The FEBS Journal

FEBS PRESS science publishing by scientists

Hypothetical protein Rv3423.1 of *Mycobacterium tuberculosis* is a histone acetyltransferase

Leny Jose[1], Ranjit Ramachandran[1], Raghu Bhagavat[2], Roshna Lawrence Gomez[1], Aneesh Chandran[1], Sajith Raghunandanan[1], Ramakrishnapillai Vyomakesannair Omkumar[3], Nagasuma Chandra[2], Sathish Mundayoor[1] and Ramakrishnan Ajay Kumar[1]

1 Mycobacterium Research Group, Tropical Disease Biology Division, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, India
2 Bioinformatics Centre, Indian Institute of Science, Bangalore, India
3 Neurobiology Group, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, India

Keywords
chromatin; histone acetyltransferase; macrophage; *Mycobacterium tuberculosis*; nucleomodulin

Correspondence
R. A. Kumar, Mycobacterium Research Group, Tropical Disease Biology Division, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram 695014, India
Fax: +91 471 2348096
Tel: +91 471 2529513
E-mail: rakumar@rgcb.res.in

(Received 14 May 2015, revised 28 September 2015, accepted 15 October 2015)

doi:10.1111/febs.13566

We isolated an 8 kDa mycobacterial hypothetical protein, Rv3423.1, from the chromatin of human macrophages infected with *Mycobacterium tuberculosis* H37Rv. Bioinformatics predictions followed by *in vitro* biochemical assays with purified recombinant protein showed that Rv3423.1 is a novel histone acetyltransferase that acetylates histone H3 at the K9/K14 positions. Transient transfection of macrophages containing GFP-tagged histone H1 with RFP-tagged Rv3423.1 revealed that the protein co-localizes with the chromatin in the nucleus. Co-immunoprecipitation assays confirmed that the Rv3423.1–histone interaction is specific. Rv3423.1 protein was detected in the culture filtrate of virulent but not avirulent *M. tuberculosis*. Infection of macrophages with recombinant *Mycobacterium smegmatis* constitutively expressing Rv3423.1 resulted in a significant increase in the number of intracellular bacteria. However, the protein did not seem to offer any growth advantage to free-living recombinant *M. smegmatis*. It is highly likely that, by binding to the host chromatin, this histone acetyltransferase from *M. tuberculosis* may manipulate the expression of host genes involved in anti-inflammatory responses to evade clearance and to survive in the intracellular environment.

Introduction

Tuberculosis (TB) is caused by *Mycobacterium tuberculosis* (MTB), an intracellular human pathogen that is disseminated through air. TB has been a prominent human disease for several thousand years. The latest report from the World Health Organization revealed that 9.6 million people developed TB in 2014, and it caused 1.5 million deaths [1]. An important challenge in the study of MTB is to understand how this bacterium escapes elimination by the immune system of the host. When the pathogen enters the lungs of a susceptible person, MTB succeeds in countering the defense mechanisms of the alveolar macrophages, and eventually manages to survive and multiply in the very cells that are designated to eliminate the bacterium. The strategies used by the pathogen are being intensely investigated in many laboratories, and consequently much is known now about the tactics employed by intracellular MTB to tackle host defense mechanisms and cause TB. Dissecting the mechanisms by which intracellular pathogens regulate host gene expression

Abbreviations
AAC(6′)-Iy, aminoglycoside 6′-N-acetyltransferase; ChIP, chromatin immunoprecipitation; GFP, green fluorescent protein; HAT, histone acetyltransferase; MTB, *Mycobacterium tuberculosis*; TB, tuberculosis.

for their own advantage is crucial in developing methods to control the diseases caused by them. Regulation of gene expression in eukaryotic cells is brought about by the concerted activation and repression of transcription through chromatin remodeling and modifications that control the accessibility of DNA regulatory sequences to transcription factors [2]. Post-translational modifications of histones (methylation, acetylation, phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, etc.) modulate sequence-specific DNA–protein interactions [3]. Chromatin modification as a mechanism by which intracellular bacteria modify their host environment has gained considerable attention in the recent past. Several bacterial pathogens, such as *Listeria monocytogenes* [4], *Clostridium perfringens* [5], *Streptococcus pneumoniae* [5], *Helicobacter pylori* [6] and *Anaplasma phagocytophilum* [7], are known to cause epigenetic modifications to manipulate host responses associated with inflammation and immunity, which play a role in progression of the disease. Recently, such bacterial factors that act directly in the nucleus have been named 'nucleomodulins' [8]. Infection by MTB has been shown to bring about changes in gene expression in the host macrophages, leading to induction or subversion of several immunomodulatory pathways [9,10]. We hypothesized that at least some of these changes may be caused by chromatin modifications induced directly by some products of the intracellular pathogen, or indirectly by recruitment of host chromatin-modifying enzymes by some bacterial factors. Such pathogenically induced reprogramming of host gene expression may either augment or diminish the host's defense mechanisms, depending on which direction the equilibrium shifts. Although MTB has been reported to cause chromatin remodeling in macrophages [11,12], the identity and function of the bacterial or cellular component(s) involved in this process are largely unknown. Here we report the isolation and identification of one such mycobacterial factor that directly binds to the host chromatin and causes hyperacetylation of histone H3 to up-regulate gene expression. We infected macrophages with MTB H37Rv, immunoprecipitated the chromatin using antibody against histone H3, and analyzed the chromatin-bound proteins by 2D gel electrophoresis followed by LC/MS/MS. A unique protein present in the chromatin complex of infected macrophages was identified as Rv3423.1, an hypothetical protein of 8.4 kDa. By homology modeling and subsequent *in vitro* assays, we demonstrated that this protein possesses histone acetyltransferase activity (HAT) on histone H3 at the K9/K14 positions.

