Microbiology (2010), **156**, 3445–3455 **DOI** 10.1099/mic.0.042275-0

ClgR regulation of chaperone and protease systems is essential for *Mycobacterium tuberculosis* parasitism of the macrophage

Megan Estorninho,[1] Hilde Smith,[2] Jelle Thole,[2] Jose Harders-Westerveen,[2] Andrzej Kierzek,[1] Rachel E. Butler,[1] Olivier Neyrolles[3] and Graham R. Stewart[1]

Correspondence
Graham R. Stewart
G.Stewart@surrey.ac.uk

[1]Division of Microbial Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford GU2 7XH, UK

[2]Central Veterinary Institute of Wageningen UR, Division of Infectious Diseases, Animal Sciences Group, Edelhertweg 15, 8219 PH Lelystad, The Netherlands

[3]Institut de Pharmacologie et de Biologie Structurale, Centre National de la Recherche Scientifique and Université Paul Sabatier (UMR 5089), 205 Route de Narbonne, 31000 Toulouse, France

Received 4 June 2010
Revised 14 July 2010
Accepted 3 August 2010

Chaperone and protease systems play essential roles in cellular homeostasis and have vital functions in controlling the abundance of specific cellular proteins involved in processes such as transcription, replication, metabolism and virulence. Bacteria have evolved accurate regulatory systems to control the expression and function of chaperones and potentially destructive proteases. Here, we have used a combination of transcriptomics, proteomics and targeted mutagenesis to reveal that the *clp* gene regulator (ClgR) of *Mycobacterium tuberculosis* activates the transcription of at least ten genes, including four that encode protease systems (ClpP1/C, ClpP2/C, PtrB and HtrA-like protease Rv1043c) and three that encode chaperones (Acr2, ClpB and the chaperonin Rv3269). Thus, *M. tuberculosis* ClgR controls a larger network of protein homeostatic and regulatory systems than ClgR in any other bacterium studied to date. We demonstrate that ClgR-regulated transcriptional activation of these systems is essential for *M. tuberculosis* to replicate in macrophages. Furthermore, we observe that this defect is manifest early in infection, as *M. tuberculosis* lacking ClgR is deficient in the ability to control phagosome pH 1 h post-phagocytosis.

INTRODUCTION

Proteases and their associated chaperones carry out an essential homeostatic role in all living cells. They catalyse the degradation and recycling of truncated, damaged or aggregated proteins. Protease systems also have a direct regulatory role in the physiological state of the cell by controlling the abundance of specific proteins involved in a variety of fundamental cellular processes such as transcription, DNA replication and cell division (Darwin, 2009; Frees *et al.*, 2007; Gottesman, 2003).

Abbreviations: CDS, coding sequences; ClgR, Clp gene regulator; EMSA, electrophoretic mobility shift assay; FAM, 5-carbofluorescein; GST, glutathione *S*-transferase; LC, liquid chromatography; pfp, percentage of false positives; qRT-PCR, quantitative real-time PCR; TAMRA, *N*,*N*,*N*′,*N*′-tetramethyl-6-carboxyrhodamine.

Accession numbers for the microarray data associated with this paper are given in the text.

Two supplementary figures and four supplementary tables are available with the online version of this paper.

Bacteria possess tens of different proteases but a significant proportion of protein degradation is performed by a small number of ATP-dependent proteases such as Clp, Lon, HsIUV and FtsH (Gottesman *et al.*, 1997). Each of these enzymes has an ATPase domain and a proteolytic domain encoded either on a single polypeptide or on two separate polypeptides. The Clp protease system is the most studied and is present in all bacteria except the Mollicutes. The archetypal Clp protease of *Escherichia coli* constitutes an assembly of two heptameric rings of the proteolytic ClpP subunits, flanked by two hexameric rings of an ATPase chaperone, either ClpA or ClpX. Additional members of the ATPase/Clp chaperone family, such as ClpC, ClpE and ClpL, exist in other bacteria (Butler *et al.*, 2006). The extent and importance of Clp-mediated proteolysis is well illustrated in *Bacillus subtilis*, where ClpP is responsible for 50 % of cellular protein turnover and is a major control factor in the transition from exponential- to stationary-phase growth (Kock *et al.*, 2004) and the downregulation of central metabolism in glucose-starved cells (Gerth *et al.*,

2008). The influence of ClpP extends to diverse areas of cellular activity, and the virulence of several bacterial pathogens, including *Streptococcus mutans*, *Streptococcus pneumoniae*, *Staphylococcus aureus* and *Listeria monocytogenes*, has been shown to rely on ClpP and Clp ATPases (Frees *et al.*, 2004; Gaillot *et al.*, 2000; Ibrahim *et al.*, 2005; Kajfasz *et al.*, 2009).

The genome of *Mycobacterium tuberculosis* encodes approximately 50 proteases, including two paralogues of ClpP and the associated ATPase chaperones ClpC and ClpX (Ribeiro-Guimarães & Pessolani, 2007). In addition, *M. tuberculosis* possesses FtsH and is one of a select group of bacteria that possess a eukaryote-like ATP-dependent proteasomal apparatus. In *M. tuberculosis*, the latter system has been shown to degrade proteins that have been post-translationally tagged with small protein modifiers in a system analogous to eukaryote ubiquitination (Darwin, 2009). The *M. tuberculosis* ClpP and FtsH enzymes appear to be essential cellular components (Sassetti *et al.*, 2001) and the proteasome is important for virulence (Darwin *et al.*, 2003).

Regulation of these proteolytic systems is extremely important. In Gram-negative bacteria, including *E. coli*, *clp* genes are regulated primarily by the alternative sigma factor σ^{32}. Low-GC Gram-positive bacteria such as *B. subtilis* have evolved a negative control strategy using the transcriptional repressor CtsR. In contrast, the high-GC actinomycetes appear to have adopted a highly conserved transcriptional activator ClgR (*clp* gene regulator), which has been shown to control *clp* genes in *Corynebacterium glutamicum* (Engels *et al.*, 2004, 2005), *Streptomyces lividans* (Bellier & Mazodier, 2004; Bellier *et al.*, 2006) and *Bifidobacterium breve* (Ventura *et al.*, 2005). A homologue of ClgR exists in *M. tuberculosis* and its expression is highly upregulated in stress conditions such as heat shock (Stewart *et al.*, 2002) and during macrophage infection (Schnappinger *et al.*, 2003). Two recent studies of *M. tuberculosis* and *Mycobacterium smegmatis* suggest that ClgR also regulates expression of ClpP1/ClpP2 and ClpC1 in mycobacteria (Barik *et al.*, 2010; Mehra & Kaushal, 2009). A wider examination of the ClgR regulon has not been made prior to the present study.

One of the central features of *M. tuberculosis* pathogenesis is an ability to enter and replicate in the macrophage immune cell of the host in a process that involves the bacterial modulation of phagosome maturation (Russell, 2001). It is clear from studies of gene transcription that the bacterial adaptation and response to the macrophage involves a dramatic remodelling of the cell (Schnappinger *et al.*, 2003; Tailleux *et al.*, 2008). The extent to which proteolysis contributes to these gene regulatory changes is only just beginning to be investigated (Barik *et al.*, 2010) and the direct contribution of proteolysis to the remodelling of the wider cellular protein content has not been examined. In this study, we investigate the regulation of proteases and chaperones by ClgR in *M. tuberculosis* and we examine how dysregulation of proteolytic control affects mycobacterial parasitism of the macrophage.

METHODS

Bacterial strains and culture conditions. *Escherichia coli* DH5α TOP10 (Invitrogen, UK) and BL21(DE3) pLysS were grown at 37 °C in Luria–Bertani broth and agar or 2× YTA broth containing hygromycin 150 μg ml^{-1} or 100 μg ampicillin ml^{-1} as appropriate. *M. tuberculosis* H37Rv was grown at 37 °C in Middlebrook 7H9 broth (Becton Dickinson) or on Middlebrook 7H11 medium (Becton Dickinson) containing 10 % albumin/dextrose (glucose)/catalase (ADC) enrichment and 10 % oleic acid/ADC (OADC) enrichment respectively and 50 μg hygromycin ml^{-1} or 15 μg kanamycin ml^{-1} as appropriate.

