Regulation of the inducible acetamidase gene of Mycobacterium smegmatis

Tanya Parish,† Eshwar Mahenthiralingam,† Philip Draper, Elaine O. Davis and M. Joseph Colston

Author for correspondence: Philip Draper. Tel: +44 181 959 3666 ext. 2346. Fax: +44 181 913 8528. e-mail: p-draper@nimr.mrc.ac.uk

The inducible acetamidase of Mycobacterium smegmatis NCTC 8159 is expressed at high levels in the presence of a suitable inducer, such as acetamide. The gene and 1.5 kb of upstream sequence had previously been sequenced. A further 1.4 kb of upstream sequence has now been determined, containing an additional ORF on the opposite strand to the acetamidase gene. This ORF has significant homologies to genes encoding regulatory proteins involved in amidase expression in other organisms. Restriction fragments from the 4 kb region were subcloned into a promoter-probe shuttle vector to locate the approximate region of the acetamidase promoter and investigate the mechanism of regulation. An inducible promoter was found to lie in the 1.4 kb region situated 1.5 kb upstream from the acetamidase coding region.

Expression of the acetamidase was studied at the protein and mRNA levels. Using immunoblotting, induction of the enzyme was demonstrated in minimal medium containing succinate plus acetamide, but not in a richer medium (Lemco broth) plus acetamide, confirming that regulation of acetamidase expression is mediated by both positive and negative control elements. After induction by acetamide, an increase above basal level could be detected after 1 h for both protein levels (using ELISA) and mRNA levels (using Northern blot analysis), indicating that control of expression is at the mRNA level. The size of the mRNA transcript detected was approximately 1.2 kb, the size of the acetamidase coding region. Since no promoter was identified immediately upstream of the coding region, this raises the possibility that a larger, primary transcript (possibly polycistronic) is cleaved to produce a stable form encoding the acetamidase protein.

Keywords: Mycobacterium smegmatis, acetamidase, gene expression, inducible gene

INTRODUCTION

Since 1985, with the creation of genomic libraries of Mycobacterium tuberculosis and M. leprae (Clark-Curtiss et al., 1985; Young et al., 1985a, b), over 50 mycobacterial antigen genes have been characterized (Young et al., 1992). Many of these genes cannot be expressed in Escherichia coli under the control of their own promoters since very few mycobacterial promoters are recognized by the E. coli transcription machinery (Clark-Curtiss et al., 1985; Das Gupta et al., 1993). Moreover, few mycobacterial promoters have been directly identified and very little is known about regulatable promoters in mycobacteria.

Amidase activity was first reported in bacterial extracts from M. phlei and M. bovis BCG (Halpern & Grossowicz, 1957). Draper (1967) demonstrated the presence of an inducible amidase enzyme in M. smegmatis NCTC 8159 which enables the organism to grow on simple aliphatic amidines as the sole carbon sources. Induction occurs using acetamide or butyramide as the substrate. The very large increase of a 47 kDa protein, subsequently shown to be the amidase, in cells grown on acetamide as compared...
to succinate, glutamate or Lecombo broth has previously been shown (Draper, 1967; Mahenthiralingam et al., 1993).

The study of the expression of the gene encoding this enzyme should provide a better understanding of how mycobacterial gene expression is regulated. In addition, the isolation and analysis of such a strong and regulatable promoter should be a useful tool to aid the expression of cloned genes in mycobacteria. The gene encoding the acetamidase has been cloned previously and a 2-8 kb fragment was sequenced (Mahenthiralingam et al., 1993).

In this study we have further investigated the mechanisms involved in the regulation of expression of the acetamidase. To identify the location of the promoter we used promoter-probe shuttle vectors containing a chloramphenicol acetyltransferase (CAT) reporter gene. The promoter region for the acetamidase-encoding gene lies in a previously unsequenced region of DNA, 1.5 kb upstream of the acetamidase gene; the nucleotide sequence of this region was determined and a novel ORF with homology to genes encoding regulatory proteins in other bacteria was identified. Northern blot analysis was used to demonstrate that induction of the chromosomal acetamidase occurs at the level of mRNA.

**METHODS**

**Bacterial strains and culture.** *M. smegmatis* NCTC 8159 and *M. smegmatis* mc²¹⁵⁵ (Snapper et al., 1990) were grown at 37 °C in shaking in L-broth (Clare & Meadow, 1995), in Middlebrook 7H9 broth (Difco), both containing 0.05% Tween 80, or in minimal medium (Kohn & Harris, 1941) without glucose, containing 3 ml trace elements 1⁻¹ (Hopwood & Wright, 1978), 0.05% Tween 80 and 2 g carbon source (acetamide, succinate or succinate) ⁻¹. *M. smegmatis* mc²¹⁵⁵ for electroporation was grown on modified Dubos liquid medium (Dubos & Davis, 1946). *Escherichia coli* strainTG2 was grown in L-broth (LB) or 2 × TY medium (Sambrook et al., 1989).

