Published in final edited form as:

Gene. 2012 May 25; 500(1): 85–92. doi:10.1016/j.gene.2012.03.044.

A screen for non-coding RNA in *Mycobacterium tuberculosis* reveals a cAMP-responsive RNA that is expressed during infection

Shaaretha Pelly, William R. Bishai, and Gyanu Lamichhane*

Johns Hopkins University School of Medicine, Center for Tuberculosis Research, 1550 Orleans Street, Baltimore, MD 21287

Abstract

A key to the success of *Mycobacterium tuberculosis* (*Mtb*) is the bacteria's ability to survive and thrive in the presence of numerous stresses mounted by the host. Small, non-coding RNAs (sRNAs) have been shown to modulate numerous stress responses in bacteria, yet to date only two studies have screened the *Mtb* transcriptome to identify sRNA. Their association with oxidative and acid stress has been demonstrated but the cellular function and role of these sRNAs in the pathogenesis of tuberculosis (TB) remains unknown. Here, we have identified an sRNA, *ncrMT1302*, in a locus involved in cAMP metabolism and demonstrate that expression of *ncrMT1302* responds to changes in pH and cAMP concentration. The differential expression of *ncrMT1302* observed in wild-type *Mtb* during growth is abolished in a strain lacking *MT1302*, an adenylyl cyclase encoding gene. We report that *ncrMT1302* is expressed in *Mtb* residing in the lungs of mice during an active infection.

Keywords

Small regulatory RNA; acid stress

1. INTRODUCTION

Mycobacterium tuberculosis (*Mtb*), the etiological agent of tuberculosis (TB), is one of the most successful pathogens today as it maintains a reservoir in one-third of the human population and kills close to 1.7 million each year (WHO, 2010). A key to the success of the TB bacillus is its ability to withstand stress and proliferate in the host under adverse conditions. Upon transmission from one host to another, *Mtb* has to survive in intra and extracellular environments, primarily in the lungs (Manabe et al., 2008), under conditions that evolve during the course of the disease. These include the acidic pH within the phagolysosomal compartment of the macrophage (Fisher et al., 2002) and the hypoxic and nutrient poor environment in granulomatous lesions (Dannenberg, 1993). *Mtb* is also capable of tolerating a wide range of antibiotics (Wallis et al., 1999). This resilience is largely due to the organism's ability to sense stress conditions and mount an appropriate

© 2012 Elsevier B.V. All rights reserved.

*Corresponding author: Gyanu Lamichhane, PhD, Assistant Professor, Johns Hopkins University, 1550 Orleans St, Rm 106, Baltimore, MD 21287, Ph: 410-502-8162, Fax: 410-614-8173, lamichhane@jhu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

response. Sigma factors and two-component signal transduction systems are widely known to coordinate differential expression of transcriptional regulators in response to stress in bacteria (Sachdeva et al., 2010). Perhaps due to this, studies of molecular pathogenesis of TB have been confined to classical genetic analysis of protein encoding genes of *Mtb* and their function. The existence of sRNAs in *Mtb* was not known until 2009 (Arnvig and Young, 2009) and therefore their roles have remained poorly understood.

Many sRNAs have been identified in *Escherichia coli* (*E.coli*) and other bacteria (Vogel et al., 2003; Gottesman, 2005). While only a small number of sRNAs have been studied, regulation by sRNAs is now widely known to contribute to stress responses and consequently to virulence (Gottesman et al., 2006). In *E.coli* for example, 6S RNA has been demonstrated to contribute to stationary phase growth by associating with the sigma 70 RNA polymerase holoenzyme (Wassarman and Storz, 2000). These sRNAs often act by base-pairing with specific target mRNAs and altering their translation or stability, while others associate with specific proteins and modulate their function (Waters and Storz, 2009).

Unlike protein encoding open reading frames (ORF), the lack of consensus sequences for the initiation and termination of transcription of sRNAs has meant difficulty in using structured bioinformatic tools to identify them (Vogel and Sharma, 2005). The first sRNAs in *E.coli* were identified by chance while studying neighboring genes and others were identified in screens of multicopy plasmid screens seeking specific phenotypes (Wassarman et al., 1999). Recently however, genome-wide searches have led to the identification of a larger repertoire of sRNAs in *E.coli*. To date, sRNAs in *Mtb* have been identified through high throughput screens: recently Arnvig *et al* reported identification of nine novel sRNAs with sizes ranging between 20 and 70 nucleotides (Arnvig and Young, 2009), and Dichiarà *et al* identified thirty four novel sRNAs in *Mycobacterium bovis* BCG and showed that a subset of these existed in *Mtb* (Dichiara et al., 2010).

As part of our long standing effort to identify essential genetic elements, we report identification of sRNAs of *Mtb*. To date, there has been no report of a role of sRNAs in the pathogenesis of TB. Here, we performed the first high throughput screen of *Mtb* RNA in sizes ranging from 70 to 200 nucleotides to identify novel sRNAs in this pathogen. We isolated an sRNA whose genomic locus is flanked by *MT1302* and *MT1303*. *MT1302* (*Rv1264*) has been reported to be an adenylyl cyclase that is active at low pH (Dittrich et al., 2006) while *MT1303* (*Rv1265*) has been shown to be upregulated in macrophages (Hobson et al., 2002) and induced by cAMP under low oxygen conditions (Gazdik and McDonough, 2005). On studying this sRNA, we find that it is in a region previously found to be bound by a transcription factor, Cmr (cAMP and macrophage regulator), that regulates cAMP induced genes when the bacilli infect macrophages (Gazdik et al., 2009). Next, we show that its expression levels are altered by acidic pH, changes in cAMP concentration and in a mutant strain lacking a functional copy of *MT1302*. This suggests a role for a small, non-coding RNA in *Mtb* in a potential cAMP modulated regulon. In addition, using the mouse model of TB, we demonstrate that this sRNA is expressed during an active *Mtb* infection.