Gene replacement of *M. tuberculosis clgR*. Regions of DNA flanking *clgR* (Rv2745c) were PCR-amplified from genomic DNA by using HF Expand polymerase and the primer pairs Rv2745A (5′-GGACTAGTACTGGTGGTGTCCAGCCAGT-3′)/Rv2745B (5′-GGACTAGTACGACCTCACGCACCAA AGC-3′) and Rv2745C (5′-GCTCTAGACGGTGGTGTCGCTGGCGGTG-3′)/Rv2745D (5′-GCTCTAGAATCGCCGCCGAAGTCTGGTTG-3′) for up- and downstream regions, respectively. The regions were cloned into the Zero Blunt TOPO vector (Invitrogen) and then *Spe*I fragments were subcloned around the hygromycin-resistance gene (*hyg*) in the suicide delivery vector pG5. PG5 carries an additional kanamycin-resistance gene to aid cloning in *E. coli* and the *sacB* gene to provide counterselection for single-crossover integration of the gene-replacement vector. The resulting plasmid, pM6, was introduced into *M. tuberculosis* H37Rv via electroporation and gene replacement transformants (Δ*clgR*) were selected as previously described (Dussurget *et al.*, 2001). Gene replacement of *clgR* with *hyg* was confirmed by Southern blot analysis of BspEI-digested genomic DNA probed with the Rv2745A/Rv2745B PCR product.

To complement the deletion of *clgR*, we chose to reintroduce a single copy of the *clgR* gene with its own promoter. The *clgR* ORF and 381 bp of upstream sequence was amplified from genomic DNA by PCR, using the primers Rv2745complA (5′-TGCGGCCGCTCGGATTGTCGATGC-3′) and Rv2745complB (5′-TGCGGCCGCTTAGGCCACCGCCAG-3′) and cloned into the Zero Blunt TOPO vector. This vector was digested with the *Not*I restriction enzyme, and the *clgR* fragment subcloned into the integrating vector pKinta (Stewart *et al.*, 2001). The resulting vector, pK2, was introduced in to *M. tuberculosis* Δ*clgR* via electroporation and transformants were selected by plating on 7H11 medium containing kanamycin.

RNA extraction. Ten millilitres of broth culture in late exponential phase was added directly to 20 ml RNAprotect Cell Reagent (Qiagen) and immediately vortexed. The bacteria were pelleted by centrifugation and resuspended in 1.2 ml Trizol (Invitrogen) with 0.5 ml of 0.1 mm silica-ceramic beads and processed at 6±5 W for 45 s in a Fastprep homogenizer. The samples were centrifuged at 13 000 r.p.m. for 15 min in a microfuge, to remove bacterial debris, and the supernatant was transferred into a fresh tube. The phases were separated by the addition of 0.6 ml chloroform, mixing and centrifugation. The aqueous phase was re-extracted with chloroform and the RNA was precipitated with propan-2-ol, washed in 70 % ethanol and dissolved in RNase-free water. The RNA was cleaned up by RNeasy purifcation (Qiagen), including on-column DNase I digestion.

RNA concentrations were assessed using a NanoDrop ND-1000 UV-Vis spectrophotometer (Thermo Scientific) and the quality of the RNA was determined using an Agilent 2100 Bioanalyser and an Agilent RNA 6000 Nano kit (Agilent).

Transcriptional profiling by microarrays. cDNA was labelled by incorporation of Cy3 or Cy5 dCTP (Amersham) during reverse transcription of RNA, essentially as described previously (Stewart

et al., 2002). RNA (10 μg) was mixed with 3 μg random hexamer oligonucleotides, heated to 95 °C and snap-cooled. In a total volume of 25 μl, the labelling reaction was initiated by the addition of 5 μl First Strand Buffer, 25 mM DTT, 1 mM each dATP, dGTP and dTTP, 0.4 mM dCTP, 2 nmol Cy3- or Cy5-dCTP and 500 U Superscript II reverse transcriptase (Invitrogen). The reaction mixture was incubated in the dark at 25 °C for 10 min and then at 42 °C for 90 min. The relevant pairs of labelled cDNA (wild-type H37Rv versus either Δ*clgR* or the complemented mutant strain) were mixed and purified using a Qiagen MinElute kit, eluting in water.

The paired cDNAs were compared by co-hybridization to whole-genome microarrays, prepared by spotting PCR products representing all predicted *M. tuberculosis* H37Rv coding sequences (CDS) onto Corning GAPS-coated slides. The microarrays were provided by the Bacterial Microarray Group at St Georges, University of London (BμG@S), and the array design is available in BμG@Sbase (accession number E-BUGS-23; http://bugs.sgul.ac.uk/A-BUGS-23) and also ArrayExpress (accession number A-BUGS-23). Dye swaps were included to avoid dye bias effects. Signal intensities of hybridization were collected using Genepix Pro 3.0 and an Axon 4000B microarray scanner and image processing was done with Bluefuse software for microarrays (Bluegnome). Fully annotated microarrays have been deposited in BμG@Sbase (accession number E-BUGS-96; http://bugs.sgul.ac.uk/E-BUGS-96) and also ArrayExpress (accession number E-BUGS-96).

Reverse transcription PCR. Quantitative real-time PCR (qRT-PCR) analysis was applied to cDNA samples to independently confirm the transcription levels of a small subset of genes highlighted by the microarray results (*acr2*, *clpP1* and *clpC*).

qRT-PCR was performed using the Applied Biosystems (ABI) Prism 7000 sequence detection system for quantitative gene expression analysis. Primer Express software (ABI) version 2.0 was used to design the TaqMan probe and primer sets, according to the parameters described by ABI (Supplementary Table S4, available with the online version of this paper). Primers were unmodified and probes were dually labelled with 5-carbofluorescein (FAM) at the 5′ end and *N*,*N*,*N*′,*N*′-tetramethyl-6-carboxyrhodamine (TAMRA) at the 3′ end. cDNA synthesis was performed with Superscript II (Invitrogen) using the manufacturer's protocol. PCR was performed with appropriate primers and probes and TaqMan Universal PCR Mastermix (ABI) as directed by the manufacturer but with half the reaction volumes.

Quantification of mRNA transcript levels was carried out with the ABI proprietary software against a gDNA standard curve. The mean ratio of expression of *acr2*, *clpP1* and *clpC* was calculated for the Δ*clgR* strain versus wild-type *M. tuberculosis*. Pearson's correlation coefficient was used to compare qRT-PCR ratios with microarray expression ratios for the gene targets.

Proteomic profiling by LC/MS-MS. Three 200 ml cultures each of wild-type and Δ*clgR M. tuberculosis* were grown in shaking cultures to OD_{600} ~0.8. The bacteria were pelleted by centrifugation, washed twice in cold PBS and resuspended at 300 mg ml^{-1} in PBS, containing a cocktail of protease inhibitors. The cells were disrupted using 0.1 mm silica-ceramic beads in a Fastprep homogenizer. Urea was added to the homogenate to a concentration of 9 M, dithiothreitol was added to 70 mM and Triton X-100 to 2 %. The samples were incubated, with mixing, for 30 min and then centrifuged to remove debris before filter-sterilization.

