Macrophage-specific *Mycobacterium tuberculosis* genes: identification by green fluorescent protein and kanamycin resistance selection

Vikas Srivastava,1 Carine Rouanet,2 Ranjana Srivastava,1 B. Ramalingam,2 Camille Locht2 and Brahm S. Srivastava1

Correspondence
B. S. Srivastava
dbrahm@gmail.com
C. Locht
camille.locht@pasteur-lille.fr

1Microbiology Division, Central Drug Research Institute, Lucknow, India
2INSERM U629, Institut Pasteur de Lille, 1 rue du Professor Calmette, F-59019 Lille Cedex, France

*Mycobacterium tuberculosis* survives and multiplies inside macrophages of its host by modulating the expression of several genes essential for *in vivo* survival. An *in vivo* expression system has been developed, based on green fluorescent protein and kanamycin resistance, to identify *M. tuberculosis* genes which appear to be up-regulated in infected macrophages. A promoter-trap shuttle vector, pLL192, was constructed, containing a streptomycin resistance gene as selection marker and an artificial bicistronic operon composed of the promoterless green fluorescent protein (gfp) gene, followed by the kanamycin resistance gene. A unique *Bam*HI site upstream of the gfp gene allowed for insertion of promoter libraries. The vector was validated by the use of known regulated or constitutive *M. tuberculosis* promoters. In addition, an *M. tuberculosis* genomic DNA library was inserted into pLL192 and then introduced into *Mycobacterium bovis* BCG. The recombinant BCG cells were then used to infect the J774A.1 murine macrophage-like cell line in the presence of kanamycin. Several recombinant BCG cells were thereby selected that were resistant to kanamycin within infected macrophages, but were sensitive to kanamycin when grown *in vitro*. The kanamycin resistance phenotype was paralleled by the fluorescence phenotype. After nucleotide sequencing, the corresponding genes were identified as mce1A, PE_PGRS63(RV3097c), Rv2232, Rv1026, Rv1635c, viuB, Rv2231(cobC) and Rv0997.

Real-time PCR analysis using RNA isolated at various time points from *M. tuberculosis* and *M. bovis* BCG grown *in vitro* and within macrophages, confirmed the up-regulation of these genes. The level of up-regulation varied from 2- to 40-fold in macrophages compared to growth *in vitro*.

INTRODUCTION

Tuberculosis is the second leading cause of mortality by infectious diseases in the human population, killing about 1.8 million people worldwide [www.who.int/mediacentre/factsheets/fs104/en/print.html (accessed on 6 November 2004)]. The problem has worsened with the spread of AIDS, which may lead to reactivation of dormant tubercle bacilli and make patients more prone to infection by non-tuberculous mycobacteria.

Tuberculosis is treated by a combination of four drugs which act on different targets. In recent years, *Mycobacterium tuberculosis* isolates resistant to one or more of these drugs have been spreading, which seriously hampers successful control measures (WHO, 1997). Novel drugs acting on new targets are therefore urgently needed.

There are several important macromolecules in the mycobacterial cell which are vital for survival of the organism in host cells and progression of the disease (Smith, 2003). These macromolecules may either be constitutively expressed or their expression may be up-regulated within the host. Finding such targets is the key to developing new drugs.

*M. tuberculosis* infects by aerosol and is ingested by alveolar macrophages. Contrary to most other invading microorganisms, which are killed within the macrophages, the tubercle bacillus survives and grows in the unfavourable conditions inside the macrophages. It has been reported that alterations in gene expression in the mycobacteria and in the macrophages take place, which allow the mycobacteria to...
survive (Barker et al., 1998; Dubnau et al., 2002; Hobson et al., 2002). Studies with other pathogens have shown that genes which are up-regulated during infection are often important for pathogenicity (Chiang et al., 1999). Therefore, the identification of mycobacterial genes that are up-regulated in infected macrophages may provide targets for drug design or vaccine development.