2. MATERIALS AND METHODS

2.1. Strain, growth conditions and stresses

A clinical isolate of *Mtb*, strain CDC1551 (Valway et al., 1998), and a previously generated *MT1302* transposon insertion mutant JHU1264-780 were used. The mutant has a Himar 1 transposon insertion in the TA dinucleotide located at +780 position from the putative translation start site (Lamichhane et al., 2003). Further details of the transposon insertion mutant are available in this database (<http://webhost.nts.jhu.edu/target/>). Strains were grown in Middlebrook 7H9 broth (Difco) supplemented with 0.5% glycerol, 10% oleic acid-

albumin-dextrose-catalase (OADC) and 0.05% Tween 80 under constant shaking at 37°C. For growth phase studies, cultures at optical density ($A_{600\text{nm}}$) of 0.3 were used for early exponential phase, 0.8 for exponential phase, exponential + 2 days for late exponential cultures, late exponential + 1 day for stationary phase and first day of visible cellular clumping for late stationary, clumped cultures. For stresses, 200 ml cultures at an $A_{600\text{nm}}$ of 0.8 were centrifuged, washed twice with phosphate buffered saline (PBS) and resuspended as described below. To induce nitric oxide stress, diethylenetriamine nitric oxide adduct (DETA-NO) was added to the broth at a final concentration of 250 μM . To starve bacilli of nutrients, cells were resuspended in PBS containing 0.05% Tween 80. To induce acid stress, cells were resuspended in 7H9 broth adjusted to a pH of 5.5 with 1M HCl. As a control, a 200 ml culture was resuspended in Middlebrook 7H9 broth. Bacilli were incubated under stress conditions for 6 hours and RNA was extracted with Trizol (Invitrogen) as described (Hatfull and Jacobs, 2000). For exposure to drugs, washed cells were resuspended in Middlebrook 7H9 broth containing 0.5 $\mu\text{g/ml}$ rifampicin or 0.05 $\mu\text{g/ml}$ isoniazid and incubated for 3 hours before extracting RNA. For exposure to cAMP, 110 μM dibutyryl cAMP (Sigma) was added to 200 ml cultures at specific growth phases and total RNA was extracted after 8 hours. Cultures without addition of dibutyryl cAMP were used as a control. All studies were carried out in duplicate.

2.2. PCR and Quantitative Reverse-Transcription PCR

5 μg total RNA was treated with 2U of TURBOTM DNA-free DNase (Applied Biosystems) according to manufacturer's instructions for 30 minutes, followed by addition of 2U of DNase for 30 minutes. All samples were tested for residual DNA by quantitative reverse transcription PCR (qRT-PCR) and confirmed to be free of DNA if no amplified product could be detected past 35 cycles. Reverse transcription was carried out using SuperScriptIII Reverse Transcriptase (Invitrogen) and a gene specific reverse primer. qRT-PCR was performed using SYBR[®] Green PCR Master Mix (Applied Biosystems) in the iCycler (BioRad) and normalized to levels of *sigA*, a housekeeping gene whose expression has been reported to remain constant throughout growth phases (Manganelli et al., 1999). For *ncrMT1302*, Primer 1A: GTGACCAATCGAAATGTCCG and Primer 2A: ACATGGGCAGACCCGCGTGA were used as forward and reverse primers respectively. Similarly, *sigA*-F: CTCGACGCTGAACCAGACCT and *sigA*-R: AGGTCTTCGTGGTCTTCGTC were used as forward and reverse primers respectively for *sigA*. For *MT1302*, Primer *MT1302_F*: GAACATCGACGATCTGTTGG and Primer *MT1302_R*: AATCTCGTCGGGGGTGAT were used. For *MT1303*, Primers *MT1303_F*: CAGGAACCGGTGTCACGTT and *MT1303_R*: CTGGCATCTGGTTGCTGAG were used. The average expression levels, normalized to *sigA*, and standard deviations were calculated using data from three technical replicates of two independent experiments. A *t*-test was used to compare the significance of differences between groups.

2.3. Northern Blotting

Total RNA was separated by electrophoresis on an 8% or 10% Tris-Borate EDTA (TBE)-Urea polyacrylamide gel (Sequagel, National Diagnostics). The gel was stained and viewed under UV with 0.5 $\mu\text{g/ml}$ ethidium bromide to ensure equal loading. The RNA CenturyTM Marker (Applied Biosystems) was used as an RNA ladder to identify 5S rRNA on the polyacrylamide gel. As 5S rRNA is 115 nucleotides in length and our largest *ncrMT1302* clone was 109 nucleotides in length, we used 5S rRNA as a loading control and size marker for the northern blot. Gels were then electroblotted onto Brightstar membranes (Applied Biosystems) and UV crosslinked. Membranes were incubated overnight in ULTRAhybOligo at 54°C with 0.1 $\mu\text{M/ml}$ biotin-tagged probes. Membranes were washed according to manufacturer's instructions (NorthernMax, Applied Biosystems) and the Brightstar Biotin kit (Applied Biosystems) was used for detection.

2.4. Dot Blot

0.5 µg plasmid DNA with 0.4M NaOH and 10mM EDTA was heated at 95°C for 20 minutes and placed on ice. The sample was then dotted onto a Brightstar membrane and dried. The membrane was rinsed in 2X SSC, air dried and UV crosslinked. Prehybridization, hybridization, washing and detection were carried out according to the DIG-High Prime DNA Labeling and Detection Starter Kit II (Roche). 20 ng/ml of DIG-11-dUTP labeled DNA probes for 5S rRNA, 23S rRNA and tRNA-Serine were used.

2.5. RNA Isolation and cloning

Mtb cultures at five growth phases, determined by A_{600nm} , were obtained, centrifuged and the bacterial pellet was resuspended in Trizol (Invitrogen). This mixture was transferred to O-ring tubes containing 0.5 ml of 0.1 mm zirconia beads (BioSpec Products). Cells were incubated at 25°C for 10 minutes and lysed by bead-beating using a mini-beadbeater at 4,800 RPM. Lysed cells were centrifuged for 5 minutes at 13,000 RPM, the supernatant was transferred to a fresh microfuge tube and RNA was then extracted as described (Hatfull and Jacobs, 2000). RNA was resuspended in 30µl of RNase free water and stored at -80°C.