Results

Growth and survival of *M. tuberculosis* in human macrophages

THP-1-derived macrophages were infected at a multiplicity of infection of 20 MTB bacilli per cell. Internalization of bacteria by macrophages was confirmed by confocal microscopy (Fig. 1A) using MTB expressing GFP. Expression of host *CORO1A*, a well established marker for viable intracellular mycobacteria [13], was analyzed by quantitative PCR. The levels of *CORO1A* were significantly higher at 24 h than at 12 h postinfection in macrophages containing live bacteria, confirming that the intracellular mycobacteria were viable (Fig. 1B). Use of heat-killed bacteria did not cause any increase in the levels of *CORO1A* in macrophages.

MTB hypothetical protein Rv3423.1 binds to the chromatin in infected macrophages

Using a modified ChIP protocol (mChIP) [14,15], we immunoprecipitated the chromatin of infected as well as non-infected macrophages with antibody against histone H3. Upon resolving the immunoprecipitates by 2D gel electrophoresis, one distinct spot was observed in the infected sample, but not in the uninfected sample (Fig. 1C). The MALDI-TOF mass spectrometry profile and peptide mass fingerprint of this protein are shown in Fig. 2A,B. Mascot database search using -NCBI-nr (National Center for Biotechnology Information-non redundant) protein database (www.matrixscience.com) of the spot identified it as mycobacterial hypothetical protein MT3532.1 in MTB CDC1551, whose corresponding number in MTB H37Rv is Rv3423.1. The Rv3423.1 protein has 76 amino acids and a predicted molecular mass of 8.4361 kDa. The amino acid sequence of the protein was submitted to the UNIPROT protein database (http://www.uniprot.org/uniprot/P9WKY5). A BLASTn search revealed that the full-length gene is present in all strains of MTB, and also in some members of the MTB complex, namely *Mycobacterium canetti*, *Mycobacterium bovis* and *Mycobacterium africanum*, with 100% homology. However, it is absent in fast-growing non-pathogenic *Mycobacterium smegmatis*.

In silico structure modeling and function prediction

The bioinformatics protocol that we employed is illustrated in Fig. 3A. A FASTA search using the