Different approaches have been used to identify microbial genes that play a role in the interaction with host cells and thus in virulence (Jacobs et al., 1991). These include signature-tagged transposon mutagenesis (Camacho et al., 1999), in vivo-induced antigen technology (Deb et al., 2002), subtractive hybridization (Plum & Clark-Curtiss, 1994), selective capture of transcribed sequences (Graham & Clark-Curtiss, 1999) and in vivo expression technology (IVET; Mahan et al., 1993). Several important genes have been identified in mycobacteria, including hspX (Yuan et al., 1998), and genes encoding isocitrate lyase (McKinney et al., 2000), mycobacterial cell entry protein (Arruda et al., 1993) and glutamine synthetase (Harth & Horwitz, 1999).

Compared to other techniques, IVET has the advantage of identifying bacterial genes in the host environment. It was originally developed to identify virulence genes of Salmonella typhimurium (Mahan et al., 1993) and this elegant approach is being applied now to other pathogenic bacteria. In the present investigation, we have used a modified IVET strategy based on a promoter-trap plasmid vector. The plasmid contains two reporter genes, gfp, encoding green fluorescent protein, and kan, encoding kanamycin resistance, in a transcriptional fusion downstream of a promoter cloning site. An M. tuberculosis promoter library was screened in macrophages on the basis of gfp expression and kanamycin resistance. Expression levels of genes belonging to the identified promoters were then individually confirmed by real-time PCR analysis.

**METHODS**

**Bacterial strains and growth conditions.** *Escherichia coli* DH5α [F′ (80dMacZ M15) recA1 endA1 gyrA96 thi-1 hsdR17 (rK mK Tc Rk] supE44 thi-1 deoR ΔlacZYA-argF-169 Inv) and NM522 [F′ proAB lacB ΔlacZ M15/supE thi Δlac-proAB Δ(lacDMS-merB)5 (rK mK] McrB; MBI Fermentas] were used for cloning of the different constructs and the library. Mycobacterium tuberculosis H37Rv, Mycobacterium smegmatis mc²155, Mycobacterium bovis BCG Pasteur strain (isolate 1173P2; WHO) were employed in this study. *E. coli* was grown in Luria–Bertani (LB) broth or plated on LB solidified with 1.5 % (w/v) agar. Mycobacteria were grown in liquid Sauton medium (Sauton, 1912), Middlebrook 7H9 medium (Difco) supplemented with ADC (0.2 % glucose, 0.5 % bovine serum albumin fraction V, 0.0003 % beef catalase) or solid Middlebrook 7H10/7H11 agar (Difco) + OADC enrichment (Becton Dickinson). All cultures were grown at 37 °C with or without shaking. Antibiotics were added into the media when required at the following concentrations: streptomycin, 10–25 µg ml⁻¹; kanamycin, 5–25 µg ml⁻¹; cycloheximide, 50 µg ml⁻¹; gentamicin, 50 µg ml⁻¹.

**Fluorescence measurement.** Fluorescence was measured by using a FACS Calibur (Becton Dickinson) and the Cell Quest software provided by the supplier. The bacterial population was gated by using side and forward scatter parameters and then analysed for fluorescence. The results are expressed as the percentage population fluorescing and the mean fluorescence intensity.