The total RNA preparation (70µg) was electrophoresed on an 8% TBE-urea polyacrylamide gel. RNA ranging from 70 to 200 nucleotides was excised and eluted in 1x TBE and extracted using DTR columns (Edge Biosystems). RNA was then cloned using the miRCatTM microRNA Cloning Kit (IDT) according to manufacturer's instructions. Briefly, a 3' linker was ligated to the gel-extracted RNA using T4 RNA ligase. RNA with the 3' linker was then reverse transcribed using a 5' primer and purified using polyacrylamide gel electrophoresis (PAGE). A second 3' linker with a different sequence from the first linker was then ligated to cDNA using T4 RNA ligase. cDNA with unique linkers at both 5' and 3' ends was PAGE purified, amplified by PCR using primers complementary to both linkers and cloned into the pCR2.1 TOPO vector (Invitrogen) according to manufacturer's instructions. Plasmid DNA was extracted using the Miniprep Kit (Qiagen) and screened for presence of cDNA corresponding to serine-tRNA, 23S rRNA and 5S rRNA using dot-blot analysis. The plasmids that lacked these sequences were chosen and sequenced using the M13 reverse primer. BlastN was used for analyzing the sequences (Altschul et al., 1997). Two DNA linkers with different but known sequences were used to capture the sRNA transcripts. Therefore, the 5' and 3' ends of the sRNAs could be readily identified and the orientation of each sRNA determined based on the unique sequences for the first linker at the 3' end and the second linker at the 5' end. Sequences of the linkers and primers can be found in the miRCatTM microRNA Cloning Kit manual (IDT DNA).

2.6. *Mtb* RNA from mouse infection

Five 4 to 5 weeks old female BALB/C mice were infected with an aerosol of *Mtb* CDC1551 to deliver ~10³ CFU to the lungs. At 3 weeks following infection, mice were sacrificed, lungs obtained and total RNA was prepared using the Trizol (Invitrogen) extraction method as described (Hatfull and Jacobs, 2000). 8µg of total RNA from each sample was used in each lane of the northern blot.

3. RESULTS

3.1. Identification of sRNAs in *Mycobacterium tuberculosis* by cloning

Previous studies to identify sRNAs in mycobacteria have focused on sRNA from *Mtb* in the size range of 20 to 70 nucleotides (Arnvig and Young, 2009) and in the ranges of <80 nucleotides or 100 to 200 nucleotides from non-pathogenic *M. bovis* BCG (Dichiara et al., 2010). We created a cDNA library using RNA isolated from *Mtb* at the late exponential phase of growth in order to identify potential sRNA species in the size range of 70 and 200

nucleotides. We selected this size range based on sRNAs identified in *E.coli* which are predominantly of these sizes (Waters and Storz, 2009). As 5S rRNA falls in this range, we removed the ribosomal RNA with a complementary, single-stranded DNA oligonucleotide followed by treatment with RNaseH. RNA molecules were size fractionated and directly cloned and sequenced using the method outlined in Figure 1. Briefly, a single stranded DNA linker was ligated to the 3' end of the RNA, the samples were then reverse transcribed and a second linker was ligated to the 3' end of the cDNA. The 5'-ligation independent method we used ensures that RNA modified at the 5' end would be captured in our screen. This is of particular importance as little is known about sRNA processing in *Mtb* and sRNA species in some organisms are known to be tri-phosphorylated at the 5' end (Pak and Fire, 2007). cDNA with known linker sequence tags on both ends were PCR amplified and cloned into a TA cloning vector to create a cDNA library.

Initial sequencing of 276 clones revealed that the cDNA library was enriched in serine-tRNA, 23S rRNA and 5S rRNA. We excluded these undesired clones from the library with a dot blot using single stranded DNA probes complementary to serine-tRNA, 23S rRNA and 5S rRNA. Analysis of the sequences revealed 238 clones of tRNA or rRNA, 26 clones represented sequences within known open reading frames (ORFs) and 12 clones represented intergenic regions (Table 1). Sequences representing ORFs (Table 2) could be degraded mRNA or regulatory byproducts of a processed mRNA. While some sRNAs may be encoded by regions within ORFs, their characterization is often more complex as it is difficult to ascertain if they are independent transcriptional units or a regulatory side product of the mRNA encoded by that ORF. Hence these clones were not investigated further. Of the 12 clones representing intergenic sequences (Table 3), 4 were novel sRNAs identified for the first time in this study and 8 contained sequences from a region between *MT1302* and *MT1303* (Figure 2). The presence of homologous sRNA in *M.bovis* BCG was recently reported by DiChiara et. al.(DiChiara et al., 2010). We annotated this sRNA as *ncrMT1302* (non-coding RNA MT1302) hereafter and selected it for further study due to its high abundance in our screen.

3.2. *ncrMT1302* levels vary with phases of growth in vitro

Next we studied the *ncrMT1302* expression profile during the growth cycle of *Mtb* in a nutrient rich defined medium *in vitro*. *Mtb* was grown in Middlebrook 7H9 broth and total RNA was extracted at five physiologically distinct growth phases. A northern blot analysis was performed to verify expression of *ncrMT1302* and the presence of any additional transcripts (Figure 3A). A band representing *ncrMT1302* was detected around the 115 nucleotide 5S rRNA band, suggesting that there is a single functional species of the sRNA of a similar size. The blot also revealed an increase in the expression of *ncrMT1302* from early exponential to stationary phase of growth.