Equal amounts of total protein (50 μg) were separated by SDS-PAGE. The proteins in the gel were stained with Coomassie brilliant blue. After staining, the gels were washed in MilliQ water and each lane was sliced into ten bands; each band was processed for in-gel digestion according to the method of Shevchenko *et al.* (1996). Briefly, bands were washed and dehydrated three times in 50 mM ammonium bicarbonate (ABC), pH 7.9 and 50 mM ABC + 50 % acetonitrile (ACN), respectively. Subsequently, cysteine bonds were reduced with 10 mM dithiotreitol for 1 h at 56 °C and alkylated with 50 mM iodoacetamide for 45 min at room temperature in the dark. After two subsequent washing/dehydration cycles the bands were dried for 10 min in a vacuum centrifuge and incubated overnight with 0.06 μg $μl^{-1}$ trypsin at 25 °C. Peptides were extracted once in 1 % formic acid and subsequently twice in 50 % ACN in 5 % formic acid. The volume was reduced to 50 μl in a vacuum centrifuge. Subsequently peptides were separated by LC-MS/MS and the resulting MS/MS spectra were searched against the genome sequence of *M. tuberculosis* H37Rv using Sequest. The resulting files were imported into Scaffold (Proteome Software). The number of spectra per protein per sample was normalized against the total number of measured spectra. Only protein identifications with a probability of >95 % in at least two out of the triplicate samples were taken for further analysis.

Expression and purification of recombinant ClgR. The ORF encoding ClgR (Rv2745c) was amplified by PCR from *M. tuberculosis* genomic DNA, using HF Expand polymerase (Roche) and the primers pGEX/Rv2745cFOR (5′-GAATTCCAATGGCGGCTTTGG-3′) and pGEX/Rv2745cREV (5′-CTCGAGTTTAGGCCACCGCCA-3′), cloned into the pCRII-Blunt-TOPO vector and transformed into TOP10 *E. coli* as directed in the Zero Blunt TOPO PCR Cloning kit (Invitrogen). The cloned fragment was excised from the TOPO vector using *Eco*RI and *Xho*I restriction enzymes and subcloned into the *Eco*RI and *Xho*I sites of pGEX-4T2 vector (Novagen) to make pGEX-4T2_Rv2745c. pGEX-4T2_Rv2745c was transformed in to *E. coli* BL21(DE3)pLysS and clones were grown overnight at 37 °C, with shaking, in 10 ml of 2× YTA medium with 100 μg ampicillin ml^{-1}.

One litre volumes of broth supplemented with 100 μg ampicillin ml^{-1} were inoculated with the overnight cultures at 1 : 100 dilution and grown to OD_{600} 0.5 at 30 °C. Expression of the protein was induced by addition of 0.4 mM IPTG and incubation at 30 °C for 4 h. The cells were pelleted by centrifugation and lysed by sonication in PBS containing Complete Protease Inhibitors (Roche). The lysate was centrifuged at 12 000 ***g*** for 15 min at 4 °C and soluble glutathione *S*-transferase (GST)-tagged ClgR was isolated using glutathione-Sepharose 4B (Amersham Biosciences), with elution in 20 mM glutathione, and assessed by SDS-PAGE. The recombinant protein readily precipitated during storage but samples were stable for short periods at 4 °C. Instability of the recombinant ClgR also precluded cleavage of the GST tag.

Electrophoretic mobility shift assays. Interaction of GST-ClgR with the putative operator sequence in the *acr2* promoter region was assessed by electrophoretic mobility shift assay (EMSA) using a DIG-labelled oligonucleotide generated by annealing Rv0251c+ (5′-TTCCACCTCGGCGTTCACCGGAAGCGAACACTGTCACACAG-3′) and Rv0251c−(5′-CTGTGTGACAGTGTTCGCTTCCGGTGAACGC-CGAGGTGGAA-3′) as directed in the Gel Shift kit (Roche Diagnostics). Binding reactions were performed at 25 °C for 15 min and the products were electrophoresed in 6 % polyacrylamide before transfer to a positively charged nylon membrane and detection using alkaline phosphatase-conjugated anti-DIG Fab fragments and CSPD chemiluminescent substrate.

Bacterial growth assays in murine bone-marrow-derived macrophages. Macrophages were derived from bone marrow cells by culture for 7 days in RPMI medium containing 10 % fetal calf serum and supplemented with 20 % L929 conditioned medium. Macrophages were seeded into 24-well tissue culture plates at a concentration of 4×10^5 cells per well and allowed to adhere overnight. *M. tuberculosis* strains were grown to mid-exponential

phase, pelleted by centrifugation, washed and resuspended in RPMI to a concentration of 4×10^5 c.f.u. ml^{-1}. The bacteria were added to the monolayers at an m.o.i. of 1 for 2 h. The monolayers were washed three times with PBS to remove free bacteria and then cultured in RPMI at 37 °C with 5 % CO_2. Intracellular survival and growth was assessed by lysis of the monolayers and enumeration of bacteria by the plating of serial dilutions on Middlebrook 7H11.

Phagosome maturation assay. *M. tuberculosis* cells were surface labelled by incubation with FITC (1 μg ml^{-1} in carbonate buffer, pH 9.2). To analyse phagosome maturation, murine bone-marrow-derived macrophages were seeded in chamber slide wells overnight. Fluorescent bacilli were added to the monolayers at an m.o.i. of 5 and in the presence of LysoTracker Red DND90. The infections were synchronized by allowing the bacteria to adhere to host cells for 30 min at 4 °C before the replacement of the inoculum with fresh RPMI medium and incubation at 37 °C in 5 % CO_2 for 1 h. The monolayers were washed three times and fixed in 2 % paraformaldehyde in PBS before visualization using a Zeiss LSM510 confocal laser scanning microscope.

Bioinformatic and statistical analyses. Consensus sequences for ClgR binding motifs were generated using MEME (Maximum Em for Motif Elicitation) (Bailey & Elkan, 1994). The *M. tuberculosis* H37Rv genome was searched for ClgR binding motifs using MAST (motif alignment and search tool) (Bailey & Gribskov, 1998). Microarray data were analysed with LIMMA (Smyth *et al.*, 2003; Smyth, 2005) and RankProd (Breitling *et al.*, 2004) R software packages. The print-tip loess procedure of the LIMMA package was used for array normalization (Yang *et al.*, 2002). The means of duplicated spots on the arrays were calculated, and between-array normalization was performed where the log-ratio intensities were scaled to have the same median absolute deviation across arrays (Yang *et al.*, 2002). Rank product (Breitling *et al.*, 2004) analysis was used to identify differentially expressed genes. The percentage-of-false-positives value (pfp) of 0.10 was used as a guide to identify genes in which expression had significantly changed. Survival of individual mutants in macrophages, and frequencies of colocalization with LysoTracker, were analysed by Student's *t*-test.

RESULTS

The ClgR regulon of *M. tuberculosis*

Bioinformatic search for ClgR binding sites. We used MEME (Bailey & Elkan, 1994) to calculate the combined consensus sequence for the ClgR operator in *C. glutamicum*, *S. lividans* and *Bif. breve* and used the MAST software to search for matching patterns throughout the *M. tuberculosis* genome. Thirty-six matching patterns were found in intergenic regions and, as previously observed in *S. lividans* and *C. glutamicum*, putative ClgR binding sites were identified in the *clpP1*/*clpP2* and *clpC* promoter regions (Table 1 shows ClgR binding sites corroborated by other data in this paper and in existing literature). Unlike in the other actinomycetes studied, the *M. tuberculosis* gene encoding the ClpB ATPase was also preceded by a putative ClgR operator. As in *C. glutamicum* (Engels *et al.*, 2005) a ClgR operator was observed in the upstream region of the PtrB protease gene. Similar to Streptomyces (Bellier & Mazodier, 2004) a ClgR binding site is also present upstream of *clgR* itself.