**Preparation of cell-free extracts for SDS-PAGE.** *M. smegmatis* was grown in 500 ml Lemo broth plus Tween 80 (1 conical flask) for 24 h prior to harvesting by centrifugation (5000 g for 5 min) and resuspended in 3 ml sterile distilled water; 0.5 ml of cells was used to inoculate 250 ml conical flasks containing 100 ml minimal medium with either acetamide, succinate or acetamide plus succinate, or 100 ml Lemo broth and Lemosm broth plus acetamide. Cultures were incubated at 37 °C and harvested after 24 h. Cell-free extracts were prepared from *M. smegmatis* cultures using the Mini-BeadBeater (BioSpec Products): 50 mg bacterial cells, suspended in 1 ml 50 mM HEPES/KOH pH 7.5, 10 mM MgCl₂, 60 mM NaN₄Cl, 10% (v/v) glycerol and 5 mM 2-mercaptoethanol, was added to 0.5 ml of 0.1 mm sterile glass beads (Stratche Scientific) and shaken in the Mini-BeadBeater for 1 min. The supernatant was recovered by centrifugation. Cell-free extracts were analysed by SDS-PAGE (Laemmli, 1970); protein concentration was measured using the BCA protein assay (Pierce).

**Immunoblotting.** Cell-free extracts were fractionated by SDS-PAGE and transferred to nitrocellulose membranes (Hybond-C, Amersham) using a semi-dry blotter (Ancos). The filter was blocked for 1 h in Tris-buffered saline (TBS; 20 mM Tris/HCl, pH 7.5, 0.5 M NaCl) containing 3% (w/v) gelatin, washed in TTBS (TBS containing 0.05% Tween 20) and then transferred to the antibody solution (1%, w/v, gelatin in TTBS) containing a 1:1000 dilution of a rabbit anti-serum prepared against the purified acetamidase enzyme (Mahenthiralingam et al., 1993) for 2 h. The filter was then washed in TTBS and incubated for 2 h in antibody solution containing a 1:1000 dilution of peroxidase-conjugated anti-rabbit IgG (Dako); after washing in TTBS and then TBS, the blot was developed by incubation in TBS containing 0.6 mg 4-chloro-1-naphthol ml⁻¹, 20% (v/v) methanol and 0.018% hydrogen peroxide for 20 min.

**ELISA of amidase.** *M. smegmatis* NCTC 8159 was grown with shaking in 5 × 100 ml minimal medium plus Tween 80 (250 ml flasks) for 24 h at 37 °C. One flask was harvested immediately; 2 ml 10% (w/v) acetamide was added to each of the remaining flasks, which were shaken at 37 °C until harvested 0.5, 1 and 2 h later. Cells were washed once with 0.1 M Tris/HCl, pH 7.2, 0.1% Tween 80, then suspended in 1 ml of the same buffer and sonicated for 10 × 1 min at 4 °C at 40 W with a 3:2 mm probe. The sonicates were centrifuged for 30 min at 11 600 g; the supernatants were removed, treated with 20 μl 1% Na₂SO₄ and kept at 4 °C. Protein was measured with the BCA reagent, with a correction to allow for interference with the assay by Tris buffer in the samples. ELISA plates (96-well Maxisorp; A/SNunc) were coated with 2 μg total protein per well in 0.1 ml 0.04 M Na₂CO₃/NaHCO₃ buffer, pH 9.6, which was dried down to the plates at 37 °C. After washing, the plates were blocked with 1% dried skimmed milk powder (Marvel) in PBS containing 0.05% Tween 20. Rabbit anti-amidase serum was used at 1:500 and doubling dilutions as the first antibody. The second antibody was peroxidase-conjugated anti-rabbit immunoglobulin (Dako) and the colour reagent was 2',2'-azino-bis(3-ethylbenzthiazoline)-6-sulphonic acid (Sigma). Colour development was continued at room temperature until the most intense wells had an A₄₀₀ value of between 1 and 2. Plates were read in an EL312 kinetics reader (Bio-Tek) and data analysed using the Instat statistics package (University of Reading).