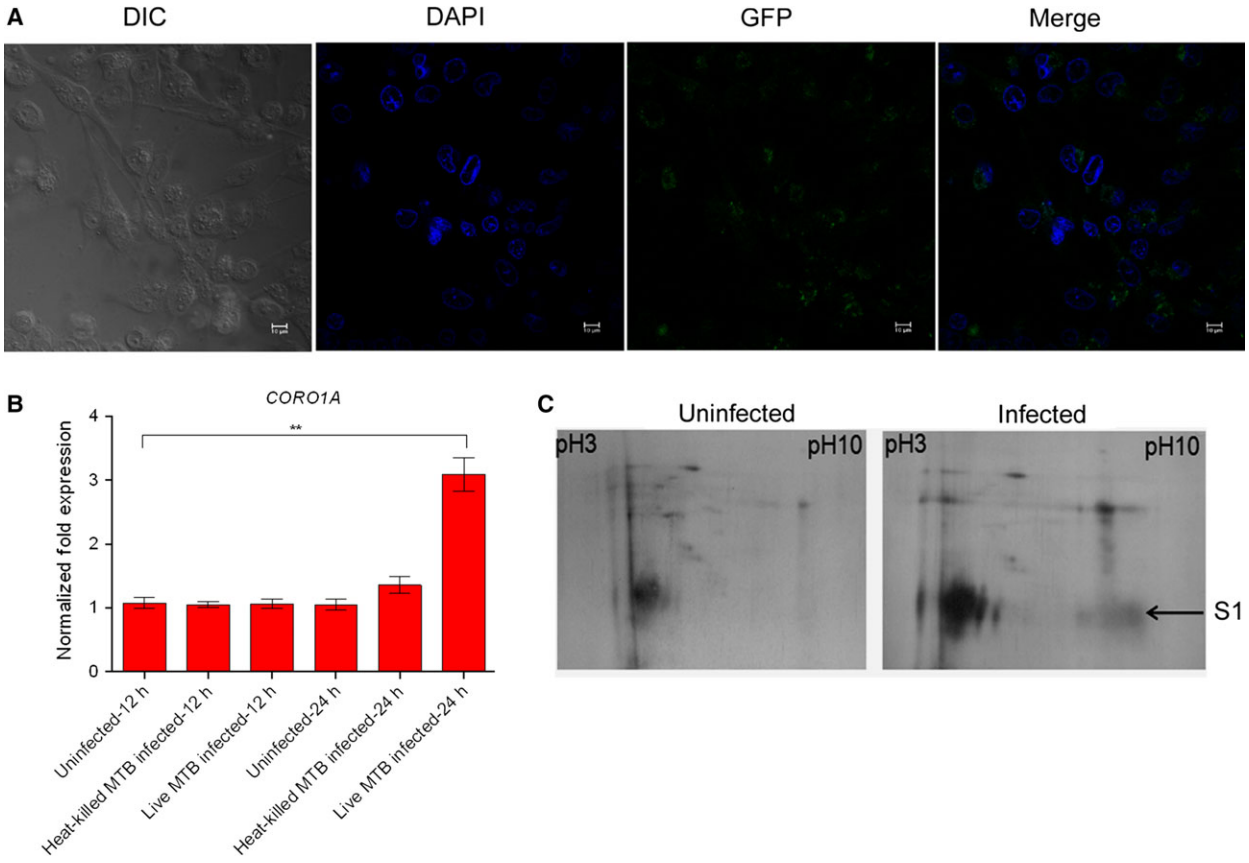
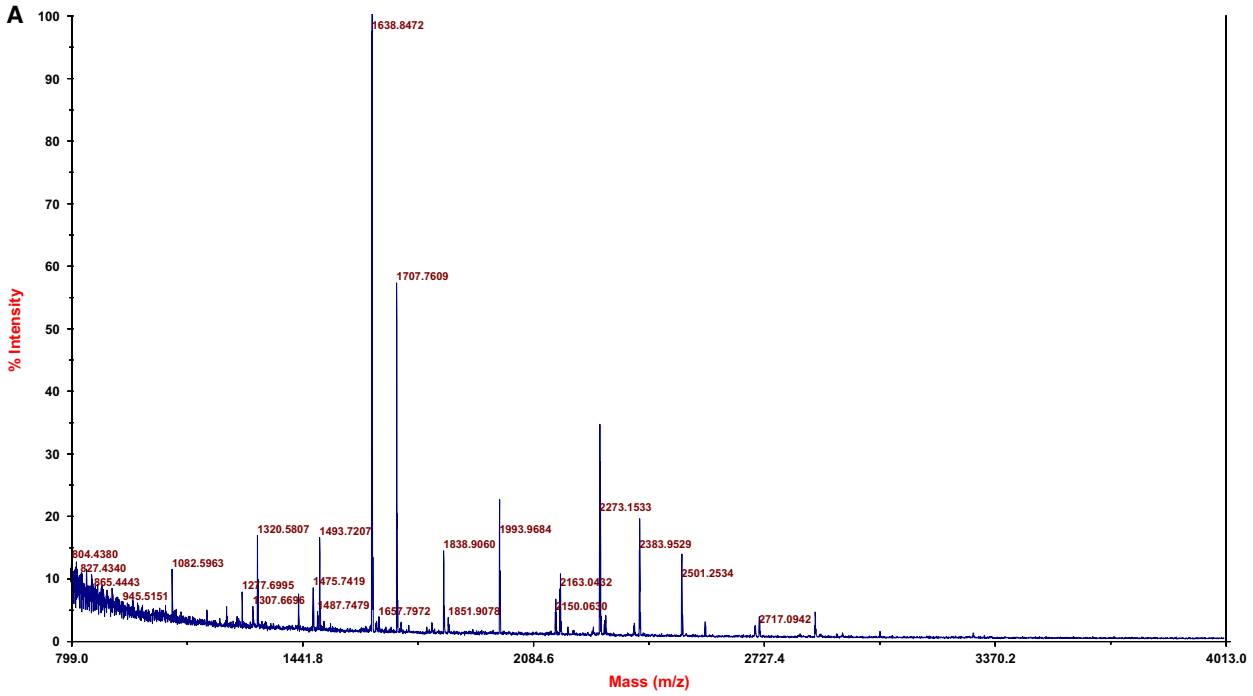


Fig. 1. Photomicrographs of intracellular MTB, and 2D gel electrophoretograms of chromatin immunoprecipitates from uninfected and infected macrophages. (A) Microscopic images of THP-1-derived macrophages infected with GFP-expressing MTB H37Rv at a multiplicity of infection of 20:1. Cell nuclei were counter-stained with DAPI, and images were captured at 40× magnification. The images shown are representatives of at least three experiments. Scale bar = 1 μm. (B) Levels of *CORO1A* transcripts in macrophages at 12 and 24 h post-infection with live MTB H37Rv at a multiplicity of infection of 20:1. The mRNA expression levels were compared with those for uninfected macrophages and macrophages infected with heat-killed MTB H37Rv. *GAPDH* was used as the endogenous control for quantitative PCR. Values are means and standard deviations from three independent experiments. (C) 2D gel electrophoresis of immunoprecipitates from uninfected and infected macrophages. THP-1-derived macrophages were infected with MTB H37Rv at a multiplicity of infection of 20:1. Chromatin from uninfected macrophages (left) and infected macrophages (right) was immunoprecipitated using antibody against histone H3. The proteins of the precipitate were subjected to isoelectric focusing using an IPG strip (linear pH gradient of 3-10) and subsequently resolved by 10% SDS/PAGE. The protein spot S1 that is unique to infected macrophages was subjected to MALDI-TOF analysis.