To further assess this observation independently, we performed quantitative reverse transcription PCR (qRT-PCR) to determine the sRNA levels and normalized these levels to *sigA*, a gene that has been reported to be expressed at constant levels throughout all growth phases in culture (Manganelli et al., 1999) (Figure 3B). We observed 200 to 12,000 fold higher levels of *ncrMT1302* transcript relative to that of *sigA*. *ncrMT1302* levels were low during the early exponential phase of growth, elevated in stationary phase but highest in exponential phase and in late stationary phase, clumped cultures. *ncrMT1302* expression was approximately 60 fold higher in clumped cultures compared to the early exponential phase of growth. The high levels of *ncrMT1302* suggest that the sRNA may be vital for adaptation to physiological conditions unique to these growth phases.

3.3. Low pH and rifampicin alter expression of *ncrMT1302*

Bacterial sRNAs have been shown to be involved in regulating response to various stresses (Gottesman et al., 2006). The increased expression of *ncrMT1302* in late stationary phase, clumped cultures, a stage of growth that is defined by stress due to limiting nutrients and increased accumulation of toxic metabolic waste (Smeulders et al., 1999), led us to hypothesize that *ncrMT1302* may be involved in responding to stresses that are prevalent *in vivo* during infection and also during exposure to antibiotics. The presence of increased levels of reactive nitrogen species (Nathan and Shiloh, 2000), limiting nutrition and low pH have been shown to prevail in tissues harboring *Mtb* (Dannenberg, 1993). Therefore, we exposed cultures of *Mtb* at the exponential phase of growth to nitric oxide, low pH and limited nutrition for six hours to assess any functional role of *ncrMT1302* under these conditions. Total RNA was extracted from *Mtb* cultures and *ncrMT1302* expression was assessed by qRT-PCR. We observed an increase and a decrease in sRNA levels under nutrient limitation and exposure to DETA-NO respectively (Figure 4A). The increase in sRNA levels in nutrient limiting conditions is consistent with the growth phase expression study in which we observed elevated levels of *ncrMT1302* during stationary phase and in clumped cultures (Figure 3). The largest difference, a 7 fold decrease in *ncrMT1302* levels, was observed in cultures with reduced pH. This suggests that expression or stability of the sRNA is most strongly associated with a pH stress response.

A number of drugs are used to treat TB. While the method of action of some drugs is understood, the complete mechanism has been elusive for many drugs. sRNA genes, such as *micF*, have been found to be upregulated on exposure to cationic peptide antimicrobials (Oh et al., 2000). Hence we sought to investigate if *ncrMT1302* could be involved in responding to stresses induced during exposure to rifampicin and isoniazid as these are the two drugs that constitute the backbone of TB treatment (WHO, 2010). Total RNA was extracted from *Mtb* cultures exposed to the minimal inhibitory concentrations (MIC) (Lorian, 2005) of either isoniazid or rifampicin for three hours and *ncrMT1302* expression was assessed by northern blot analysis (Figure 4B). RNA isolated from *Mtb* cultured in the absence of these drugs was used as a negative control and 5S rRNA was used as a loading control. Northern blots revealed a diminished level of expression of *ncrMT1302* upon exposure to rifampicin while there was no change on exposure to isoniazid. This was independently confirmed by qRT-PCR (data not shown). Rifampicin is known to bind the RpoB subunit of RNA polymerase and inhibit new transcription (Campbell et al., 2001). Therefore the quantity of any RNA species in a cell following exposure to rifampicin is expected to remain unchanged and any changes due to degradation would reflect its stability. The relative decrease of *ncrMT1302* compared to 5S rRNA indicates that it is less stable than the ribosomal RNA.

3.4. *ncrMT1302* expression is altered by cAMP and the absence of *MT1302*

sRNAs, such as GadY, regulate genes at the same genetic locus (Opdyke et al., 2004). The two genes flanking the *ncrMT1302* locus are *MT1302*, which encodes an adenylyl cyclase that converts ATP to cAMP (Dittrich et al., 2006), and *MT1303*, which encodes a hypothetical protein whose transcription is induced by cAMP under low-oxygen conditions (Gazdik and McDonough, 2005) (Figure 7). Therefore, it is reasonable to postulate that the locus may be involved in cAMP metabolism. We thus hypothesized that *ncrMT1302* expression could be affected by changes in cAMP concentration. To study this, we exposed cultures of wild-type (*WT*) *Mtb* to 110 μ M dibutyryl-cAMP (dbcAMP) as described (Gazdik and McDonough, 2005) for 8 hours, using cultures at the same growth phase but incubated in 7H9 broth without dbcAMP as controls. Total RNA was extracted from these cultures and *ncrMT1302* transcript levels were studied using qRT-PCR. During early exponential phase, the level of *ncrMT1302* was low and remained unchanged on exposure to dbcAMP. *ncrMT1302* levels were elevated during exponential phase and the addition of dbcAMP

reduced transcript levels 5 fold, while during stationary phase, when *ncrMT1302* expression was lower, dbcAMP increased transcript levels 5 fold (Figure 5A). Hence, cAMP supplementation results in the same fold change but is able to both increase and decrease *ncrMT1302* levels. Furthermore, *ncrMT1302* levels in untreated exponential phase cultures and cAMP supplemented stationary cultures are similar, suggesting that an optimal cellular concentration of the sRNA is required for it to function. This data implies that there is specific cAMP moderated regulation of *ncrMT1302* expression.