**Construction of promoter-probe vectors.** Promoter-probe vectors for use in mycobacteria were constructed based on the *E. coli* vector pKK232-8 (Pharmacia Biotech), which contains a promoterless CAT gene. A 2.6 kb Hpal-EcoRV fragment derived from pAL5000 (Rauzier et al., 1988) required for replication in mycobacteria was isolated from pYUB12 (Snapper et al., 1988). This was cloned into pKK232-8 linearized by partial digestion with PstI, to give pEJ94 with the insert in the PstI site derived from pBR322. Following this a 1.2 kb SmaI fragment carrying the kanamycin-resistance gene of Tn5 was isolated from pUC4-KIXX (Pharmacia Biotech) and cloned into pEJ94 linearized by partial digestion with Drai to permit selection in mycobacteria, giving pEJ106. The introduction of this kanamycin-resistance fragment led to the addition of two incomplete HindIII sites flanking a XhoI fragment required for HpaI-EcoRV fragment derived from pYUB12 (Snapper et al., 1988). This was cloned into pKK232-8 linearized by partial digestion with PstI, to give pEJ94 with the insert in the PstI site derived from pBR322. Following this a 1.2 kb SmaI fragment carrying the kanamycin-resistance gene of Tn5 was isolated from pUC4-KIXX (Pharmacia Biotech) and cloned into pEJ94 linearized by partial digestion with Drai to permit selection in mycobacteria, giving pEJ106. The introduction of this kanamycin-resistance fragment led to the addition of two incomplete HindIII sites flanking a XhoI site; these were removed by partial digestion with HindIII followed by isolation of near-full-size DNA, filling in of the sticky ends using Klenow polymerase, and religation. A suitable clone containing a unique HindIII site in the polylinker preceding the CAT gene was identified and termed pEJ108 (see Fig. 3).

**Subcloning of mycobacterial DNA.** Many of the procedures used in the manipulation and cloning of *M. smegmatis* DNA were as described by Sambrook et al. (1989). Restriction fragments of the previously cloned *M. smegmatis* acetamidase
DNA (Mahenthiralingam et al., 1993) were prepared using the appropriate restriction enzymes together with commercial buffers (Life Technologies) and subcloned into pUC18 (Norrander et al., 1983) and subsequently the vectors pEJ106 or pEJ108. Ligations were transformed into *E. coli* TG2; competent cells were prepared by the CaCl₂ method (Cohen et al., 1972). Electroporation of *M. smegmatis* mc²155 was carried out according to the method of Snapper et al. (1988) and transformants were selected on Middlebrook 7H11 agar (Difco) containing 10% (v/v) Dubos medium supplement (Difco) and 10–15 μg kanamycin ml⁻¹. Recombinant plasmids were isolated from transformants using the alkaline lysis method (Burnboim & Doly, 1979); for mycobacterial transforms the initial lysis step was effected for 16 h at 37 °C.

CAT enzyme assays. Mycobacterial transformants were grown in minimal medium containing acetamide or succinate and 15 μg kanamycin ml⁻¹. Cells were harvested from 1.5 ml of culture, washed and resuspended in 130 μl 0.01 M Tris/HCl, pH 8; cell-free extracts were prepared by sonication on ice for 2 x 15 s, with a 15 s rest, using a microprobe (3.2 mm diameter, Branson Ultrasonics). Endogenous acetylases were inactivated by incubation at 60 °C for 10 min, the suspensions were centrifuged for 3 min (10000 × g) and the supernatants recovered. Reactions were set up containing a 44 μl sample of cell-free extract plus 6 μl 2 M Tris/HCl, pH 8, to bring the concentration of Tris to 0.25 M, or 5 μl standards (0, 0.1, 0.05, 0.01 and 0.005 units CAT; Promega) prepared by serial dilution in 0.25 M Tris/HCl (pH 8), made up to 110 μl with 0.25 M Tris/HCl, pH 8, and 10 μl of 1:10 dilution of [³²P]chloramphenicol (2020 MBq mmol⁻¹; Amersham), with 5 μl n-butyl CoA (5 mg ml⁻¹) added to initiate the reaction. Reaction mixtures were incubated at 37 °C for 3 h, then 350 μl xylene was added and the samples were vortexed for 30 s, then centrifuged for 3 min (10000 × g). The organic phase was recovered, extracted twice with 100 μl 0.25 M Tris/HCl, pH 8, and finally a 200 μl sample was measured in a Beckman scintillation counter (model LS2800). Total protein concentration was assayed for each extract and the results are expressed as units of CAT enzyme activity per mg total protein (units mg⁻¹).