Sequence Annotated by Structure (SAS) tool [16] resulted in identification of 1S5K, an aminoglycoside 6-N-acetyltransferase [AAC(6′)-ly] from *Salmonella enterica*, as a template, with a Smith–Waterman score of 70, 26.8% sequence identity and 100% query coverage. Alternately, the fold recognition procedure resulted in identification of 1J2Z, an UDP-*N*-acetylglucosamine acyltransferase from *Helicobacter pylori*, as a template with an @TOME [17] score of 88.65, 18.5% sequence identity and 100% query coverage. Most hits resulting from the @TOME search corresponded to acyltransferases, indicating that acyltransferase activity was the most probable function. Nucleotide binding activity, which is observed in acyl transferases [18], was also suggested for this protein. From the FASTA and @TOME results, we built 3D structure models for the query sequence using 1S5K and 1J2Z as templates. Thus the bioinformatics approaches suggested acyl and acetyltransferases as possible templates, and transferase and nucleotide binding activities as possible functions. The *S. enterica* template 1S5K showed the presence of the ligand acetyl CoA at the binding site, and thus the model built on 1S5K could be annotated with the binding of acetyl CoA by structurally superposing the model onto the template and mapping the binding site based on 1S5K (Fig. 3B). The AAC(6′)-ly of *S. enterica* is a HAT in addition to being an aminoglycoside *N*-acetyltransferase [19]. There is a



B

Gel Idx/Pos	78/D3	Instr./Gel Origin	AK069/MALDI 2008.03.31 9192	Process Status	Analysis Succeeded
Plate [#] Name	[1] MALDI 2008.03.31 9192	Instrument Sample Name		Spectra	1

Rank	Protein Name	Species	Accession No.	Protein MW	Protein PI	Protein Score	Protein Score C. I. %	Pep. Count
1	hypothetical protein MT3532.1 [Mycobacterium tuberculosis CDC1551]	Mycobacterium tuberculosis CDC1551	gi|13883362	8436.1	9.47	48	88.391	6

Peptide Information

Calc. Mass	Obsrv. Mass	± da	± ppm	Start Seq.	End Seq.	Sequence	Modification
801.4036	801.4358	0.0322	40	28	35	GAGLQPCR	
806.4011	806.4449	0.0438	54	59	65	MMVGVNR	
816.4144	816.4413	0.0269	33	36	42	RDCPVAR	
822.396	822.4551	0.0591	72	59	65	MMVGVNR	Oxidation (M)[1]
838.3909	838.4739	0.083	99	59	65	MMVGVNR	Oxidation (M)[1,2]
1320.6809	1320.5804	−0.1005	−76	57	67	CRMMVGVNRVR	
1434.6471	1434.7535	0.1064	74	1	12	MSFGFPTFSQNR	Oxidation (M)[1]
1638.7944	1638.8469	0.0525	32	13	27	FTEQYSGLCPIAPGR	

Fig. 2. Protein identification by MALDI-TOF MS and peptide mass fingerprinting. (A) Peptide mass fingerprint of spot S1. The protein spot was cut from the 2D gel, 'in-gel' trypsin-digested, purified, and analyzed by MALDI-TOF MS. (B) Protein and peptide summary report for spot S1, generated by GPS Explorer. The observed masses were queried to the Mascot database. A protein score > 40 was considered to indicate a statistically significant match ($P < 0.05$). The observed masses of six peptides significantly matched the theoretical masses of tryptic digests of Rv3423.1 with a window of error of < 100 p.p.m.

significant structural similarity between the bacterial AAC(6′)-ly and the Hpa2 HAT of yeast [19]. This *Salmonella* enzyme is able to acetylate calf thymus histone preparations and also the 21 amino acid human histone H3 N-terminal peptide [19,20]. The *in silico* prediction that Rv3423.1 protein has significant sequence similarity with the *S. enterica* AAC (6′)-ly, and the fact that the protein was found bound to the chromatin of infected macrophages, strongly suggest HAT activity for the protein.

Rv3423.1 protein acts as a histone acetyltransferase *in vitro*

The *Rv3423.1* gene was amplified from the genomic DNA of MTB H37Rv, cloned and expressed as an N-terminal fusion in a pET SUMO expression system (Invitrogen, Carlsbad, CA, USA) in *Escherichia coli* (Fig. 4A). The fusion protein was purified using Ni-NTA affinity chromatography, and was cleaved using SUMO protease to yield the 8.4 kDa native protein