Transcriptional profile of Δ*clgR M. tuberculosis*. To investigate if the other MAST-predicted ClgR operators were functional binding motifs, we examined the transcriptional profile of a *clgR* mutant strain of *M. tuberculosis*. The Δ*clgR* strain was generated by gene replacement and verified by Southern blot analysis (Supplementary Fig. S1). Whole-genome transcription profiles of the wild-type strain and Δ*clgR* were generated by microarray analysis of RNAs extracted from late-exponential-phase broth cultures. Transcription data are available at BμG@Sbase (accession number E-BUGS-96; http://bugs.sgul.ac.uk/E-BUGS-96) and also ArrayExpress (accession number E-BUGS-96) and are depicted graphically in Fig. 1(a). The wild-type and mutant transcriptomes were normalized and compared by rank product analysis. This revealed considerable differences in gene expression between the strains. In other actinomycetes, ClgR acts as a transcriptional activator and thus we expected genes under its control to have lower levels of transcription in the Δ*clgR* strain. Twenty-six genes were downregulated in the mutant at a percentage of false positives (pfp) value of 0.1 (Supplementary Table S1). Cross-referencing of the MAST-predicted operator regions

Table 1. ClgR binding sites in *M. tuberculosis*

Transcription unit	Gene products/function	ClgR binding site (5′–3′)	Position (bp) upstream of ATG
Rv0251c–Rv0249c	Acr2 chaperone, oxidoreductase, succinate dehydrogenase	ACAGTGTTCGCTTCCGGTGAACGCCGAG	96
Rv0384c*	ClpB chaperone	CGATTCTAAGCTCGCGGCGAAACGGCGG	99
Rv0782	PtrBb protease	GACTTCTTCGCACTGGGCGCGGCCAGCG	46
Rv1043c	Protease	TCGGTGTTCGCTTCACGCGAACTAGGCG	156
Rv2461c–Rv2460c	ClpP1 and ClpP2 proteases	CCGTATGACGCTGTAAGCGAACGCGCCG	89
Rv2745c–Rv2744c	ClgR regulator, 35 kDa protein	AACCCCACCGCCGGCGGCGTTCACCTGA	32
Rv3269	Chaperone	TCCTGTGACGCAGGCCGCGTCCAGCAGG	105
Rv3596c*	ClpC1 protease-associated chaperone	CATCGGTTCGCCGCCAGCGAACGCGGCA	203

*Microarray analysis of transcription indicated all promoters except *clpB* (Rv0384c) and *clpC1* (Rv3596c) to be downregulated in an *M. tuberculosis* Δ*clgR* strain.

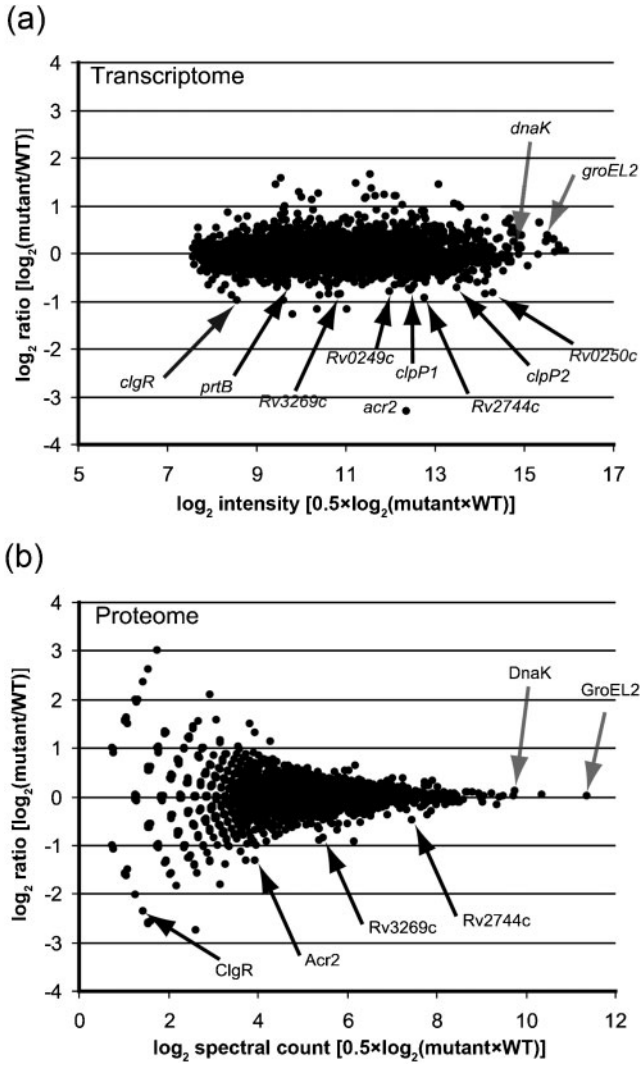

Fig. 1. Gene expression profile of *M. tuberculosis* lacking the ClgR transcriptional activator. (a) Scatter plot showing transcriptional profile derived by whole-genome microarray: log Δ*clgR*/wild-type signal ratios are plotted against log total signal intensity, where log ratios are centralized such that mean log mutant and wild-type signals are equal to zero. (b) Scatter plot showing protein profiles derived by LC/MS-MS: log Δ*clgR*/wild-type ratios of protein spectral counts are plotted against log total spectral count, where log ratios are centralized such that mean log mutant and wild-type spectral counts are equal to zero.

with the Δ*clgR* transcriptional profile revealed that four of these genes, Rv1043c, Rv2744c, *ptrBb* and *clpP2*, were downstream of predicted ClgR operators (Table 1), thus confirming the functionality of the predicted operators in these genes. *clpP2* is transcribed bicistronically with a paralogous protease gene, *clpP1*, which also had lower expression in the Δ*clgR* mutant compared to the wild-type (0.63-fold, pfp 0.13). These data are consistent with qRT-PCR data on *clpP1/2* and *clpC1* in *M. tuberculosis* over-expressing ClgR (Mehra & Kaushal, 2009) but also indicate a wider role for ClgR as a regulator of multiple protease systems, including PtrB and Rv1043c, which encodes a serine protease with similarity to the HtrA family of dual chaperone proteases. Rv2744c encodes the 35 kDa antigen (O'Connor *et al.*, 1990) and is the second gene of a predicted operon with *clgR* (Roback *et al.*, 2007). Downregulation of Rv2744c may be due to the polar effect of the gene replacement of *clgR*, although the reversal of this downregulation in the complemented mutant strain suggests otherwise (Supplementary Table S2).

We used the sequences of the mycobacterial ClgR operators in Rv1043c, Rv2744/45c, *ptrBb* and *clpP1/clpP2* to rescreen for ClgR binding sites in the intergenic regions of downregulated genes in the Δ*clgR* mutant. Imperfect palindromic motifs with similarity to the *clgR* operator (Table 1) were identified in the chaperone gene, *acr2*, and upstream of Rv3269, which encodes a protein with 28 % identity to *M. tuberculosis* GroEL2 and approximately 40 % identity to short GroEL family proteins from many other mycobacteria, such as *M. vaccae and M. gadium*. The predicted chaperone function of these genes is also consistent with ClgR acting as a regulator of systems involved in protein homeostasis/degradation. Rv3269 is predicted to be transcribed bicistronically with *ctpC*, encoding a metal cation-transporting p-type ATPase. However, *ctpC* was not significantly downregulated in the Δ*clgR* mutant (0.85-fold change, $P=0.20$).

The combined results of motif searches for ClgR operators and microarray analysis suggest that at least ten genes are directly regulated by ClgR in *M. tuberculosis* (Table 1). It is of note that neither *clpC1* nor *clpB* was demonstrated to be differentially regulated in the Δ*clgR* mutant, despite the prediction of ClgR binding sites in upstream sequences. However, another study showed that *clpC1* was upregulated by overexpression of ClgR (Mehra & Kaushal, 2009) and there is considerable evidence from other bacteria that Clp ATPases are regulated by ClgR.