**DNA sequencing.** Restriction enzyme fragments from the DNA to be sequenced upstream of the gene were subcloned into the single-stranded bacteriophage vectors M13mp18 and M13mp19 (Norrander et al., 1983). A complete HaeIII digest of the largest restriction fragment (819 bp) was used to produce small fragments which were then cloned into M13mp18. The order of the blunt-ended HaeIII fragments along the sequence was determined by running sequencing reactions for the 800 bp fragment clones; a short stretch of bases was found to be missing so primers (P1 and P2) were synthesized to cover this region. The sequences of the primers were as follows. P1, 5' ATGTAGAGGTGCTGGTCGAA 3' at position 745-759; P2, 5' AGGACGCCTGTTGTTCCAAC 3' at position 316-335; and P5, 5' GACGCGATGGCCTTG 3' at position 1271-1290. Primers P3 (5' CTTGAAGCGGAAT-CTCGATGGTGA 3' at position 1160-1178) and P5 (5' GACGCGATGGCCTTG 3' at position 745-759) were synthesized to sequence across the SstI and HindIII restriction sites and finally a 200pl sample was measured in a Beckman scintillation counter (model LS2800). Total protein concentration was assayed for each extract and the results are expressed as units of CAT enzyme activity per mg total protein (units mg⁻¹).

**RNA extraction.** Approximately 300 mg wet weight of mycobacteria was resuspended in 1 ml guanidinium extraction buffer (6 M guanidinium isothiocyanate, 25 mM sodium acetate, pH 5.2, 0.1 mM dithiothreitol and 1% sarkosyl; all solutions were made up with diethyl-pyrocarbonate-treated deionized water). The suspension was added to an equal volume of heat-sterilized 0.1 mM zirconia beads (Biospec) contained in a 2 ml screw-cap microcentrifuge tube (Sarstedt). Mycobacteria were ruptured by a 3 min pulse on the MiniBeadBeater device. Debris and beads were sedimented by brief centrifugation (10000 g for 3 min) and the cleared lysate (500–700 μl) retained. The remaining mycobacterial pellet and beads were briefly re-extracted on the BeadBeater (30 s pulse) with fresh guanidinium buffer and the resulting extract pooled with the first.

The lysate was extracted once with an equal volume of phenol/chloroform (1:1), followed by extraction with an equal volume of chloroform. Total nucleic acids were precipitated with ethanol (5 μl 5 M NaCl and 250 μl ethanol were added per 100 μl extract) for 1 h at −70 °C, collected by centrifugation, washed with ethanol, dried and dissolved in 500 μl fresh guanidinium buffer. After repeating phenol/ chloroform and chloroform extractions, nucleic acids were precipitated with ethanol and dissolved in 100 μl 10 mM MgCl₂, 20 mM Tris/HCl, pH 8. DNA was digested for 1 h at 37 °C by addition of 5 units RNase-free DNase (Boehringer Mannheim). After removal of DNase by phenol/chloroform extraction, RNA was precipitated with ethanol and dissolved in a minimum volume of deionized water. RNA concentration was determined by absorbance at 260 nm.

**Northern blot analysis.** Formamidine-denatured RNA was separated on gels (1% agarose) containing formaldehyde (Sambrook et al., 1989). RNA molecular size markers (Boehringer Mannheim), stained by addition of 1 μg ethidium bromide ml⁻¹, were included as size standards. RNA was transferred by capillary elution and cross-linked by UV irradiation to nylon membranes (Hybond-N) according to the manufacturer's instructions. Membranes were prehybridized for 2 h in 50% formamide, 5 x SSPE (20 x SSPE is 3 M NaCl, 0.02 M EDTA, 0.2 M NaH₂PO₄, H₂O), 2 x Denhardt’s (50 x Denhardt’s is 1% Ficoll, 1% polyvinylpyrrolidone, 1% bovine serum albumin fraction V), 0.1% SDS at 42 °C prior to overnight hybridization (42 °C) with a 600 bp SalI internal fragment of the acetamidase gene which had been labelled with 100 μCi (37 MBq) [α-³²P]dATP (Dupont) using a random-primed DNA labelling kit (Boehringer Mannheim). After hybridization, filters were washed at 42 °C for 20 min in 1 x SSC (20 x SSC is 3 M NaCl, 0.3 M sodium citrate), 0.1%
Acetamidase induction

Expression of the acetamidase from M. smegmatis NCTC 8159 is induced in cells grown in minimal medium plus acetamide as compared to those grown on succinate or Lemco broth (Draper, 1967; Mahenthiralingam et al., 1993). We wished to determine whether the acetamidase would be induced in the presence of two carbon sources (succinate and acetamide) or in Lemco broth containing acetamide. We used two strains of M. smegmatis, namely NCTC 8159 and M. smegmatis mc²155; results for M. smegmatis NCTC 8159 are shown. M. smegmatis mc²155 is a strain with a high efficiency of transformation which is widely used in genetic studies; it was important to confirm that its acetamidase is similarly inducible to that of NCTC 8159, since we wished to use it for promoter activity assays.