cAMP is involved in the function and expression of the two genes adjacent to *ncrMT1302* (Dittrich et al., 2006; Gazdik et al., 2009). Having identified cAMP as a modulator of *ncrMT1302* levels, we speculated that *ncrMT1302* may be functionally linked to genes at this locus. *MT1302* encodes the only adenylyl cyclase in *Mtb* that functions selectively under low pH conditions (Dittrich et al., 2006). Since we observed low pH to affect *ncrMT1302* expression, we hypothesized *ncrMT1302* to be functionally related to *MT1302*. To explore this, we compared levels of *ncrMT1302* between *WT Mtb* and a transposon mutant strain lacking a functional copy of *MT1302* (*MT1302::Tn*). As the mutant harbors a Himar 1 transposon insertion at the +780 nucleotide position from the putative translation start site (Lamichhane et al., 2003) and is 408 nucleotides downstream of the *ncrMT1302* locus, it is sufficient to disrupt *MT1302* function without interrupting the small RNA gene. This transposon has, on average, 30 stop codons in all frames, and therefore effectively disrupts *MT1302* translation (Rubin et al., 1999). We grew cultures of both strains and compared *ncrMT1302* expression levels at the following four growth phases: early exponential phase, exponential phase, stationary phase and clumped cultures. Total RNA was extracted and assessed using northern blotting and qRT-PCR. The qRT-PCR results show that *ncrMT1302* was expressed in the mutant lacking *MT1302* but there was no significant differential expression over different phases of growth (Figure 5B). Furthermore, the level of expression is lower compared to that of the *WT* strain. While the level of expression of *ncrMT1302* at early exponential and stationary phases were similar in the *WT* and *MT1302::Tn* strains, we observed a 5 fold increase in expression in the *WT* compared to the *MT1302::Tn* mutant during the exponential phase and in clumped cultures. The reduced expression of the sRNA in the mutant suggests that *MT1302* may be functionally linked to *ncrMT1302*.

3.5. Mouse model of TB demonstrates that *ncrMT1302* is expressed during infection

Macrophages infected with *Mtb* use acidified phagosomes as a form of defense against the bacteria (Rohde et al., 2007). Since we observed a reduced expression of *ncrMT1302* during acidic stress we hypothesized that the sRNA could have a functional role during infection. To test this hypothesis, we isolated total RNA from the lungs of naïve mice (negative control) and mice infected with *Mtb*, and performed northern blotting analysis with a probe against *ncrMT1302*. As *Mtb* RNA encompassed only a small fraction of the RNA extracted, we could assume that by loading the same amount of total RNA, there would be similar levels of host RNA in both uninfected and infected tissue derived RNA. 5S rRNA was used as a positive control for the presence of bacterial RNA and was detected only in the samples obtained from the lungs of *Mtb* infected mice. The analysis however detected an abundance of *ncrMT1302* RNA in the lungs of infected mice, indicative of a role of this sRNA during infection in the murine model of TB (Figure 6). While 5s rRNA and *ncrMT1302* were not detected in the total RNA samples from uninfected mice, the abundance of eukaryotic RNA resulted in nonspecific bands.

4. DISCUSSION

The intergenic sequences of many bacteria have now been demonstrated to contain sequences that encode small, regulatory RNAs which function in virulence, cell cycle and

numerous stress responses (Livny and Waldor, 2007; Padalon-Brauch et al., 2008). The first study to predict *Mtb* sRNA genes was carried out by Livny *et al* who predicted 56 sRNA genes in *Mtb* through the sRNAPredict2 program (Livny et al., 2006). However none of these were experimentally verified or identified in the studies that followed suggesting that bioinformatics analysis alone was not sufficient to identify sRNA genes in *Mtb*. While our studies were underway, Arnvig *et al* reported identification of 9 sRNA species through a cloning scheme in *Mtb* in the 20 to 70 nucleotide range (Arnvig and Young, 2009). They demonstrated that 3 of these sRNAs were induced by oxidative stress and a fourth, F6, was induced by both oxidative and acid stress. Overexpression of two of these sRNAs proved lethal in *Mtb* and over expression of F6 resulted in a slow growth phenotype. This indicates an important role for sRNAs in *Mtb* physiology and also suggests that sRNAs could be involved in regulation of a more global stress response as indicated by the response of F6 to multiple stresses. Soon thereafter, Dichiara *et al* identified sRNAs in *M.bovis* BCG, a non-virulent mycobacteria. Using a screen similar to the one reported here together with computational prediction, Dichiara *et al* reported the discovery of 34 novel sRNAs. This included an approximately 120 base sequence, *Mcr11*, from the *M.bovis* BCG transcriptome that is identical in sequence to the *ncrMT1302* RNA we report here. Our annotation is based on the existing unique annotation of the upstream open reading frame. They demonstrated that *Mcr11* was expressed in an *in vitro* culture of *Mtb* and its expression increased from late log to stationary phase of growth.

Here, we identify five small RNA candidates through a high throughput screen of an *Mtb* small RNA library consisting of RNA between 70 and 200 nucleotides (Table 1). Eight independent clones in our sRNA library harbored the *ncrMT1302* sequence, while other sRNA candidate sequences were observed only once. We propose two likely explanations for this observation: (a) *ncrMT1302* is stable with a half-life longer than other sRNA, or (b) it is expressed in higher abundance. The quantitative assessments we carried out show that *ncrMT1302* is indeed expressed at high levels in *Mtb* relative to *sigA* and that *ncrMT1302* is differentially expressed, with levels peaking in clumped cultures which represent late stationary phase. Our results correlate with previous reports as small regulatory RNAs have been implicated in the presence of environmental stresses during growth and this is most evident during stationary phase (Gottesman et al., 2006).

Northern blot and qRT-PCR analyses identified a low pH and rifampicin exposure as conditions that affect *ncrMT1302* levels in the cell. This is consistent with published reports that sRNAs function to coordinate stress responses in bacteria (Gottesman et al., 2006). A reduced pH is faced by the bacterium in the macrophage phagosome, suggesting that *ncrMT1302* could play a role in sensing and responding to pH changes. This is of particular importance as *Mtb* must be able to counteract acidification of phagolysosomes during host invasion (Pethe et al., 2004). The smaller expression changes observed under nutrient limitation and DETA-NO treatment could be due to downstream effects of these specific stress responses and not directly linked to *ncrMT1302* function. Furthermore, the increase in *ncrMT1302* levels under nutrient limitation could explain the elevated levels of the sRNA we observed in clumped cultures. Lastly, as rifampicin inhibits synthesis of new RNA, the reduced level of *ncrMT1302* is an indication of its stability relative to 5S rRNA whose levels remain relatively high upon rifampicin exposure (Campbell et al., 2001).