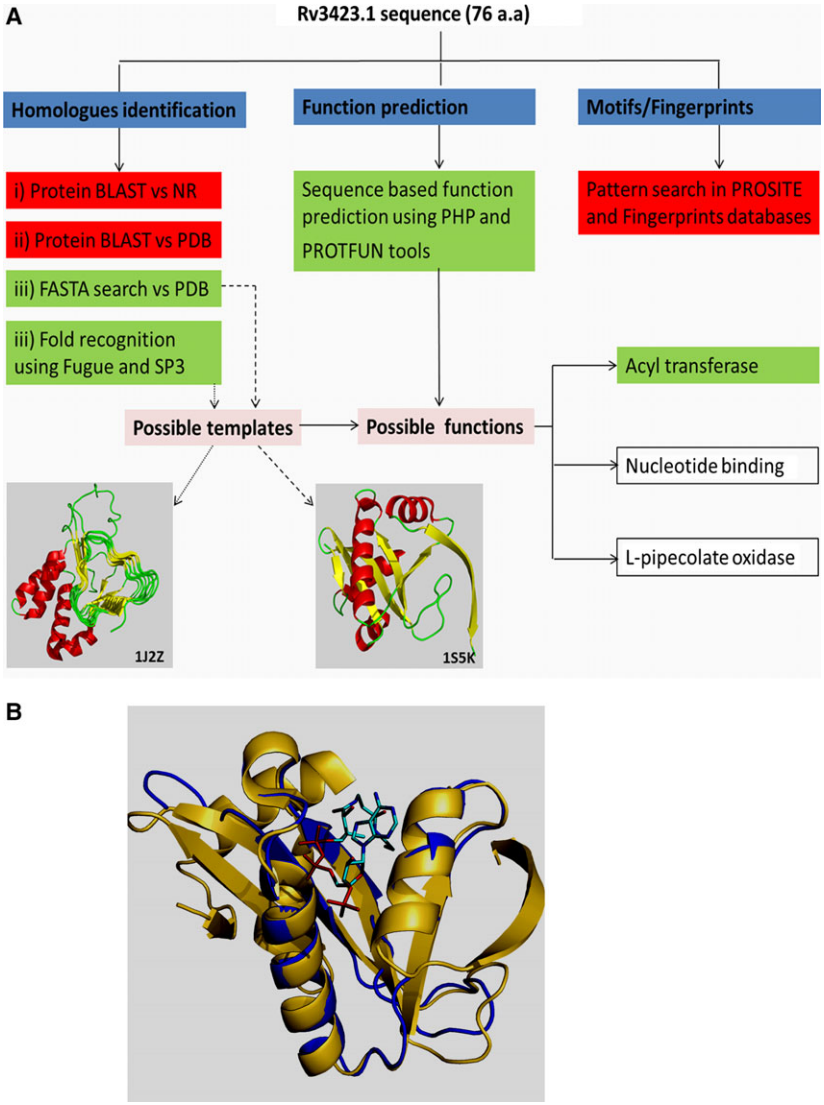


Fig. 3. Homology modeling and prediction of the function of the Rv3423.1 protein. (A) The major steps involved in the modeling are shown in blue boxes, tools/results that gave conclusive answers are shown in green boxes, and tools/results that did not give any clues are shown in red boxes. Ribbon representations of the two probable models identified in the study are shown in separate panels with the PDB IDs of the structures chosen as templates. (B) Structural superposition of template 1S5K (*S. enterica* AAC(6′)-Iy) and the model based on it. Both structures are shown in ribbon representation (yellow for 1S5K and blue for the model). The binding pocket of acetyl CoA (sticks) is also shown.

(Fig. 4B). The purified native protein was subjected to various HAT assays to validate the *in silico* prediction.

Incorporation of the radioactive acetyl group on histones from ^{3}H-acetyl CoA

First we tested the HAT activity of the purified Rv3423.1 protein using calf thymus histones as substrate in the presence of ^{3}H-labeled acetyl CoA. A significant amount of radioactivity was incorporated into calf thymus histones and retained on the P81 filter disc, corresponding to ^{3}H-acetyl groups added to histones, when recombinant Rv3423.1 was used as the HAT enzyme. The results also indicated that Rv3423.1 readily acetylates histones in a concentration-dependent manner (Fig. 5A). We used recombinant Rv2059 protein (a periplasmic solute-binding protein of MTB), cloned and purified under similar conditions, as a mock control. When corresponding amounts of Rv2059 protein were used in the assay, only a basal level of radioactivity was retained on the filter disc, indicating that the high levels of incorporation of the ^{3}H-acetyl group onto histones are indeed due to the HAT activity possessed by Rv3423.1 protein.

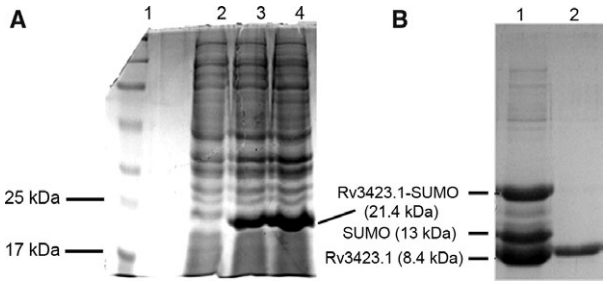


Fig. 4. Over-expression of the Rv3423.1–SUMO fusion protein in *E. coli*, and SUMO protease cleavage. (A) Over-expression of SUMO-tagged Rv3423.1 after induction with isopropyl-β-D-thiogalactopyranoside at 28 °C. Crude cell lysates (20 μg) were resolved by SDS/PAGE. Lane 1, molecular weight markers; lane 2, un-induced sample; lane 3, induced for 4 h; lane 4, induced for 8 h. (B) The recombinant Rv3423.1–SUMO fusion protein was subjected to cleavage of the tag by protease for 16 h, and native protein was purified. Lane 1, recombinant fusion protein after protease cleavage; lane 2, eluted native Rv3423.1 protein.

