differences in the transcriptional machinery (Triccas et al., 2001). We thus do not rule out the possibility that the *rel* promoter can be under regulation in *M. tuberculosis*. Also there could be other regulatory elements that cannot be detected by the assays presented here. We expect this system would find a wide range of application.

Acknowledgements

The authors wish to thank Council for Scientific and Industrial Research, India for financial support (NMITLI). V.J. is thankful to the CSIR for the award of senior research fellowship. Initial part of this work was carried out by Saaket Verma.

References

Bannantine, J.P., Barletta, R.G., Thoen, C.O., Andrews Jr., R.E., 1997. Identification of *Mycobacterium paratuberculosis* gene expression signals. Microbiology 143, 921–928.

Bashyam, M.D., Kaushal, D., Das Gupta, S.K., Tyagi, A.K., 1996. A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. J. Bacteriol. 178, 4847–4853.

Chatterji, D., Ojha, A.K., 2001. Revisiting the stringent response, ppGpp and starvation signaling. Curr. Opin. Microbiol. 4, 160–165.

Chawla, M., Das Gupta, S.K., 1999. Transposition-induced structural instability of *Escherichia coli*-mycobacteria shuttle vectors. Plasmid 41, 135–140.

Cirillo, J.D., Barletta, R.G., Bloom, B.R., Jacobs Jr., W.R., 1991. A novel transposon trap for mycobacteria: isolation and characterization of IS1096. J. Bacteriol. 173, 7772–7780.

Cole, S.T., et al., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393, 537–544.

Dellagostin, O.A., Esposito, G., Eales, L.J., Dale, J.W., McFadden, J., 1995. Activity of mycobacterial promoters during intracellular and extracellular growth. Microbiology 141, 1785–1792.

Feltens, R., Feltens, R., Gößringer, M., Willkomm, D.K., Urlaub, H., Hartmann, R.K., 2003. An unusual mechanism of bacterial gene expression revealed for the Rnase P protein of *Thermus* strains. Proc. Natl. Acad. Sci. 100, 5724–5729.

Fowler, A.V., Zabin, I., 1983. Purification, structure and properties of hybrid beta-galactosidase proteins. J. Biol. Chem. 258, 14354–14358.

Harrison, G.P., Mayo, M.S., Hunter, E., Lever, A.M., 1998. Pausing of reverse transcriptase on retroviral RNA templates is influenced by secondary structures both 5′ and 3′ of the catalytic site. Nucleic Acids Res. 26, 3433–3442.

Jackson, M., et al., 1996. The *Mycobacterium tuberculosis* purine biosynthetic pathway: isolation and characterization of the *pur*C and *pur*L genes. Microbiology 142, 2439–2447.

Jain, S., Kaushal, D., DasGupta, S.K., Tyagi, A.K., 1997. Construction of shuttle vectors for genetic manipulation and molecular analysis of mycobacteria. Gene 190, 37–44.

Knipfer, N., Nooruddin, L., Shrader, T.E., 1998. Development of an α-complementation system for mycobacterial promoter analysis. Gene 217, 69–75.

Kumar, D., Srivastava, R., Srivastava, B.S., 1998. Genetic rearrangements leading to disruption of heterologus gene expression in mycobacteria: an observation with *E. coli* β-galactosidase in *Mycobacterium smegmatis* and its implication in vaccine development. Vaccine 16, 1212–1215.

Mechold, U., Murphy, H., Brown, L., Cashel, M., 2002. Intramolecular regulation of the opposing (p)ppGpp catalytic activities of Relseq, the Rel/Spo enzyme from *Streptococcus equisimilis*. J. Bacteriol. 184, 2878–2888.

Miller, J.H., 1972. Experiments in Molecular Genetics. Cold Spring Harbor Press, Cold Spring Harbor, NY, USA.

Mulder, M.A., Zappe, H., Steyn, L.M., 1997. Mycobacterial promoters. Tuber. Lung Dis. 78, 211–223.

Ojha, A.K., Mukherjee, T.K., Chatterji, D., 2000. High intracellular level of guanosine tetraphosphate in *Mycobacterium smegmatis* changes the morphology of the bacterium. Infect. Immun. 68, 4084–4091.

Ojha, A.K., Verma, S., Chatterji, D., 2002. Synthesis of an unusual polar glycopeptidolipid in glucose-limited culture of *Mycobacterium smegmatis*. Microbiology 148, 3039–3048.

Parish, T., Mahenthiralingam, E., Draper, P., Davis, E.O., Colston, M.J., 1997. Regulation of the inducible acetamidase gene of *M. smegmatis*. Microbiology 143, 2267–2276.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., Jacobs Jr., W.R., 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. 4, 1911–1919.

Stover, C.K., et al., 1991. New use of BCG for recombinant vaccines. Nature 351, 456–460.

Timm, J., Lim, E.M., Gicquel, B., 1994a. *Escherichia coli* mycobacteria shuttle vectors for operon fusion and gene fusion to lacZ: the pJEM series. J. Bacteriol. 176, 6749–6753.

Timm, J., et al., 1994b. Transcription and expression analysis, using *lac*Z and *phoA* gene fusions, of *Mycobacterium fortuitum* β-lactamase gene cloned from a natural isolate and a high level β-lactamase producer. Mol. Microbiol. 12, 491–504.

Triccas, J.A., Britton, W.J., Gicquel, B., 2001. Isolation of strong expression signals of *Mycobacterium tuberculosis*. Microbiology 147, 1253–1258.