MT1302 encodes an adenylyl cyclase with an auto-inhibitory N-terminal domain (Tews et al., 2005) that makes it the only adenylyl cyclase in *Mtb* to function selectively at a low pH. *MT1303* encodes a hypothetical protein that is upregulated during macrophage infection (Monahan et al., 2001) and by acidic pH in *M. bovis* BCG but not in *Mtb* (Gazdik et al., 2009). Our observation that *ncrMT1302* levels are sensitive to low pH correlates well with the hypothesis that *ncrMT1302* could be functionally linked to genes at this locus. cAMP is

another modulator of this locus-with MT1302 being an adenylyl cyclase and *MT1303* being a cAMP-induced gene under low oxygen conditions (Gazdik and McDonough, 2005). Our study demonstrates that cAMP supplementation modulates ncrMT1302 levels and this too is consistent with the three genes at this locus being functionally linked. This hypothesis is further supported by the reduced expression and loss of differential expression of ncrMT1302 in the absence of a functional *MT1302* gene.

cAMP regulates expression of numerous genes in *Mtb* by binding to cNMP binding transcription factors. The best characterized of these is the CRP (cAMP receptor protein) transcription factor (McCue et al., 2000). cAMP binding enhances its DNA binding properties (Bai et al., 2005) enabling it to both positively and negatively regulate genes in *Mtb* such as *whiB1*, which has an activating and repressing CRP binding site (Stapleton et al., 2010). This paradigm supports the cAMP induced changes we observed in ncrMT1302 expression. Although we found no CRP binding domains upstream of *ncrMT1302*, there are 11 proteins with cNMP binding domains in *Mtb* (McCue et al., 2000) that could potentially regulate expression of ncrMT1302 in response to cAMP. A second cAMP associated transcription factor is Cmr (cAMP and macrophage regulator) (McCue et al., 2000). A previous study has reported that Cmr induces *MT1303* transcription by binding to the upstream region of *MT1303* (Gazdik et al., 2009). The study identified a binding region of Cmr using a gel shift assay. We mapped homology to this sequence to derive the coordinates of this site in Figure. 7 and show that this region significantly overlaps the *ncrMT1302* sequence. As such the *ncrMT1302* RNA sequence is highly homologous to the Cmr DNA binding sequencing, suggesting that Cmr could be a potential binding target for the ncrMT1302 sRNA.

Herein, we have assigned a stress response to an sRNA in *Mtb* by demonstrating a reduced expression of ncrMT1302 under low pH condition. This is the first report to identify a potential functional linkage of an sRNA in *Mtb* and demonstrate expression of an sRNA *in vivo* using the mouse model. As cAMP contributes to virulence and a pH stress response is vital for the survival of the bacillus, this study demonstrates a new avenue for studying TB disease through sRNAs.

Acknowledgments

We thank Hani Zaher for expert advice on RNA related studies, Maia Schoonmaker for providing RNA from infected animals and Cara Smith for assistance with cDNA library screening and dot-blot analysis. A gift of the miRCATTM RNA cloning Kit from Integrated DNA Technologies (IDT) is appreciated. This work was funded by NIH/NIAID AI30036, AI079590, AI37856, AI36973 and DP2OD008459. SP, WRB and GL planned the studies, SP conducted them, SP, WRB and GL analyzed the data and SP and GL wrote the manuscript.

Abbreviations

cAMP	Cyclic adenosine monophosphate
cDNA	Complementary DNA
CFU	Colony forming unit
Cmr	cAMP and macrophage regulator
cNMP	Cyclic nucleotide monophosphate
CRP	cAMP receptor protein
dbcAMP	Dibutyryl-cAMP
DETA-NO	Diethylenetriamine nitric oxide adduct

M.bovis BCG	Mycobacterium bovis Bacille Calmette-Guérin
Mtb	Mycobacterium tuberculosis
ncrMT1302	Non-coding RNA MT1302
ORF	Open reading frame
PBS	Phosphate buffered saline
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
sRNA	Small, non-coding RNA
TB	Tuberculosis
WT	Wild-type

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–402. [PubMed: 9254694]
- Arnvig KB, Young DB. Identification of small RNAs in *Mycobacterium tuberculosis*. *Mol Microbiol.* 2009; 73:397–408. [PubMed: 19555452]
- Bai G, McCue LA, McDonough KA. Characterization of *Mycobacterium tuberculosis* Rv3676 (CRPMt), a cyclic AMP receptor protein-like DNA binding protein. *J Bacteriol.* 2005; 187:7795–804. [PubMed: 16267303]
- Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell.* 2001; 104:901–12. [PubMed: 11290327]
- Dannenberg AM Jr. Immunopathogenesis of pulmonary tuberculosis. *Hosp Pract (Off Ed).* 1993; 28:51–8. [PubMed: 8419415]
- Dichiara JM, Contreras-Martinez LM, Livny J, Smith D, McDonough KA, Belfort M. Multiple small RNAs identified in *Mycobacterium bovis* BCG are also expressed in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. *Nucleic Acids Res.* 2010
- Dittrich D, Keller C, Ehlers S, Schultz JE, Sander P. Characterization of a *Mycobacterium tuberculosis* mutant deficient in pH-sensing adenylate cyclase Rv1264. *Int J Med Microbiol.* 2006; 296:563–6. [PubMed: 17005450]
- Fisher MA, Plikaytis BB, Shinnick TM. Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J Bacteriol.* 2002; 184:4025–32. [PubMed: 12081975]
- Gazdik MA, Bai G, Wu Y, McDonough KA. Rv1675c (cmr) regulates intramacrophage and cyclic AMP-induced gene expression in *Mycobacterium tuberculosis*-complex mycobacteria. *Mol Microbiol.* 2009; 71:434–48. [PubMed: 19040643]
- Gazdik MA, McDonough KA. Identification of cyclic AMP-regulated genes in *Mycobacterium tuberculosis* complex bacteria under low-oxygen conditions. *J Bacteriol.* 2005; 187:2681–92. [PubMed: 15805514]
- Gottesman S. Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet.* 2005; 21:399–404. [PubMed: 15913835]
- Gottesman S, McCullen CA, Guillier M, Vanderpool CK, Majdalani N, Benhammou J, Thompson KM, FitzGerald PC, Sowa NA, FitzGerald DJ. Small RNA Regulators and the Bacterial Response to Stress. *Cold Spring Harbor Symposia on Quantitative Biology.* 2006; 71:1–11.
- Hatfull, GF.; Jacobs, WR, Jr. RNA Preparation-Trizol. American Society for Microbiology; Washington DC: 2000.