NADH generation

Histone acetylation by HATs releases free CoA, which can reduce NAD to NADH with the help of pyruvate dehydrogenase. In a cell-free assay, the NADH generated as a result of acetylation of a 21 amino acid N-terminal peptide of histone H3 was quantified as an indirect measure of HAT activity, and was expressed as relative absorbance per nanogram of protein. The HAT activity of Rv3423.1 increased gradually from the 1st hour till the 4th hour (Fig. 5B). At the 3rd and 4th hours, the extent of acetylation was comparable to that by the positive control (HeLa nuclear extract). Recombinant Rv2059 protein, used as a mock control, did not show any detectable HAT activity.

ELISA

Antibody against acetyl histone H3 (K9/14) specifically detected the acetyl 21 amino acid N-terminal peptide of histone H3. Our data strongly suggest that Rv3423.1 protein is able to acetylate histone H3 at positions K9 and K14 (Fig. 5C). The efficiency of acetylation by Rv3423.1 protein was comparable to that by the positive control (HeLa nuclear extract). In addition, the acetyltransferase activity of the protein was totally abolished by anacardic acid (10 μM), an inhibitor of HAT activity.

Immuno dot blot

To confirm the above results, we performed an immuno dot blot assay with antibody against acetyl histone H3 (K9/K14). The acetylation was evident when 250 ng protein was used in the reaction, which correlated with the results from ELISA. The extent of acetylation increased significantly when the amount of the protein was doubled. In agreement with the ELISA results, anacardic acid at 10 μM completely inhibited the HAT activity of Rv3423.1 (Fig. 5D).

Together, the results from these four assays unequivocally prove that Rv3423.1 is a HAT.

Rv3423.1 gene is expressed by intracellular MTB

After infecting the macrophages with MTB H37Rv, we analyzed the expression of the *Rv3423.1* gene by intracellular bacteria using quantitative PCR. We found that expression of the gene marginally increased (1.5-fold) from 12 to 24 h post-infection (Fig. 6). The level of *Rv3423.1* expression was found to not to decrease up to 60 h, after which we were unable to follow the expression due to death of infected cells.

Co-localization of Rv3423.1 protein with chromatin

We co-transfected THP-1-derived macrophage cells with constructs carrying histone H1–eGFP and Rv3423.1–DsRed fusions, and analyzed the fluorescence by confocal microscopy. As expected, H1–eGFP fluorescence was exclusively localized in the nucleoplasm. Although a significant amount of Rv3423.1 protein was found to be present in the cytoplasm, it was also observed inside the nucleus. To confirm its nuclear localization, we analyzed Z-stacked images of transfected cells, and found that Rv3423.1–DsRed was present as large distinct spots distributed randomly inside the nucleus against the uniform green fluorescence of the chromatin (Fig. 7). A movie composed of stacked images clearly showed a 3D distribution of the protein inside the nucleus of transfected macrophages, in association with the chromatin (Movie S1), which confirmed the co-localization.

Rv3423.1 physically interacts with histone H3

Using various *in vitro* assays, we have shown that Rv3423.1 acetylates histone H3. To perform its HAT activity, it is necessary that Rv3423.1 physically interacts with histone H3. We performed a co-immunoprecipitation assay to demonstrate this interaction. The protein complex was immunoprecipitated with antibody against histone H3, and the bound proteins were probed for the presence of Rv3423.1 by Western analysis using antibodies raised against Rv3423.1 protein. The results clearly showed that Rv3423.1 is