Microarray analysis of *M. tuberculosis* Δ*clgR* that was complemented by reintroduction of the *clgR* gene to the attB phage integration site showed that *clgR* transcript levels were restored to wild-type levels and the transcript levels of all ten ClgR-regulated genes except Rv3269 were restored towards wild-type levels (Supplementary Table S2). Validation of a sample of the microarray results was provided by real-time RT-PCR analysis of the transcript levels of *acr2*, *clpP1* and Rv3596c using eight biological replicates. The results are shown in Supplementary Fig. S2 and reveal a good correlation between microarray and RT-PCR-derived expression ratios (Pearson correlation coefficient=1.0)

Interaction of GST-ClgR with the *acr2* promoter region. The putative binding motif upstream of *acr2* (Table 1, central region CGC-N_5-GTG) did not conserve the central CGC-N_5-GCG palindromic motif of the operators upstream of other downregulated genes and known regulated genes in other actinomycetes. This prompted the experimental

validation of ClgR binding to the *acr2* upstream sequence. We were unable to purify recombinant *M. tuberculosis* ClgR as an N-terminal His-tagged recombinant protein in *E. coli* but were able to express and purify a C-terminal glutathione *S*-transferase fusion protein (GST-ClgR). We used electrophoretic mobility shift assay (EMSA) to demonstrate that GST-ClgR specifically interacted with a 40 bp oligonucleotide corresponding to the putative ClgR binding motif (Fig. 2). The percentage of the labelled oligonucleotide shifted was low but the binding was highly specific as GST-ClgR did not interact with a control oligonucleotide in which the sequence had been scrambled, and binding could be competitively inhibited by an excess of unlabelled oligonucleotides. *acr2* encodes a chaperone (Stewart *et al.*, 2005) and is the lead gene of an apparent operon (Roback *et al.*, 2007) that also includes a gene of unknown function (Rv0250c) and Rv0249c, predicted to encode a subunit of succinate dehydrogenase. All three genes in the *acr2* operon were downregulated in the Δ*clgR* strain (Supplementary Table S1).

MEME was used to generate a consensus sequence for the *M. tuberculosis* ClgR operator using genes that showed evidence of downregulation in the Δ*clgR* strain and including the *clpC1* but excluding the *clpB* motifs (Fig. 2b).

Proteomic profile of Δ*clgR M. tuberculosis*. To determine if the transcriptional changes observed in the Δ*clgR* strain were reflected in the proteome of the bacterium, we performed LC-MS/MS on protein extracts of wild-type and Δ*clgR* bacteria grown in triplicate to late-exponential phase in broth culture. More than 2000 proteins were identified; the comparison of mean spectral counts between the mutant and wild-type is depicted graphically in Fig. 1b and tabulated in the Supplementary Table S3. Of the ten genes identified by microarray and motif searches as being regulated by ClgR, the protein products of six were identified by LC-MS/MS but only the 35 kDa antigen (Rv2744c) encoded downstream of *clgR* and the two chaperones, Acr2 and Rv3269, were less abundant in the Δ*clgR* mutant (mutant/wild-type fold change 0.71, 0.40 and 0.55 respectively). Other chaperones such as DnaK, which is co-regulated with Acr2 by the HspR repressor, and GroEL2 were not expressed differently between the strains at either the transcriptional or the protein level (Fig. 1a, b). One reason for the only partial correlation between transcription and protein levels is that protein abundance is affected by many parameters other than transcript abundance, e.g. rates of protein translation and degradation. In addition, the sensitivity of the protein profiling technique employed here does not match that of the transcriptional profiling, and thus the moderate differences in transcription observed may result in changes in protein abundance below the sensitivity limits of our assay.

ClgR regulation is essential for macrophage infection

Survival of Δ*clgR* in macrophages. In broth culture, wild-type *M. tuberculosis*, Δ*clgR* and Δ*clgR* complemented with a single copy of the *clgR* gene grew at similar rates to similar final cell density (Fig. 3a). To test the importance of ClgR during intracellular growth in host cells we infected murine bone-marrow-derived macrophages with wild-type, Δ*clgR* and Δ*clgR* complemented strains at an m.o.i. of 1. All three bacterial strains were considerably depleted during the first 48 h of infection but the Δ*clgR* strain survived

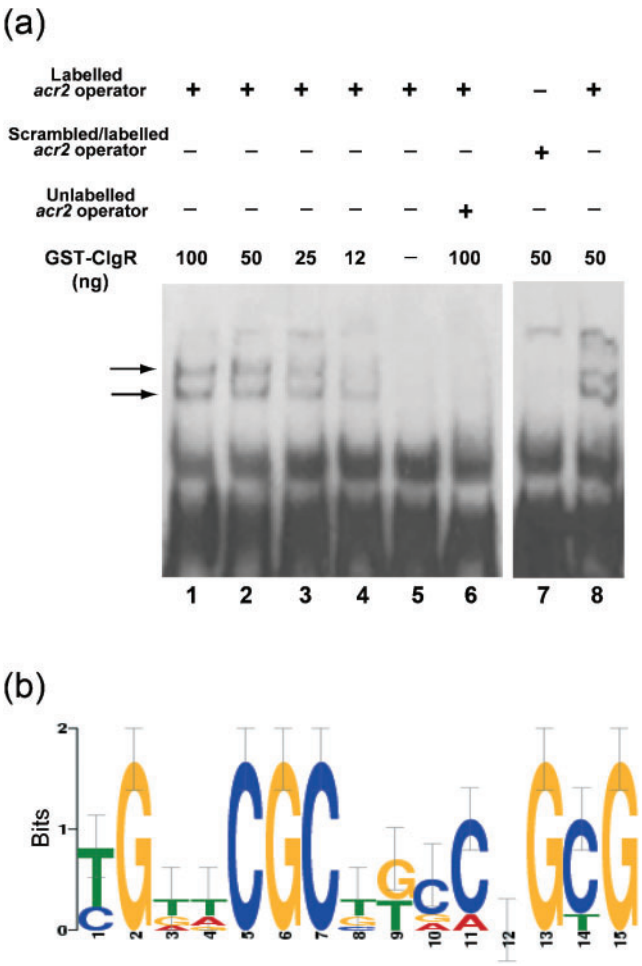

Fig. 2. ClgR–DNA interaction. (a) GST-ClgR binds the palindromic motif in the Acr2 promoter region. EMSA detection was done using a fixed amount of DIG-labelled Rv0251c probe incubated with reaction mixtures containing 100 ng GST-Rv2745c purified protein (lane 1); 50, 25 and 12.5 ng GST-Rv2745c (lanes 2, 3, and 4 respectively); DIG-labelled Rv0251c probe only (lane 5); 100 ng GST-Rv2745c with DIG-labelled Rv0251c probe and 100-fold excess of unlabelled probe (lane 6); 50 ng GST-Rv2745c with a scrambled DIG-labelled Rv0251c probe (lane 7); and 50 ng GST-Rv2745c and DIG-labelled Rv0251c probe (lane 8). (b) The consensus binding site of ClgR demonstrated as a sequence logo to show the relative frequency of each base at each sequence position. All binding sites identified by a combination of MEME and microarray (Table 1) were used to generate the logo. The error bars indicate an approximate 95 % confidence interval.