**Construction of promoter-trap vector pLL192.** The *E. coli–Mycobacterium* shuttle vector pMV261 (Stover et al., 1991), containing an *E. coli* (oriE) and a mycobacterial (oriM) origin of replication, as well as a kanamycin resistance gene and the mycobacterial hsp60 promoter, was chosen to construct pLL192, as shown in Fig. 1. The gene encoding streptomycin resistance (Sm') was amplified by PCR from pHP450 (Prentki & Krisch, 1984), by using primers Strep-F1 (5′-AATTGGGTTACCGGTGATTTGATGC-3′) and Strep-R1 (5′-CCGAAACGCGGAGGAAGCCTTAA-3′). The amplified fragment was directly inserted into the PCR-II-TOPO vector (TOPO TA cloning; Invitrogen). The Sm' gene to be used as a selection marker was then recovered by a BamHI/Asp718 restriction and inserted into pMV261 previously digested with the same enzymes. After transformation of *E. coli* NM522, Sm' clones were isolated. The promoterless gfp gene was amplified by PCR from pJFX2 (Triccas et al., 1999) by using primers GFP Forw2 (5′-GGGATCTGACGTTAGTTAAGAAG-3′) and GFPBlunt (5′-TAAGATATCATTTGTATGCTTGATCAGA-3′) containing a BamHI and an EcoRI site (underlined), respectively. The promoterless kan gene was amplified by PCR from pMV261 by using the primers KanR1 (5′-GGGTCTGACGCTAGTGATC-3′) and KanBlunt (5′-GGGTCTGACGCTAGTTAAGAAG-3′), the latter containing a HindIII site (underlined). The two amplified fragments were then ligated with BamHI/EcoRI and HindIII/SpeI and inserted in a single step into the BamHI/SpeI-digested Sm' pMV261 derivative. The GFPBlunt and KanBlunt primers were defined such that, after cloning, the start codon of the kan gene overlaps the stop codon of the gfp gene. The final construct, named pLL192, contains an artificial bicistronic operon composed of the promoterless gfp gene, followed by the kan gene. The presence of a unique BamHI site immediately upstream of this transcriptional fusion allows for the insertion of promoter libraries.

**Construction of the M. tuberculosis promoter-trap library.** *M. tuberculosis* H37Rv genomic DNA was isolated as described previously by Marmur (1961) and Baulard et al. (1996), and partially digested with Sau3AI. After agarose gel electrophoresis, fragments of 100–800 bp were purified (QIAquick Gel Extraction Kit; Qiagen) and inserted into dephosphorylated BamHI-digested pLL192. After

---

**Fig. 1.** Plasmid map of pLL192.
Macrophage infection. The J774A.1 mouse macrophage cell line was grown in RPMI (Gibco-BRL) or Dulbecco’s modified Eagle medium (DMEM; Gibco-BRL) supplemented with 10% heat-inactivated fetal calf serum (FCS; Gibco-BRL), at 37°C in a 5% CO₂ humidified atmosphere. The cells were seeded in 24-well cell culture clusters (Corning) supplemented with 10% (v/v) ADC. Glycerol (15%, v/v) was added and the mycobacteria were stored at −80°C. Before use, the recombiant mycobacterial pool was grown in Sauton medium to mid-exponential phase.

Table 1. Primers used for promoter and real-time PCR

<table>
<thead>
<tr>
<th>Gene/amplicon</th>
<th>Primer sequences (5’–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetamidase gene upstream</td>
<td>CGGGATCCGAAGTGACCGGTCTCAAGCGTC CGGGATCCAACTACCTCGGGGATGTTGGA CTCCGGGCCTGTCAGTTCA</td>
</tr>
</tbody>
</table>
| Rv0467/icl gene upstream | GAAGATCTCTGGCTGAAACCGCGGCTGG GAAGATCTAGAACAACCTCCAGGTC
| 16S rRNA         | TCCCGGGGCTTGTACA CCACGTGCTGGGTTGTA |
| Rv2232           | CACGAGTGGGCTGGGTTGTA |
| Rv1026           | GGTCCGCGTGGCTGGAAAGC GGCCACCGCGAAAGATGAAT |
| Rv3097c/PE_PGRS63 | GCGATCCGGCTTTGTTTTC CTCCGGCGCTTTGTTTTC |
| Rv0169/mcelA     | ACATGGTGGTGGTGTGTAAGTATGTG CCGCTCTAGGCTAGCCGATC |
| Rv2895c/viuB     | GACCTGCCCCACCGCGAG ACGGCGCGCAGTACCC |
| Rv0997           | GGGCCGATAGCGGTGGGATGGAT GGGCTTGGCGTGGGTTG |
| Rv1635c          | CGGCCGGCTGGTGGTTTTC CGGCCGGTGGTGGTTGC |
| Rv2231/cobC      | GACGCGGCTGGGAAATCGAG CTCGCGGGGAACCCATGACC |
heat-inactivated according to the manufacturer’s instructions. About 200–400 ng RNA was used per reaction. In each set of reactions, 16S rRNA was used for normalization. Each reaction was repeated three times with three independent RNA samples in a Cepheid SmartCycler II. Primers used in real-time PCR are listed in Table 1. A control without reverse transcriptase was included in each set of reactions, and a DNA melting point curve analysis was also done for each sample. Relative expression levels were determined using the $2^{-\Delta\Delta Ct}$ method of Livak & Schmittgen (2001), assuming that the PCR efficiencies were the same as seen from the slope of the reaction curves. The results were expressed as -fold induction of expression compared to the in vitro control.