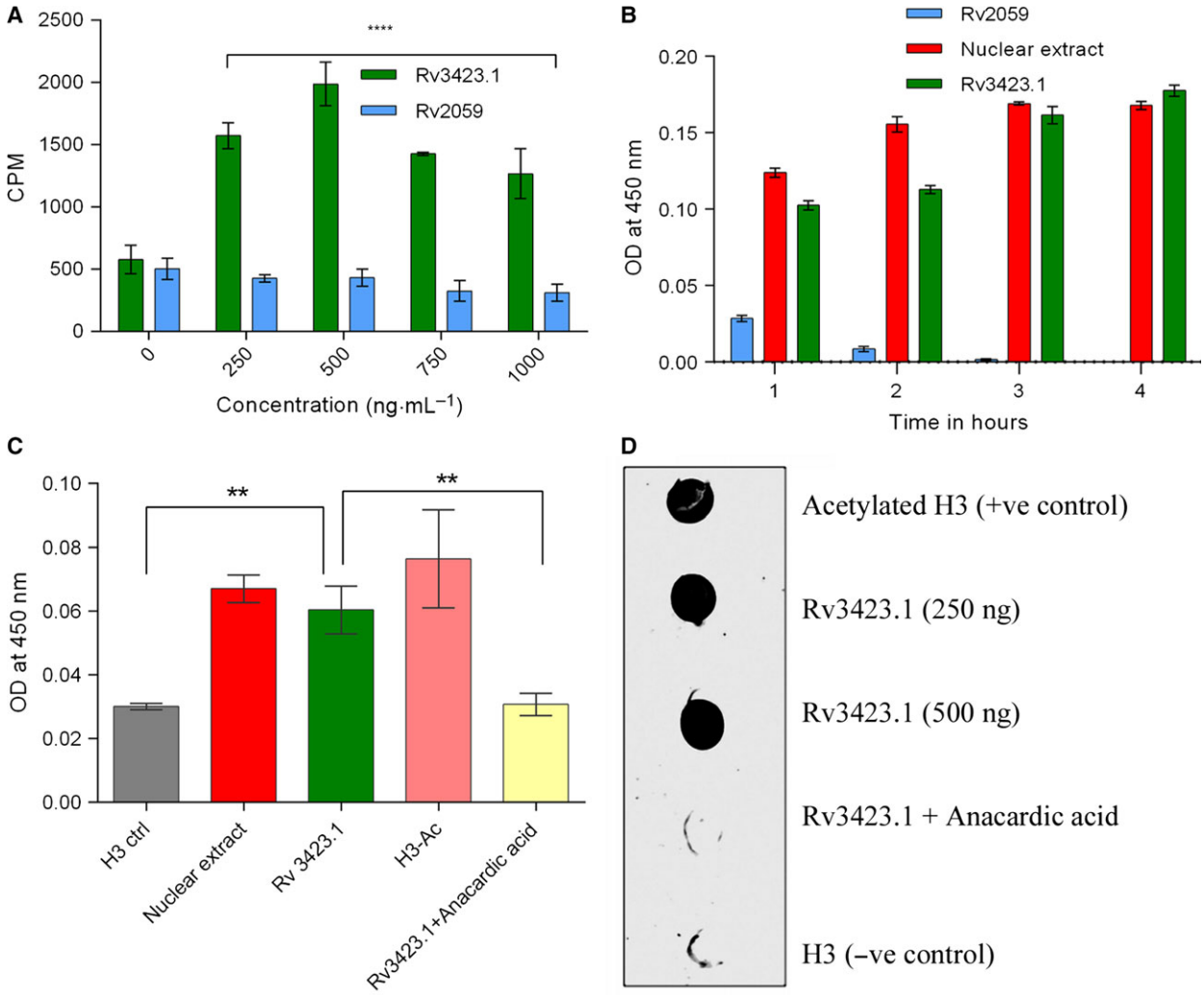


Fig. 5. Rv3423.1 is a histone acetyltransferase *in vitro*. (A) Filter binding assay. The ability of the recombinant Rv3423.1 protein to incorporate a ^{3}H-labeled acetyl group into calf thymus core histones using radioactive acetyl CoA was measured by liquid scintillation counting. The reaction mix was incubated at 37 °C for 30 min with various concentrations of the protein, ranging from 0 to 1000 ng·mL^{-1}, and the incorporated radioactivity trapped on the P81 filter was counted. A non-relevant recombinant mycobacterial protein (Rv2059) was used as a mock control. Values are means ± SEM from three independent experiments at each concentration. (B) Generation of NADH. HAT substrates provided in the HAT assay kit were incubated with Rv3423.1 (250 ng·mL^{-1}) at 37 °C for up to 4 h. HeLa cell nuclear extract was used as a positive control. A non-relevant mycobacterial protein (Rv2059) was used as a mock control. Absorbance values at 450 nm after background subtraction at four time points are plotted. Values are means ± SEM from three independent experiments. (C) ELISA. Histone H3 peptide substrate was incubated with Rv3423.1 protein (250 ng·mL^{-1}) at 30 °C for 90 min. Acetyl histone H3 was detected using antibody against acetyl H3 K9/K14. Non-acetyl histone H3, acetyl histone H3 and HeLa nuclear extract served as controls. Anacardic acid (10 μM) is a HAT inhibitor. Absorbance values at 450 nm after background subtractions are plotted. Values are means ± SEM from three independent experiments. (D) Immuno dot blot. The reaction mix was prepared as described for the ELISA, and incubated at 30 °C for 90 min. The mix was then spotted onto poly(vinylidene difluoride) membrane and air-dried. The dot blot was probed using antibody against acetyl H3 K9/K14 and alkaline phosphatase-conjugated secondary antibody. A representative image of three independent Western dot blots is shown. Anacardic acid (10 μM) was used to inhibit HAT activity.

associated with histone H3 (Fig. 8A) in infected cells. Non-specific IgG antibodies did not co-precipitate the Rv3423.1 protein. In a reverse analysis, we immunoprecipitated the proteins of the lysate of infected cells using antibodies against Rv3423.1, and the bound proteins were probed using antibody against histone H3. Antibodies against IgG were used as a non-specific control. The result confirmed that Rv3423.1 protein of MTB and histone H3 interact with each other in MTB-infected macrophages (Fig. 8B).

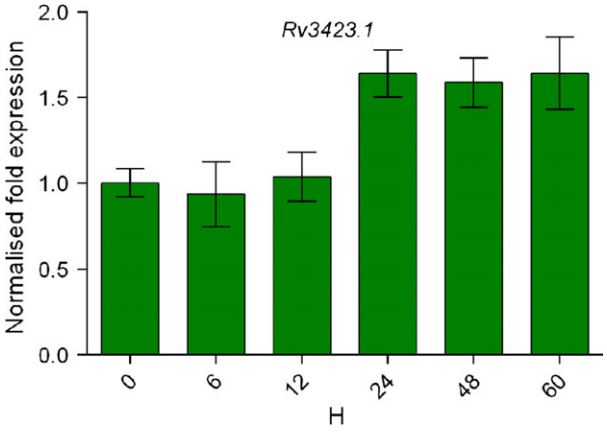


Fig. 6. *Rv3423.1* is up-regulated at 24 h after infection. THP-1-derived macrophages were infected with MTB H37Rv at a multiplicity of infection of 20 : 1. The levels of *Rv3423.1* transcript inside infected macrophages were analyzed by quantitative PCR at 0, 6, 12 and 24 h after infection. The endogenous gene *SigA* served as a control. Values are means ± SEM from three independent experiments performed in triplicate.