Fig. 1(a) shows the results using Coomassie Brilliant Blue staining to visualize total protein and Fig. 1(b) the corresponding Western blot assay to probe specifically for the acetamidase. The cell-free extract prepared from cells grown on minimal medium with either acetamide alone or succinate plus acetamide showed a large amount of protein of the expected size for the acetamidase (47 kDa), demonstrating that induction had occurred irrespective of the presence of succinate. The extracts from Lemco broth and Lemco broth plus acetamide did not have high levels of the protein, although a small increase could be seen in the presence of the inducer. The complementary Western blot (Fig. 1b) confirmed the identity of this protein as the acetamidase.

The time-course of induction of the amidase by acetamide in the presence of succinate was measured by ELISA, which allowed the detection of small increments in the ‘basal’ amidase level in uninduced cells. A statistically significant increase was apparent after 30 min, though this was small. However, the increase after 1 h was clear (data not shown), which agrees with the measurements of mRNA (see below).

The induction experiments indicate that the expression of the acetamidase is not only controlled by induction with an inducer (in this case, acetamide), but that some form of repression is also involved in the regulation of this gene, as shown by the poor induction in Lemco broth. A small amount of expression occurs in Lemco broth or minimal medium in the absence of added inducer showing that expression is not completely shut off. The results obtained for M. smegmatis mc²155 were similar to those from strain NCTC 8159 (results not shown), demonstrating that this strain also possesses an inducible acetamidase enzyme which does not differ grossly in its physiological regulation from that of NCTC 8159. This result enabled us to use M. smegmatis mc²155 as the host for plasmid constructs.

RNA extraction and identification of acetamidase transcript by Northern blotting

To determine whether amidase expression from the chromosome is controlled at the level of transcription, we carried out Northern blot analysis. Our method allowed the rapid isolation of stable, intact RNA (Fig. 2a), Total RNA extracted from acetamide-induced and uninjured M. smegmatis was analysed using a probe internal to the acetamidase coding region; specific acetamidase mRNA was detected after 1 h of induction in acetamide minimal medium (Fig. 2b). This indicates that regulation of the acetamidase occurs at the transcriptional level and that increased expression of the enzyme is due either to increased transcription of the gene or to increased stability of the mRNA. The size of the transcript (1·2 kb) approximates to the size of the known coding region of the gene. The RNA extraction procedure used produced structurally intact RNA with little degradation of the acetamidase-specific transcript. Although there is a basal level of acetamidase protein in uninduced cells no mRNA could be detected by blotting. This may indicate that mRNA is only present in small amounts in uninduced cells and is thus below the limit of detection of this method. The increased level of acetamidase after 1 h induction noted in the ELISA experiment is consistent with the detection of specific mRNA after a similar period. A small amount of
Regulation of *Mycobacterium smegmatis* acetamidase

Fig. 2. Northern blot analysis of acetamidase transcript. (a) Qualitative analysis of extracted RNA. Ethidium-bromide-stained RNA separated by non-denaturing gel electrophoresis (1% agarose). Lanes are as follows: R, RNA size markers (sizes shown on the left); 0, RNA (30 μg) extracted from 7H9 grown *M. smegmatis*; 1, RNA (30 μg) extracted from *M. smegmatis* induced in acetamide medium for 1 h; 18, RNA (30 μg) extracted from *M. smegmatis* grown for 18 h in acetamide; D, 1 kb DNA ladder (sizes shown on the right). (b) Northern blot of *M. smegmatis* RNA samples separated by denaturing gel electrophoresis (0, 1, 18; shown in panel) which have been probed for the acetamidase gene transcript. The approximate size of the inducible acetamidase transcript is indicated on the right of the panel.

Smearing was visible in both lanes containing RNA from induced mycobacteria, but not in the lane containing RNA from uninduced mycobacteria. Some faint larger bands can be distinguished in the induced lanes, which may indicate that the acetamidase transcript is cleaved from a larger precursor RNA.

Constructs to investigate acetamidase regulation

Two related promoter-probe vectors (pEJ106 and pEJ108) were used to identify the promoter region and other regulatory regions (Fig. 3). Both are derived from the promoter-probe vector pKK232-8 (Brosius, 1984), which contains a promoterless CAT gene. This reporter gene is flanked by efficient transcription terminators to prevent transcriptional read-through from other vector genes; in addition, translational stop codons are present in all three reading frames upstream from the initiation codon of the CAT gene to prevent translational read-through.