RESULTS

Construction of vector pLL192 and its validation with known promoters

To identify mycobacterial promoters specifically expressed under defined conditions, we constructed pLL192 (Fig. 1). This Escherichia coli–Mycobacterium shuttle vector contains two promoterless reporter genes, one encoding a mutated form of the green fluorescent protein (gfp) and the second corresponding to the kanamycin resistance gene (kan). These two genes are organized into a single transcription unit located immediately downstream from a unique BamHI site, allowing for the cloning of promoter libraries. To validate pLL192 as a promoter-trap vector, the gfp::kan fusion was placed under the control of three different mycobacterial promoters, namely the constitutive hsp60 promoter (Stover et al., 1991), the icl (isocitrate lyase) promoter (McKinney et al., 2000), known to be induced under specific physiological conditions, and the acetamidase gene promoter (Payton et al., 1999), whose induction is acetamide-dependent. The resulting plasmids were electroporated into M. smegmatis mc²155 and M. bovis BCG and analysed for fluorescence and kanamycin sensitivity. Empty pLL192 was used as a negative control. Fluorescence was observed with bacteria harbouring the hsp60 promoter in pLL192 (Fig. 2a) and these bacteria were resistant to up to 25 μg kanamycin ml⁻¹, compared to the negative control for which more than 80% killing was observed at 1 μg kanamycin ml⁻¹. When the strain containing the pLL192 derivative with the icl promoter was grown in Middlebrook 7H9 medium containing 10 mM acetate or glucose, fluorescence intensity was found to depend on the presence or absence of acetate or glucose (Fig. 2b). Finally, the construct containing the acetamidase gene promoter showed fluorescence intensities proportional to the concentrations of acetamide used in the media (0.2–2.0 mg ml⁻¹; Fig. 2c).

Macrophage infections and isolation of ex vivo-induced promoters

A library of Sau3AI-digested M. tuberculosis chromosomal DNA was generated using pLL192. The library was then introduced into M. bovis BCG and the recombinant BCG strains, along with the control BCG containing pLL192 without insert, were used to infect macrophages. Kanamycin (5 μg ml⁻¹) was found to kill effectively the intracellular BCG control containing pLL192 without insert. This correlated well with the lack of fluorescence (data not shown). Macrophages infected with BCG control containing pLL192 without insert were grown in the presence or absence of kanamycin. After incubation, the macrophages were lysed and the bacteria were plated onto Middlebrook 7H10.
agar with or without kanamycin. Approximately $1.5 \times 10^4$
colonies grew on the plates when the bacteria came from
macrophages grown in the absence of kanamycin and plated
onto Middlebrook 7H10 agar without kanamycin. However,
no colonies were observed for the control BCG after growth
in macrophages cultured in the presence of kanamycin.