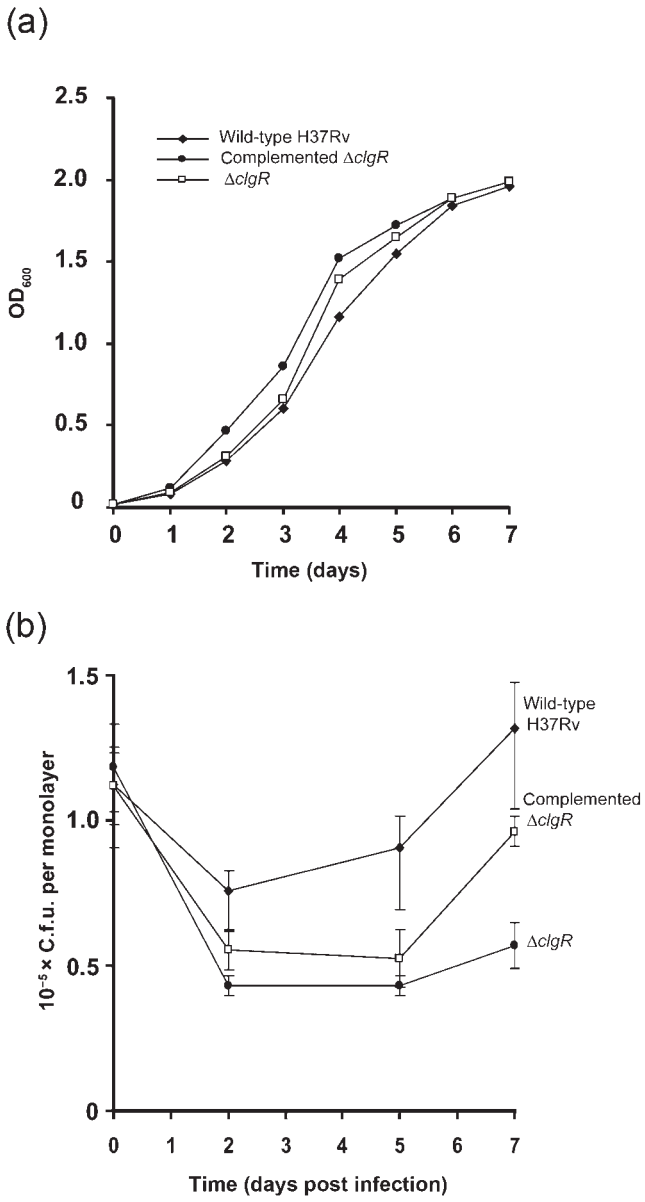

Fig. 3. Reduced growth of *M. tuberculosis* lacking the ClgR transcriptional activator in murine bone-marrow-derived macrophages. (a) In Middlebrook 7H9 broth, *M. tuberculosis* Δ*clgR* grows at a similar rate and to a similar density as wild-type *M. tuberculosis* and *M. tuberculosis* Δ*clgR* complemented with a single copy of *clgR* integrated into the chromosomal *attB* site. (b) *M. tuberculosis* Δ*clgR* has a reduced capacity to replicate in murine bone-marrow-derived macrophages compared to wild-type *M. tuberculosis* and *M. tuberculosis* complemented Δ*clgR*. Results shown are representative of two independent experiments. Error bars represent SE.

significantly less well than the wild-type ($P<0.01$) (Fig. 3b). By 7 days post-infection both the wild-type *M. tuberculosis* and the complemented Δ*clgR* strain showed significant intracellular growth but the Δ*clgR* mutant strain did not ($P<0.01$ Δ*clgR* versus wild-type, $P<0.01$ Δ*clgR* versus complemented Δ*clgR*).

Acidification of the Δ*clgR* phagosome. We also examined the maturation of the mycobacterial phagosomes at 1 h post-infection using the acidotropic dye LysoTracker Red DND99. The Δ*clgR* mutant strain demonstrated a significantly reduced ability to arrest the acidification of its phagosome compared to wild-type *M. tuberculosis* ($P<0.01$) (Fig. 4). The mutant phenotype was reversed back to wild-type by the reintroduction of the *clgR* gene into the complemented strain. Although the reduced ability to control phagosome maturation was only modest, it was highly consistent. The results shown represent the means of three fully independent and blinded experiments performed with independent cultures at different times and involving scoring of approximately 250 phagosomes for each strain. There was no difference in the viability of broth cultures of the three strains.

DISCUSSION

When a bacterium is phagocytosed by a macrophage it is exposed to a radically different and dynamic environment in the phagosome. Phagosomal conditions include controlled availability of nutrients, production of reactive oxygen by the phagosome oxidase, acidification by the vacuolar ATPase proton pump and accumulation of a variety of hydrolytic enzymes that effect the degradation of the phagosomal contents. *M. tuberculosis* is one of a small group of micro-organisms that have evolved to replicate in the phagosome. Studies of bacterial gene transcription provide a window to the extent of cellular remodelling that occurs upon internalization of *M. tuberculosis*. More than 200 genes have altered expression immediately following

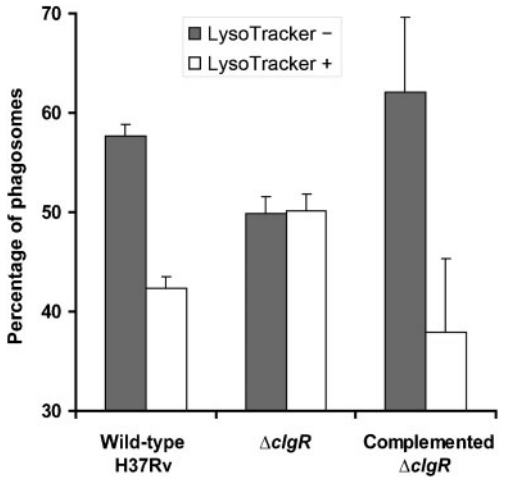

Fig. 4. Mycobacterial control of phagosome acidification is impaired by deletion of *clgR*. The graph shows the frequency of colocalization of mycobacterial phagosomes in murine bone-marrow-derived macrophages with the aciditrophic dye LysoTracker Red DND99. Results shown are the means ± SEM of three independent experiments.

phagocytosis and a further 400 transcriptional changes occur during the next 24 h of infection (Schnappinger *et al.*, 2003). However, changes in gene expression are only one way to alter the protein component of a cell. Equally important is the effect of proteases that sculpt the proteome by degradation of proteins. Regulation of proteases is vitally important to avoid the uncoordinated destruction of cellular proteins; this occurs at the level of substrate selection and also by controlling the expression of proteases via transcriptional regulators. Substrate selection involves the recognition of specific sequences or tags, often involving a chaperone, which unfolds the protein and delivers it to the proteolytic active sites sequestered within a compartmentalized multi-subunit protease. In the present report we describe a genome-wide analysis of the ClgR regulation of protein homeostatic systems in *M. tuberculosis*. We investigate the action of the ClgR transcriptional activator, defining the regulatory circuit that controls the expression of several protease and chaperone systems and the importance of this during infection of the macrophage.

In several actinobacteria, ClgR is the primary regulator for transcription of the compartmental ATP-dependent protease sytems, Clp and Lon (Bellier & Mazodier, 2004; Engels *et al.*, 2004, 2005; Ventura *et al.*, 2005). Recent studies have also indicated that ClgR controls Clp proteases in mycobacteria (Barik *et al.*, 2010; Mehra & Kaushal, 2009). To further our knowledge of the action of ClgR in *M. tuberculosis*, we generated a consensus ClgR operator sequence using experimentally determined ClgR binding sites from *S. lividans*, *C. glutamicum* and *Bif. breve*. We searched the *M. tuberculosis* genome to identify possible ClgR operators and then cross-referenced this dataset with the transcriptional profile of an *M. tuberculosis* strain deleted of ClgR. This identified three proteases, ClpP1P2, PtrB and Rv1043c, to be ClgR-regulated. ClpP proteases function as multimers consisting of two heptameric rings which act in concert with a hexameric assembly of Clp ATPase subunits (ClpC or ClpX in *M. tuberculosis*) that select proteins for unfolding and translocation to the ClpP multimer. In *E. coli*, ClpP degrades a range of proteins selected on the basis of C-terminal recognition sequences or an ssrA tag (Flynn *et al.*, 2003; Gottesman *et al.*, 1997). In *B. subtilis*, ClpP targets an even greater range of proteins and represents the major agent of protein turnover in the cell, although in this case specific recognition motifs have not been identified and the mechanism of substrate selection remains to be determined (Gerth *et al.*, 2008; Kock *et al.*, 2004). The targets of ClpP proteolysis in *M. tuberculosis* are not well defined, although ClpP has been shown to degrade the phosphorylated SigE-specific anti-sigma factor RseA (Barik *et al.*, 2010). In contrast to ClpP, the other *M. tuberculosis* proteases controlled by ClgR are ATP-independent. Rv1043c encodes a protease of the HtrA family of multimeric serine proteases, which typically have dual function, swapping between chaperone and proteolytic modes depending on stress conditions (Iwanczyk *et al.*, 2007; Meltzer *et al.*, 2008; Spiess *et al.*, 1999), and are usually associated with the cell wall. PtrBb encodes the second part of the PtrB oligopeptidase B which cleaves peptide bonds on the C-terminal side of lysyl and argininyl residues. The structure and regulation of *ptrB* in H37Rv appear to be different to all other members of the *M. tuberculosis* complex. In *M. tuberculosis* CDC1551 or *Mycobacterium bovis*, for example, there exists a single *ptrB* gene that in H37Rv has been split in two by a frame-shift mutation, giving *ptrBa* and *ptrBb*. The predicted ClgR binding site upstream of H37Rv *ptrBb* lies within the ORF of the upstream *ptrBa* and, in CDC1551 and *M. bovis*, within the ORF of the single *ptrB*. We have validated the binding of ClgR to this sequence by EMSA (data not shown). Further investigation will be needed to determine if the binding of ClgR to this intra-ORF sequence in CDC1551 and *M. bovis* has a regulatory function. Certainly, we show that in H37Rv, only *ptrBb* and not *ptrBa* is regulated by ClgR.