To confirm that the complete acetamidase fragment cloned from *M. smegmatis* does contain the inducible promoter and all cis-acting elements responsible for induction, the 4.2 kb BamHI fragment of pAM1 (Mahenthiralingam et al., 1993) was subcloned into the vector pEJ106 and designated pAGAN11. Two other smaller fragments were subcloned into pEJ108: the 2.8 kb HindIII (previously sequenced) fragment, and the 1.4 kb BamHI–HindIII fragment, as indicated in Table 1 and Fig. 4.

Extracts of *M. smegmatis* carrying these constructs, grown on either acetamide or succinate as sole carbon source, were assayed for CAT activity. CAT activity for the control extracts from the parental strain *M. smegmatis* mc2155 was essentially zero with no vector (less than 0.01 units mg⁻¹), and was very low with the vector pEJ108 (less than 0.1 units mg⁻¹) when grown on either medium.

For pAGAN11, containing the 4.2 kb fragment, CAT activity was much higher in acetamide-grown extracts than succinate-grown extracts (Fig. 4). This confirmed that the construct contains a promoter able to direct expression of the CAT gene and that this promoter is inducible by acetamide, with induction being 13-fold. It also indicates that there is no terminator present in the DNA immediately downstream of the acetamidase coding region. For pAGAN60, containing the 2.8 kb fragment including the amidase gene itself, no CAT activity was detected in either acetamide- or succinate-grown extracts, indicating that this construct does not...
Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Relevant genotype*</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. smegmatis</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCTC 8159</td>
<td>Efficient transformation mutant</td>
<td>NCTC strain, London</td>
</tr>
<tr>
<td>mc2155</td>
<td></td>
<td>Snapper et al. (1988)</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG2</td>
<td>supE hsdA5 thi Δ(lac–proAB) Δ(srl–recA)306::Tn10(Tet') F' [traD36 proAB' lacI lacZΔM15]</td>
<td>Sambrook et al. (1989)</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUC18</td>
<td>4152 bp BamHI fragment of <em>M. smegmatis</em> genomic DNA in pUC18</td>
<td>Norrander et al. (1983)</td>
</tr>
<tr>
<td>pAMI1</td>
<td>Promoter selection vector, <em>E. coli</em> and mycobacterial origins of replication, BamHI cloning site, Ap+ Km+</td>
<td>Mahenthiralingam et al. (1993)</td>
</tr>
<tr>
<td>pEJ106</td>
<td>As pEJ106, but with unique HindIII site</td>
<td>This paper</td>
</tr>
<tr>
<td>pEJ108</td>
<td>4152 bp BamHI acetamidase fragment of pAMI1 in pEJ106, containing acetamidase and 2886 bp upstream region</td>
<td>This paper</td>
</tr>
<tr>
<td>pAGAN11</td>
<td>2793 bp HindIII fragment of pAMI1 in pEJ108, containing acetamidase and 1521 bp upstream region</td>
<td>This paper</td>
</tr>
<tr>
<td>pAGAN60</td>
<td>1365 bp BamHI–HindIII fragment of acetamidase upstream region in pEJ108</td>
<td>This paper</td>
</tr>
</tbody>
</table>

*Ap*, ampicillin resistance; Km*, kanamycin resistance.

Fig. 4. (a) ORFs identified upstream from the acetamidase. ORFs are shown as black arrows in the direction of transcription. ORFs 1–3 were identified previously (Mahenthiralingam et al., 1993); ORFc was identified in the present work. (b) Constructs tested for CAT activity. Restriction fragments were prepared from the plasmid pAMI1 (Mahenthiralingam et al., 1993), containing the acetamidase structural gene and 2.8 kb of upstream sequence, and subcloned into the vectors pEJ106 (for the 4 kb BamHI fragment) and pEJ108 (other fragments). Cell-free extracts were prepared from *M. smegmatis* transformants containing recombinant pAGAN plasmids and assayed for CAT activity; transformants were grown on minimal medium containing either acetamide or succinate as the sole carbon source. For each value, three transformants were used to prepare extracts and separate CAT assays carried out three times for each extract (mean results and standard errors are shown). Restriction sites are marked on the diagram; B, BamHI; H, HindIII. The CAT gene was present at the right-hand end of all the constructs, as depicted in the figure.