When BCG cells containing the promoter library were used
to infect macrophages, grown in the presence of kanamycin
and subsequently plated onto Middlebrook 7H10 agar
without kanamycin, 14 colonies were detected. Eight of
these failed to grow on Middlebrook 7H10 agar in the
presence of kanamycin. These eight clones were then used
again to infect macrophages grown with or without kana-
mycin. No difference in the survival of the BCG clones was
observed between these two conditions, confirming that
these clones could survive in macrophages in the presence of
kanamycin. However, they remained sensitive to kanamycin
when grown in vitro on Middlebrook 7H10 agar, suggesting
up-regulation of the corresponding promoters in macro-
phages ex vivo. When analysed for fluorescence, all eight
colones were found to fluoresce strongly ex vivo, but showed
no or weak fluorescence in vitro, as shown for one example
in Fig. 3.

Plasmid DNA from each of the eight clones was isolated
and the inserts at the BamHI site were sequenced. The
sequences thus obtained were then compared to the genome
sequence of M. tuberculosis (Cole et al., 1998) to identify
the corresponding genes, and they are listed in Table 2.

Real-time PCR analysis using RNA isolated at various
time points from M. tuberculosis and M. bovis BCG grown
in vitro and within macrophages was then carried out to
confirm the up-regulation of these genes. The results are
shown in Fig. 4(a, b). After normalization to the 16S rRNA,
2- to 40-fold induction of the specific genes was observed
both in M. tuberculosis and M. bovis BCG in infected
macrophages, compared to in vitro growth. However, some

Table 2. Genes identified by IVET selection

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function of product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv2232</td>
<td>Conserved hypothetical protein</td>
</tr>
<tr>
<td>Rv1026</td>
<td>Conserved hypothetical protein; could be involved in adaptive phase of bacterial response to amino acid starvation</td>
</tr>
<tr>
<td>Rv3097c/PE_PGRS63</td>
<td>PE-PGRS family protein; probable triacyl glycerol lipase (esterase/lipase)</td>
</tr>
<tr>
<td>Rv0169/mce1A</td>
<td>Involved in host-cell invasion (entry and survival in macrophages)</td>
</tr>
<tr>
<td>Rv2231/cobC</td>
<td>Possible aminotransferase involved in cobalamin biosynthesis</td>
</tr>
<tr>
<td>Rv1635c</td>
<td>Conserved transmembrane protein</td>
</tr>
<tr>
<td>Rv0997</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>Rv2895c/viub</td>
<td>Homologue of vibriobactin utilization protein (ViuB) of Vibrio cholerae, possibly involved in mycobactin utilization</td>
</tr>
</tbody>
</table>

Fig. 3. GFP fluorescence of (a) vector alone (grey area) and clone Rv2232 (black area) in vitro and (b) vector alone (grey area) and clone Rv2232 (black area) within macrophages after 24 h of infection.
difference in gene expression was observed between *M. tuberculosis* and *M. bovis* BCG.

All the genes showed up-regulation at varying levels, and it varied at different time points. PE_PGRS63 (Rv3097c) showed the strongest induction in both the strains, but decreased at 48 h post-infection.