- Hobson RJ, McBride AJ, Kempell KE, Dale JW. Use of an arrayed promoter-probe library for the identification of macrophage-regulated genes in *Mycobacterium tuberculosis*. *Microbiology*. 2002; 148:1571–9. [PubMed: 11988532]
- Lamichhane G, Zignol M, Blades NJ, Geiman DE, Dougherty A, Grosset J, Broman KW, Bishai WR. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*. 2003; 100:7213–8. [PubMed: 12775759]
- Livny J, Brencic A, Lory S, Waldor MK. Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. *Nucleic Acids Res*. 2006; 34:3484–93. [PubMed: 16870723]
- Livny J, Waldor MK. Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol*. 2007; 10:96–101. [PubMed: 17383222]
- Lorian, V. *Antibiotics in Laboratory Medicine*. 5. Lippincott Williams & Wilkins; 2005.
- Manabe YC, Kesavan AK, Lopez-Molina J, Hatem CL, Brooks M, Fujiwara R, Hochstein K, Pitt ML, Tufariello J, Chan J, McMurray DN, Bishai WR, Dannenberg AM Jr, Mendez S. The aerosol rabbit model of TB latency, reactivation and immune reconstitution inflammatory syndrome. *Tuberculosis (Edinb)*. 2008; 88:187–96. [PubMed: 18068491]
- Manganelli R, Dubnau E, Tyagi S, Kramer FR, Smith I. Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol Microbiol*. 1999; 31:715–24. [PubMed: 10027986]
- McCue LA, McDonough KA, Lawrence CE. Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*. *Genome Res*. 2000; 10:204–19. [PubMed: 10673278]
- Monahan IM, Betts J, Banerjee DK, Butcher PD. Differential expression of mycobacterial proteins following phagocytosis by macrophages. *Microbiology*. 2001; 147:459–71. [PubMed: 11158363]
- Nathan C, Shiloh MU. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A*. 2000; 97:8841–8. [PubMed: 10922044]
- Oh JT, Cajal Y, Skowronska EM, Belkin S, Chen J, Van Dyk TK, Sasser M, Jain MK. Cationic peptide antimicrobials induce selective transcription of *micF* and *osmY* in *Escherichia coli*. *Biochim Biophys Acta*. 2000; 1463:43–54. [PubMed: 10631293]
- Opdyke JA, Kang JG, Storz G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol*. 2004; 186:6698–705. [PubMed: 15466020]
- Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, Margalit H, Altuvia S. Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res*. 2008; 36:1913–27. [PubMed: 18267966]
- Pak J, Fire A. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science*. 2007; 315:241–4. [PubMed: 17124291]
- Pethe K, Swenson DL, Alonso S, Anderson J, Wang C, Russell DG. Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci U S A*. 2004; 101:13642–7. [PubMed: 15340136]
- Rohde KH, Abramovitch RB, Russell DG. *Mycobacterium tuberculosis* invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe*. 2007; 2:352–64. [PubMed: 18005756]
- Rubin EJ, Akerley BJ, Novik VN, Lampe DJ, Husson RN, Mekalanos JJ. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proc Natl Acad Sci U S A*. 1999; 96:1645–50. [PubMed: 9990078]
- Sachdeva P, Misra R, Tyagi AK, Singh Y. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J*. 2010; 277:605–26. [PubMed: 19951358]
- Smeulders MJ, Keer J, Speight RA, Williams HD. Adaptation of *Mycobacterium smegmatis* to stationary phase. *J Bacteriol*. 1999; 181:270–83. [PubMed: 9864340]
- Stapleton M, Haq I, Hunt DM, Arnvig KB, Artymiuk PJ, Buxton RS, Green J. *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J Biol Chem*. 2010; 285:7016–27. [PubMed: 20028978]

- Tews I, Findeisen F, Sinning I, Schultz A, Schultz JE, Linder JU. The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme. *Science*. 2005; 308:1020–3. [PubMed: 15890882]
- Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, Huttenhofer A, Wagner EG. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res*. 2003; 31:6435–43. [PubMed: 14602901]
- Vogel J, Sharma CM. How to find small non-coding RNAs in bacteria. *Biol Chem*. 2005; 386:1219–38. [PubMed: 16336117]
- Wallis RS, Patil S, Cheon SH, Edmonds K, Phillips M, Perkins MD, Joloba M, Namale A, Johnson JL, Teixeira L, Dietze R, Siddiqi S, Mugerwa RD, Eisenach K, Ellner JJ. Drug tolerance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 1999; 43:2600–6. [PubMed: 10543735]
- Wassarman KM, Storz G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell*. 2000; 101:613–23. [PubMed: 10892648]
- Wassarman KM, Zhang A, Storz G. Small RNAs in *Escherichia coli*. *Trends Microbiol*. 1999; 7:37–45. [PubMed: 10068996]
- Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell*. 2009; 136:615–28. [PubMed: 19239884]
- WHO. WHO global tuberculosis control report 2010. Summary. *Cent Eur J Public Health*. 2010; 18:237. [PubMed: 21361110]

Highlights

- We identified small non-coding RNAs in *M. tuberculosis*, the pathogen that causes TB
- We characterized the role of one of these RNAs in the physiology of the pathogen
- This RNA, ncrMT1302, responds to changes in pH and cAMP
- This RNA is expressed by *M. tuberculosis* in the lungs of infected mice

Figure 1.
Mtb sRNA cloning scheme.