contain a functional promoter. CAT activity for pAGAN13 transformants in acetamide-grown extracts was comparable to the activity in acetamide-grown extracts with the pAGAN11 construct (Fig. 4). However, the level of activity in succinate-grown extracts was much higher than seen with pAGAN11, so the induction ratio was only twofold. This indicated that the inducible promoter lies within the 1.4 kb upstream, previously unsequenced, region and that some form of negative control had been lost in pAGAN13.
Regulation of Mycobacterium smegmatis acetalidase

Nucleotide sequence

Once it had been shown that the 1.4 kb region situated 1.5 kb upstream of the acetalidase gene contained the inducible fragment, the position was sequenced. The complete sequence is presented in Fig. 5. A total of 1371 bp of nucleotide sequence of the acetalidase upstream region was determined. The G+C content (67.7 mol%) is consistent with the recorded figure for total M. smegmatis genomic DNA of 65–67 mol% (Bradley, 1971). The region sequenced terminates at the sole HindIII restriction site and overlaps the region

Fig. 5. Nucleotide sequence of the 1371 bases of upstream sequence with the putative ORFc amino acid sequence below. A putative start codon was identified on the basis of the alignment with homologous sequences shown in Fig. 6; a potential ribosome-binding site which partially overlaps this start codon is underlined. The stop codon of ORFc is shown in Fig. 5.

Fig. 6. Multiple alignment of the predicted product of ORFc. Three sequences with homology to the ORFc product were identified by searching the Swiss Protein and GenBank databases and aligned using the program PILEUP. Conserved residues present in at least three of the sequences are shown in bold; invariant residues are marked with an asterisk. ORFc, putative gene for M. smegmatis regulatory protein; nhhc, positive regulator of high-molecular mass nitrile hydratase from Rhodococcus rhodochrous J1 (Komeda et al., 1996a); nhlc, positive regulator of low molecular mass nitrile hydratase and amidase expression from Rhodococcus rhodochrous J1 (Komeda et al., 1996b); amic, negative regulator of amidase expression from Pseudomonas aerugiosa (Wilson & Drew, 1991).
sequenced previously (Mahenthiralingam et al., 1993) by six nucleotides; this was confirmed using primer P3 to sequence across the HindIII site (Fig. 5).

The sequence was analysed and an ORF was identified on the opposite strand to the acetamidase coding region (called ORFc; Figs 4 and 5). A database search with ORFc identified three proteins with significant homology: the products of nhbC and nhIC from Rhodococcus rhodochrous J1 and the product of amic from Pseudomonas aeruginosa. The alignment of these four proteins generated by the program PILEUP is shown in Fig. 6. The three non-mycobacterial proteins have been identified as being involved in the regulation of amidase expression; AmiC is a negative regulator of the amidase operon in P. aeruginosa (Wilson & Drew, 1991), whereas NhHC and NhIC are both positive regulators in Rhodococcus rhodochrous J1. NhHC is a positive regulator of the high-molecular-mass nitrile hydratase (Komeda et al., 1996a) and NhIC is involved in the regulation of both the low-molecular-mass nitrile hydratase and amidase expression (Komeda et al., 1996b). The AmiC protein has been shown to bind to amides and another regulatory protein (AmiR), thus acting as an amide sensor and mediating its regulatory effect via protein–protein interaction rather than binding directly to the DNA (Wilson et al., 1993). These homologies suggest that ORFc encodes a regulatory protein involved in sensing amides.

**DISCUSSION**

Experiments using the promoter-probe vector confirmed that the promoter and other elements required for the induction of acetamidase expression are found within the 4.2 kb cloned fragment of M. smegmatis genomic DNA. The promoter was found to lie in the 1.4 kb region, which is 1.5 kb upstream from the acetamidase coding region. In addition the results demonstrated that some kind of negative control element was lacking in this region and therefore must lie in the 1.5 kb region immediately upstream of the coding region (or be disrupted in pAGAN13). The high CAT activity of transformants containing constructs pAGAN11 and pAGAN13 confirmed that the promoter is a strong one in comparison with other mycobacterial promoters (data not shown) and was able to drive transcription through the acetamidase gene and into the CAT gene. The constructs possessed a small amount of DNA downstream of the acetamidase coding region (56 bp) and these results demonstrated that this region does not contain a transcription terminator.

Three other ORFs had been identified in the previously sequenced region, encoding putative polypeptides of about 15, 10 and 20 kDa (ORF1, ORF2 and ORF3 respectively; Mahenthiralingam et al., 1993). ORF1 has recently been identified as having homology with a regulatory gene (slyA) in E. coli which is responsible for switching on a previously silent haemolysin gene and appears to be similar to a diverse group of genes that encode regulatory proteins, including MarR, which regulates the multiple-antibiotic-resistance response in E. coli, and Hpr, a regulator of oxidative stress response production in Bacillus subtilis (Oscarsson et al., 1996). We therefore predict that ORF1 encodes a regulatory protein.