The Clp ATPase chaperones ClpC and ClpB also contain good matches for the ClgR binding site but were not downregulated in our transcriptional analyses of the *clgR* mutant. It is possible that ClgR is not active at these sites under the conditions used in the transcriptional assay. Consistent with this, ClpC1 was upregulated in *M. tuberculosis* that overexpressed ClgR (Mehra & Kaushal, 2009). Indeed, the transcriptional and proteomic comparison of wild-type and mutant strains under various stress conditions may reveal ClgR regulation of these and, potentially, other genes.

Two other chaperone genes, encoding chaperones Acr2 and Rv3269, were identified as being regulated by ClgR. Each of these contained a ClgR binding motif, and both mRNA transcripts and proteins were present at lower levels in the *clgR* mutant compared to the wild-type. Acr2 is a member of the α-crystallin family of molecular chaperones that are widespread through cellular life and function primarily to stabilize proteins during stress conditions (Van Montfort *et al.*, 2001). The expression of the *M. tuberculosis acr2* gene increases more than any other gene during heat shock (Stewart *et al.*, 2002) or following macrophage infection (Schnappinger *et al.*, 2003) and it is required for pathogenesis in a murine model of tuberculosis (Stewart *et al.*, 2005). The *acr2* gene (also referred to as *hsp*) is regulated by the alternative sigma factors SigE (Manganelli *et al.*, 2001) and SigH (Manganelli *et al.*, 2002), by the MprAB two-component system (Pang & Howard, 2007) and by the HspR transcriptional repressor (Stewart *et al.*, 2002). The discovery here that *acr2* expression is controlled by ClgR adds a further layer of control over this gene and is indicative of the importance to the bacterium of controlling expression of Acr2, which is both a dominant antigen (Wilkinson *et al.*, 2005) and a virulence factor (Stewart *et al.*, 2005).

Our study also indicates that, as in *Streptomyces*, ClgR regulates transcription of its own operon in a positive

feedback loop. In other bacteria, control systems regulating the expression of proteolysis also feature a post-translational autoregulation involving the proteolytic degradation of the transcriptional regulator. The *B. subtilis* CtsR and *E. coli* σ^{32} are both degraded by proteases under their transcriptional control (Arsène *et al.*, 2000; Krüger *et al.*, 2001). In Streptomyces, ClgR is itself degraded by ClpP and is recognized via two C-terminal alanines, in line with the ssrA tag of *E. coli* (Bellier *et al.*, 2006). The *Mycobacterium* ClgR C-terminal residues Ala-Val-Ala-COOH are also consistent with the *E. coli* Clp recognition motif (Flynn *et al.*, 2003), although experimental demonstration of Clp protease substrate selection in mycobacteria has yet to be performed.

Transcription of ClgR, and members of its regulon, is induced following phagocytosis of *M. tuberculosis* by macrophages (Schnappinger *et al.*, 2003). To test if this response reflects an important functional role during intracellular parasitism, we compared the intracellular growth of wild-type *M. tuberculosis* with the *clgR* deletion mutant, which would be unable to activate expression of the regulon. Both strains were significantly depleted during the first 24 h of infection but the *clgR* strain fared less well than the wild-type. Although the wild-type strain was then able to produce intracellular growth, the *clgR* mutant could not. By examining the acidification of the mycobacterial phagosome we were able to see that even at 1 h post-infection, the *clgR* mutant was at a disadvantage and was less able to inhibit phagosome maturation. This attenuation could be explained if the rapid upregulation of the ClgR-dependent proteases and chaperones is necessary for protein homeostasis during exposure to the macrophage antimicrobial mechanisms, such as the oxidative burst. Consistent with this, ClgR is regulated by SigE, which is necessary for resistance to oxidative stress and intramacrophage growth (Manganelli *et al.*, 2001). However, the relationship between ClgR and SigE is more complicated, as the function of SigE is itself positively regulated by ClgR via the ClgR-dependent activation of ClpP and subsequent ClpP degradation of the SigE-specific antisigma factor RseA (Barik *et al.*, 2010). Alternatively, or additionally, the induction of proteases and chaperones by ClgR may directly function in the remodelling of the proteome during adaptation to the intracellular environment. In an analogous manner, *B. subtilis* uses ClpP to directly sculpt the proteome, degrading enzymes of central metabolism during the transition to anaerobic growth or carbon-limited conditions (Gerth *et al.*, 2008). It is probable that proteolysis is a key element in the adaptation to any new environment. Examination of bacterial protein abundance in protease or protease regulator mutants during macrophage infection will be necessary to assess the direct role of proteases during the intracellular adaptation of *M. tuberculosis*.

The present study provides evidence for the ClgR-dependent regulation of at least 10 genes, of which the majority are involved in protein stabilization, disassembly or degradation. We demonstrate that the induction of these genes is an essential element of the successful parasitism of the macrophage host cell. The potential of protease systems as effective antimicrobial targets has been previously demonstrated (Brötz-Oesterhelt *et al.*, 2005). These studies showed that acyldepsipeptides act on ClpP in Gram-positive bacteria to effect the dysregulation of proteolysis and cell death. Thus, understanding the regulation of proteolysis in *M. tuberculosis* and other pathogens is an important topic. Further dissection of the ClgR regulon and definition of the contribution of its constituent parts to virulence may be useful during the development of new antibiotics that target proteolytic control systems.

ACKNOWLEDGEMENTS

This work was supported by the European Community (TB-MACS, grant no. LSHP-CT-2006-037732) and by a University of Surrey studentship, awarded to M. E. We thank the BμG@S at St Georges Hospital Medical School for the supply of and assistance with microarrays.

REFERENCES

Arsène, F., Tomoyasu, T. & Bukau, B. (2000). The heat shock response of *Escherichia coli*. *Int J Food Microbiol* **55**, 3–9.

Bailey, T. L. & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In Second International Conference on Intelligent Systems for Molecular Biology, pp. 28–36. Menlo Park, CA: AAAI Press.

Bailey, T. L. & Gribskov, M. (1998). Combining evidence using *p*-values: application to sequence homology searches. *Bioinformatics* **14**, 48–54.

Barik, S., Sureka, K., Mukherjee, P., Basu, J. & Kundu, M. (2010). RseA, the SigE specific anti-sigma factor of *Mycobacterium tuberculosis*, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol Microbiol* **75**, 592–606.

Bellier, A. & Mazodier, P. (2004). ClgR, a novel regulator of *clp* and *lon* expression in *Streptomyces*. *J Bacteriol* **186**, 3238–3248.

Bellier, A., Gominet, M. & Mazodier, P. (2006). Post-translational control of the *Streptomyces lividans* ClgR regulon by ClpP. *Microbiology* **152**, 1021–1027.

Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* **573**, 83–92.

Brötz-Oesterhelt, H., Beyer, D., Kroll, H. P., Endermann, R., Ladel, C., Schroeder, W., Hinzen, B., Raddatz, S., Paulsen, H. & other authors (2005). Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082–1087.

Butler, S. M., Festa, R. A., Pearce, M. J. & Darwin, K. H. (2006). Self-compartmentalized bacterial proteases and pathogenesis. *Mol Microbiol* **60**, 553–562.

Darwin, K. H. (2009). Prokaryotic ubiquitin-like protein (Pup), proteasomes and pathogenesis. *Nat Rev Microbiol* **7**, 485–491.

Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N. & Nathan, C. F. (2003). The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* **302**, 1963–1966.