Figure 2.
Genetic locus of the largest clone obtained from the intergenic region between *MT1302* and *MT1303*.

(a)

(b)

Figure 3. Expression of ncrMT1302 during various phases of growth *in vitro*

Detection of ncrMT1302 at various phases of growth *in vitro*. (A) Northern blot of ncrMT1302 (top panel) and 5S rRNA (bottom panel) as a loading control. (B) qRT-PCR showing relative expression levels of ncrMT1302 calculated using the sigA transcript for normalization of RNA amounts. * $p < 0.05$.

(a)

(b)

Figure 4. Level of ncrMT1302 is altered by low pH and rifampicin exposure

(A) qRT-PCR data showing relative levels of ncrMT1302 normalised to sigA under different stress conditions. * $p < 0.05$. (B) Detection of ncrMT1302 (top panel) and 5S rRNA (bottom panel) transcripts isolated from *Mtb* cultures exposed to rifampicin, isoniazid and 7H9 broth using northern blotting.

(a)

(b)

Figure 5. ncrMT1302 expression is altered by cAMP and the absence of MT1302

(A) Cultures of *WT Mtb* at early exponential, mid-exponential and stationary phases of growth were exposed to 110uM dibutytyl cAMP for 8 hours and ncrMT1302c expression was assessed using qRT-PCR. Black bars (-dbcAMP) and grey bars (+dbcAMP). (B) ncrMT1302 levels were compared between *WT Mtb* and a mutant lacking a functional copy of *MT1302*. RNA levels were measured using qRT-PCR at different growth phases in the *WT* (grey) and *MT1302::Tn* mutant (black) strains. * $p < 0.05$.

Figure 6. ncrMT1302 is expressed during infection in the lungs of mice
Total RNA was extracted from *Mtb* infected and uninfected mouse lungs and probed for ncrMT1302 (top panel) and 5S rRNA (bottom panel) using northern blotting.

Figure 7. Functional genomic schematics of the *ncrMT1302* locus

Past studies (bold line) have shown that MT1302 is an adenylyl cyclase that generates cAMP at low pH and that cAMP induces expression of MT1303 under low-oxygen conditions. Cmr (cAMP and macrophage regulator) is a transcription factor that binds to the promoter region of *MT1303* and this binding site overlaps the *ncrMT1302* region we have identified. Our studies (dotted line) have shown that a low pH and changes in cAMP concentration affect the expression of *ncrMT1302*. Mapped coordinates are based on the *Mtb* CDC1551 genome.

Table 1

Genetic loci of 276 clones sequenced from a cDNA library representing *Mtb* at late exponential phase of growth.

Number of Clones	Description	Percentage of clones (%)
238	tRNA, rRNA	86.2
26	Intragenic	9.4
12	Intergenic	4.4

Table 2

Description of small RNA loci in the intragenic regions of the *Mtb* genome.

Columns 1 and 2 : Gene Annotation and Identification in the *Mtb* CDC1551 genome. Columns 3 and 4: Coordinates of the 5' and 3' ends of the cloned sRNA relative to the origin of replication in the *Mtb* genome CDC1551.

Gene Annotation	Gene ID	5' coordinate	3' coordinate
Hypothetical Protein	MT0013	13234	13135
Hypothetical Protein	MT0117	130898	130974
Virulence factor MCE family protein	MT0180	201940	202007
Hypothetical Protein	MT0184	207254	207203
PPE Family Protein	MT0269	309156	309085
Phosphate transport regulatory protein	MT0843	913882	913928
Thiolase	MT0882	955365	955412
Undecaprenyl Diphosphate Synthase	MT1118	1211612	1211654
Hypothetical Protein	MT1157.1	1248876	1248807
Hypothetical Protein	MT1183	1276639	1276529
Putative lipoprotein	MT1203	1296067	1296008
Hypothetical Protein	MT1356	1472964	1473049
PE PGRS family protein	MT1367	1489846	1489796
OXPP cycle protein OpcA	MT1493	1624836	1624790
Glycogen operon protein	MT1615	1771027	1771084
Hypothetical Protein	MT1876	2071171	2071055
Hypothetical Protein	MT1935	2133853	2133912
Hypothetical protein	MT1935	2133853	2133953
Hypothetical Protein	MT1995	2193835	2193945
Helicase	MT2162	2364224	2364176
Conserved Hypothetical Protein	MT2245	2450674	2450621
Hypothetical Protein	MT3761	4092036	4092016
ABC transporter, permease	MT3863	4194200	4194264
ABC transporter, permease	MT3863	4194200	4194264
Hypothetical Protein	MT3885	1412590	1412672
Conserved transmembrane protein	MT4029	4390088	4390156

Table 3

Description of small RNA loci located in the intergenic regions of the *Mtb* genome.

Columns 1 and 2 : Coordinates of the 5' and 3' ends of the cloned sRNA relative to the origin of replication in the *Mtb* genome CDC1551. Columns 3 and 4: Gene Annotation and Identification of the 5' flanking gene. Columns 5 and 6: Gene Annotation and Identification of the 3' flanking gene. Column 7: Size of the cDNA insert cloned.

5' coordinate	3' coordinate	5' Gene ID	5' Gene Annotation	3' Gene	3' Gene Annotation	Insert size (nucleotides)
1219936	1219988	MT1124	Pantothenate kinase	MT1125	Ser Hydroxymethyltransferase	52
1340065	1340112	MT1234	PPE family prot	MT1235	Hypothetical protein	47
1412645	1412590	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	56
1412633	1412593	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	40
1412695	1412587	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	109
1412693	1412593	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	100
1412637	1412587	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	51
1412673	1412590	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	84
1412672	1412590	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	83
1412637	1412590	MT1302	Adenylate cyclase, putative	MT1303	Hypothetical protein	48
2622446	2622492	MT2412	HP (EsxP)	MT2413	Hypothetical protein	46
4307121	4307214	MT3949	Possible Bacterioferritin	MT3952	Hypothetical protein	94