**DISCUSSION**

The identification of *M. tuberculosis* genes expressed intracellularly within macrophages may contribute significantly to the development of new strategies to control tuberculosis. In this investigation we have used a new promoter probe vector with two reporter genes in transcriptional fusion, independent of added inducers, and constructed an expression vector with two reporter genes in transcriptional fusion, as judged to be effective against intracellular bacteria (Kremer *et al.*, 1995; Maurin & Raoult, 2001). The vector was validated by using different known constitutive and inducible promoters. The combination of GFP and selection by kanamycin allowed us to select *M. tuberculosis* promoters which were activated and up-regulated within macrophages compared to mycobacteria grown *in vitro*. Vectors combining GFP with the counterselectable SacB protein (Triccas *et al.*, 1999), or kanamycin pre-selection with SacB post-selection (Saviola *et al.*, 2003), were employed to select *M. tuberculosis* genes expressed within the macrophages. The simultaneous expression of fluorescence and kanamycin resistance phenotypes conferred by pLL192 in the presence of an activated promoter makes the selection process easy in a single step and has the advantage of usability *in vivo*. In the first round, eight clones with the desired phenotype were identified. The activities of the reporter genes under the control of these promoters were minimal *in vitro* as judged by weak or absent fluorescence and kanamycin sensitivity on 7H10 agar, but were high within macrophages. This was confirmed and quantified by RT-PCR. In all cases, expression was up-regulated several-fold within macrophages. Expression of PE_PGRS63 was higher in macrophages than during *in vitro* culture. Within the macrophages, its expression was maximal after 24 h (the first time point) and then decreased. The expression of *mce1A* was observed within 24 h of infection and was similar in both *M. tuberculosis* and *M. bovis* BCG. The presence of *mce1A* in many species of mycobacteria, including mycobacteria other than *tuberculosis*-causing species, suggests that it is an important gene for entry and survival (Arruda *et al.*, 1993).

Rv1026 is a conserved hypothetical protein whose exact function is not yet known, but it may be involved in an adaptive response to amino acid starvation and has been classified as an essential gene by Himar1-based transposon mutagenesis of *M. tuberculosis* (Sassetti *et al.*, 2003). Hence, up-regulation of Rv1026 might reflect the adaptive response to stress within the macrophage.

Acquisition of iron is an essential feature of many bacterial pathogens as iron is required for growth and survival. Mycobacteria have evolved various mechanisms to acquire iron (De Voss *et al.*, 1999), but the role of the potential mycobactin utilization protein ViuB encoded by Rv2895c has not been studied. However, it shows homology to the vibriobactin utilization protein (ViuB) of *Vibrio cholerae* (Butterton & Calderwood, 1994), and may thus be involved in intracellular removal of iron from iron–mycobactin complexes.

Rv2232 is a putative phosphatase whose function remains to be defined. The expression of Rv2232 has been shown to be up-regulated significantly in SCID mice, but not in BALB/C mice (Talaat *et al.*, 2004), suggesting that this gene has a role in immunocompromised hosts. Rv2231c (CobC) is a putative transerase involved in cobalamin biosynthesis. Using Himar1-based transposon mutagenesis in *M. tuberculosis* H37Rv, it was demonstrated that Rv2231c corresponds to an essential gene in H37Rv (Sassetti *et al.*, 2003). The role of Rv1635c and Rv0997 is not known.

The deduced functions of selected genes are consistent with conditions encountered within the macrophages. The induction of their expression observed *ex vivo* may therefore represent a strategy employed by *M. tuberculosis* to promote...
the infection and to enhance intracellular survival. This investigation has led to the selection of promoters that are activated over 48 h after the beginning of the infection. Thus it might have facilitated identification of promoters which are induced early and then down-regulated, which is supported by the RT-PCR data. However, the survival of the bacilli and their multiplication inside macrophages might also involve other genes displaying an up-regulated expression during extended periods of time. Therefore, it might be fruitful to subject the promoter library to selection at different time points, up to 6–7 days, within the macrophages. We appreciate that the present investigation is not a comprehensive study because the promoter library may not be complete as random restriction by Sau3AI has most certainly led to the loss of some promoter fragments involved in intracellular growth. Fragmenting chromosomal DNA by other means may thus be helpful to construct more representative libraries to be used in the context of this vector for the identification of promoters up- or down-regulated under specific environmental conditions.

ACKNOWLEDGEMENTS

The work reported in this paper has been supported by a project jointly awarded to Brahman S. Srivastava and Camille Locht by the Indo-French Centre for Promotion of Advanced Research, New Delhi. The Indian investigators are thankful to the Director of the Institute for laboratory facilities and Mr A. L. Vishwakarma for excellent technical support with flow cytometry. V. S. gratefully acknowledges the Junior and Senior Research Fellowships of the Indian CSIR. This is communication no. 6917 of the CDRI.

REFERENCES