Dussurget, O., Stewart, G., Neyrolles, O., Pescher, P., Young, D. & Marchal, G. (2001). Role of *Mycobacterium tuberculosis* copper-zinc superoxide dismutase. *Infect Immun* **69**, 529–533.

Engels, S., Schweitzer, J. E., Ludwig, C., Bott, M. & Schaffer, S. (2004). *clpC* and *clpP1P2* gene expression in *Corynebacterium glutamicum* is controlled by a regulatory network involving the transcriptional regulators ClgR and HspR as well as the ECF sigma factor σ^{H}. *Mol Microbiol* **52**, 285–302.

Engels, S., Ludwig, C., Schweitzer, J. E., Mack, C., Bott, M. & Schaffer, S. (2005). The transcriptional activator ClgR controls transcription of genes involved in proteolysis and DNA repair in *Corynebacterium glutamicum*. *Mol Microbiol* **57**, 576–591.

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T. & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol Cell* **11**, 671–683.

Frees, D., Chastanet, A., Qazi, S., Sorensen, K., Hill, P., Msadek, T. & Ingmer, H. (2004). Clp ATPases are required for stress tolerance, intracellular replication and biofilm formation in *Staphylococcus aureus*. *Mol Microbiol* **54**, 1445–1462.

Frees, D., Savijoki, K., Varmanen, P. & Ingmer, H. (2007). Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol Microbiol* **63**, 1285–1295.

Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. (2000). The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol Microbiol* **35**, 1286–1294.

Gerth, U., Kock, H., Kusters, I., Michalik, S., Switzer, R. L. & Hecker, M. (2008). Clp-dependent proteolysis down-regulates central metabolic pathways in glucose-starved *Bacillus subtilis*. *J Bacteriol* **190**, 321–331.

Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol* **19**, 565–587.

Gottesman, S., Wickner, S. & Maurizi, M. R. (1997). Protein quality control: triage by chaperones and proteases. *Genes Dev* **11**, 815–823.

Ibrahim, Y. M., Kerr, A. R., Silva, N. A. & Mitchell, T. J. (2005). Contribution of the ATP-dependent protease ClpCP to the autolysis and virulence of *Streptococcus pneumoniae*. *Infect Immun* **73**, 730–740.

Iwanczyk, J., Damjanovic, D., Kooistra, J., Leong, V., Jomaa, A., Ghirlando, R. & Ortega, J. (2007). Role of the PDZ domains in *Escherichia coli* DegP protein. *J Bacteriol* **189**, 3176–3186.

Kajfasz, J. K., Martinez, A. R., Rivera-Ramos, I., Abranches, J., Koo, H., Quivey, R. G., Jr & Lemos, J. A. (2009). Role of Clp proteins in expression of virulence properties of *Streptococcus mutans*. *J Bacteriol* **191**, 2060–2068.

Kock, H., Gerth, U. & Hecker, M. (2004). The ClpP peptidase is the major determinant of bulk protein turnover in *Bacillus subtilis*. *J Bacteriol* **186**, 5856–5864.

Krüger, E., Zuhlke, D., Witt, E., Ludwig, H. & Hecker, M. (2001). Clp-mediated proteolysis in Gram-positive bacteria is autoregulated by the stability of a repressor. *EMBO J* **20**, 852–863.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K. & Smith, I. (2001). The *Mycobacterium tuberculosis* ECF sigma factor σ^{E}: role in global gene expression and survival in macrophages. *Mol Microbiol* **41**, 423–437.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K., Dubnau, E., Gomez, M. & Smith, I. (2002). Role of the extracytoplasmic-function sigma factor σ^{H} (H) in *Mycobacterium tuberculosis* global gene expression. *Mol Microbiol* **45**, 365–374.

Mehra, S. & Kaushal, D. (2009). Functional genomics reveals extended roles of the *Mycobacterium tuberculosis* stress response factor σ^{H}. *J Bacteriol* **191**, 3965–3980.

Meltzer, M., Hasenbein, S., Hauske, P. & other authors (2008). Allosteric activation of HtrA protease DegP by stress signals during bacterial protein quality control. *Angew Chem Int Ed Engl* **47**, 1332–1334.

O'Connor, S. P., Rumschlag, H. S. & Mayer, L. W. (1990). Nucleotide sequence of the gene encoding the 35-kDa protein of *Mycobacterium tuberculosis*. *Res Microbiol* **141**, 407–423.

Pang, X. & Howard, S. T. (2007). Regulation of the alpha-crystallin gene *acr2* by the MprAB two-component system of *Mycobacterium tuberculosis*. *J Bacteriol* **189**, 6213–6221.

Ribeiro-Guimarães, M. L. & Pessolani, M. C. (2007). Comparative genomics of mycobacterial proteases. *Microb Pathog* **43**, 173–178.

Roback, P., Beard, J., Baumann, D., Gille, C., Henry, K., Krohn, S., Wiste, H., Voskuil, M. I., Rainville, C. & Rutherford, R. (2007). A predicted operon map for *Mycobacterium tuberculosis*. *Nucleic Acids Res* **35**, 5085–5095.

Russell, D. G. (2001). *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nat Rev Mol Cell Biol* **2**, 569–577.

Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2001). Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* **98**, 12712–12717.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D. & other authors (2003). Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J Exp Med* **198**, 693–704.

Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem* **68**, 850–858.

Smyth, G. K. (2005). Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, pp. 397–420. Edited by R. Gentleman, V. Carey, W. Huber, R. Irizarry & S. Dudoit. New York: Springer.

Smyth, G. K., Yang, Y. H. & Speed, T. (2003). Statistical issues in cDNA microarray data analysis. *Methods Mol Biol* **224**, 111–136.

Spiess, C., Beil, A. & Ehrmann, M. (1999). A temperature-dependent switch from chaperone to protease in a widely conserved heat-shock protein. *Cell* **97**, 339–347.

Stewart, G. R., Snewin, V. A., Walzl, G., Hussell, T., Tormay, P., O'Gaora, P., Goyal, M., Betts, J., Brown, I. N. & Young, D. B. (2001). Overexpression of heat-shock proteins reduces survival of *Mycobacterium tuberculosis* in the chronic phase of infection. *Nat Med* **7**, 732–737.

Stewart, G. R., Wernisch, L., Stabler, R., Mangan, J. A., Hinds, J., Laing, K. G., Young, D. B. & Butcher, P. D. (2002). Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* **148**, 3129–3138.

Stewart, G. R., Newton, S. M., Wilkinson, K. A., Humphreys, I. R., Murphy, H. N., Robertson, B. D., Wilkinson, R. J. & Young, D. B. (2005). The stress-responsive chaperone alpha-crystallin 2 is required for pathogenesis of *Mycobacterium tuberculosis*. *Mol Microbiol* **55**, 1127–1137.

Tailleux, L., Waddell, S. J., Pelizzola, M., Mortellaro, A., Withers, M., Tanne, A., Castagnoli, P. R., Gicquel, B., Stoker, N. G. & other authors (2008). Probing host pathogen cross-talk by transcriptional profiling of both *Mycobacterium tuberculosis* and infected human dendritic cells and macrophages. *PLoS ONE* **3**, e1403.

Van Montfort, R., Slingsby, C. & Vierling, E. (2001). Structure and function of the small heat-shock protein/alpha-crystallin family of molecular chaperones. *Adv Protein Chem* **59**, 105–156.

Ventura, M., Zhang, Z., Cronin, M., Canchaya, C., Kenny, J. G., Fitzgerald, G. F. & van Sinderen, D. (2005). The ClgR protein

regulates transcription of the *clpP* operon in *Bifidobacterium breve* UCC 2003. *J Bacteriol* **187**, 8411–8426.

Wilkinson, K. A., Stewart, G. R., Newton, S. M., Vordermeier, H. M., Wain, J. R., Murphy, H. N., Horner, K., Young, D. B. & Wilkinson, R. J. (2005). Infection biology of a novel alpha-crystallin of *Mycobacterium tuberculosis*: Acr2. *J Immunol* **174**, 4237–4243.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J. & Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**, e15.

Edited by: S. V. Gordon