ORF2 shows homology to fmdB of Methylophilus methylotrophus, which encodes a regulatory protein involved in formamidase expression, and the acetamidase coding region itself shows striking homology (57% identity) to the formamidase coding region, fmdA (Wyborn et al., 1996). This suggests that ORF2 is also involved in amidase regulation. (Formamidase is an excellent substrate for the mycobacterial amidase; Draper, 1967.) ORF3 is homologous to amidase and urease transport genes, namely amiS of P. aeruginosa (Wilson et al., 1995), amiS2 of R. rhodochrous, urel of Helicobacter pylori (Cussac et al., 1992) and urel of Streptococcus salivarius (Chen et al., 1996); the former two genes have been identified as encoding probable membrane transport proteins. Thus it seems unlikely that ORF3 is involved in gene regulation, but more probably has a function in amide transport.

The fact that the amidase promoter is located so far upstream from the acetamidase coding region, and the arrangement of and spacing between the ORFs, add further weight to the idea that the acetamidase gene is part of an operon, where the acetamidase is transcribed as part of a polycistronic message. The ORFs so far identified by sequence homologies encode three regulatory proteins (ORF1, 2 and c) and one transport protein (ORF3). The genes encoding regulatory proteins found in this region are similar to those of amidase operons from Methylophilus methylotrophus, P. aeruginosa and R. rhodochrous; however their arrangement in Mycobacterium smegmatis is unique.

From the arrangement of ORFs in the sequence, we would expect a polycistronic message to be transcribed. However, Northern blotting identified a transcript in induced cells with a size of 1.2 kb spanning the amidase coding region only. In addition, the promoter assays did not reveal any promoter activity located in the 2.8 kb HindIII fragment covering this region (pAGAN60). This discrepancy could be explained if there is a promoter located in this fragment which fails to respond to induction due to the absence of a control site or regulatory protein. Alternatively, a polycistronic mRNA may be transcribed from an upstream promoter, which is then processed to generate the 1.2 kb mRNA species encoding the amidase alone. The predicted secondary structure of the mRNA in a polycistronic message reveals a potential processing site with a stem–loop structure just upstream of the amidase coding region; cleavage of RNA would be predicted to occur in one of the loops. After such processing the upstream message would be degraded from the 3' end and would be less stable than the amidase coding message. Another possibility exists of the initiation of transcription at a large distance from the promoter, although this seems unlikely; alternatively the inducible promoter may be
required for the expression of one of the small ORFs which would be in turn needed for promoter activity close to the acetamidase gene itself. However, in the latter case there would still be one copy of this ORF present on the chromosome which would be expressed. This should be able to direct a small amount of expression from the plasmid-encoded copies of such a promoter. Since no promoter activity was detected, this situation seems unlikely.

Results from the growth of M. smegmatis on various media indicate that acetamidase induction is controlled by a form of catabolite repression; whether this is in addition to the repression seen in the CAT assays is unclear. A small amount of expression occurs in Lemco broth or minimal medium in the absence of added inducer showing that expression is not completely shut off. There may be two independent methods of negative control operating on acetamidase expression: repression in the absence of inducer (corresponding to succinate medium) and repression in the presence of inducer and a catabolite repressor in a rich medium (Lemco plus acetamide). Whether these two are independent or operate via the same mechanism is unclear. Catabolite repression has been demonstrated in mycobacteria (Ratledge, 1982); such effects arise from the reduction in the level of cAMP in the presence of glucose (Padh & Venkitasubramanian, 1976). The molecular basis is well known in E. coli, where cAMP combines with the catabolite receptor protein (CRP; also known as CAP, the catabolite activator protein) and this complex stimulates transcription of the genes concerned. However, it is not known whether the same mechanism is responsible for catabolite repression in mycobacteria.

Although the precise mechanism of amidase regulation is still unknown, we have provided evidence that regulation of acetamidase expression is by two different systems: repression under induced conditions and induction by acetamide. We have identified another coding region for a putative regulatory protein (ORFc). We have confirmed that induction is at the level of induction by acetamide. We have identified another coding region for a putative regulatory protein (ORFc).

REFERENCES


ACKNOWLEDGEMENTS

T.P. was supported by a Postgraduate Research Studentship from the Medical Research Council of Great Britain. E.M. was supported by funds provided by the Canadian Diseases Network. We are grateful to Neil Stoker for critical reading of the manuscript.


Received 2 December 1996; revised 5 February 1997; accepted 24 February 1997.