Rv3423.1 is secreted by MTB H37Rv but not H37Ra

We analyzed the cell lysates and culture filtrates from mid-log phase cultures of aerobically grown MTB virulent strain H37Rv and avirulent strain H37Ra for the presence of Rv3423.1 protein. Western blot analysis of the H37Rv strain showed expression of the protein in both the cell lysate as well as the culture filtrate (Fig. 9A). In contrast, we were unable to detect Rv3423.1 in the culture filtrates of H37Ra, even though the protein was present in the lysate (Fig. 9B). This suggests that H37Rv, but not H37Ra, is able to secrete the protein into the extracellular environment.

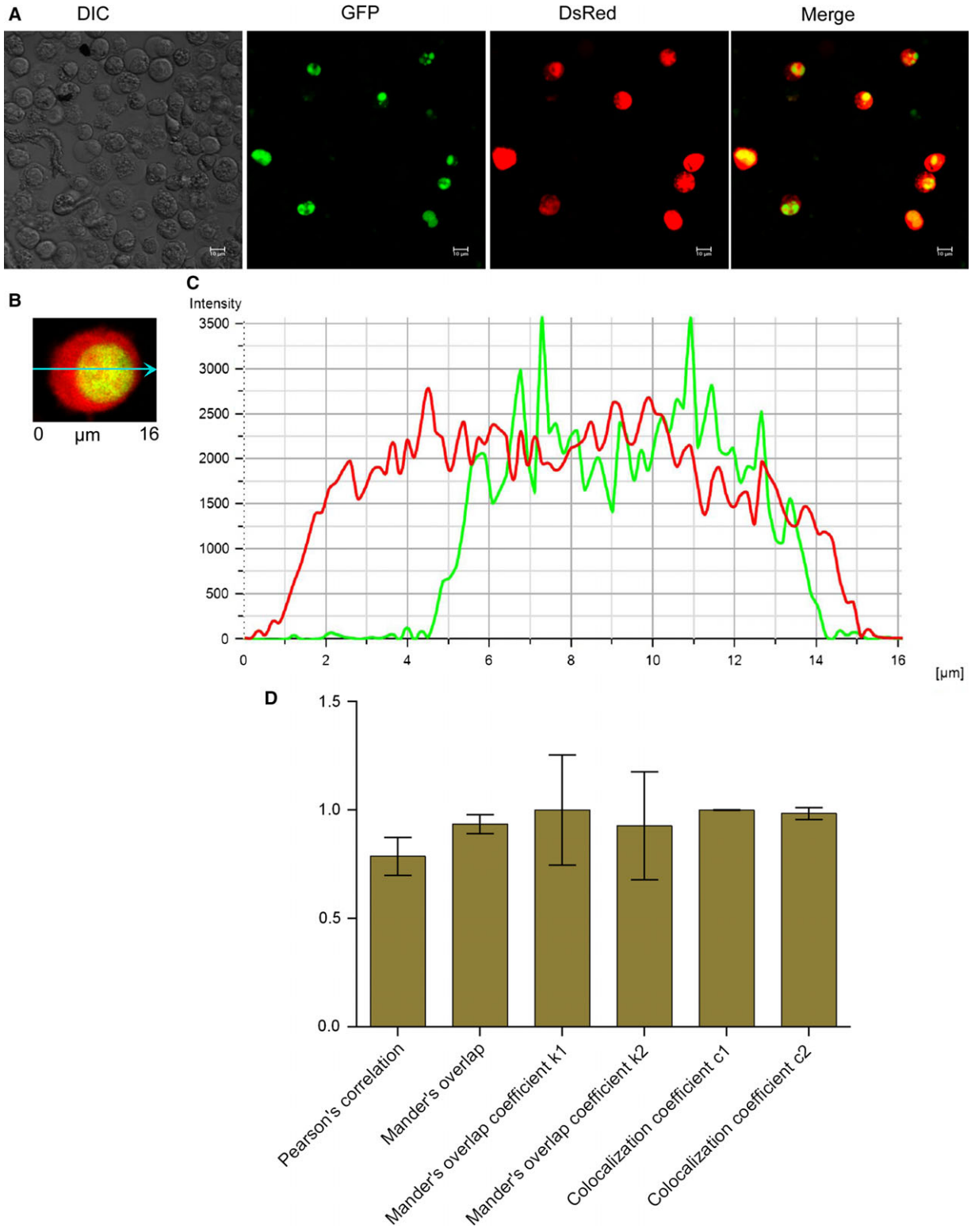
Rv3423.1 expression enhances the intracellular multiplication of recombinant *M. smegmatis*

Rv3423.1 was amplified from MTB H37Rv genomic DNA and cloned into mycobacterial constitutive expression vector pBEN. The recombinant plasmid was introduced into *M. smegmatis* by electroporation, and growth of the bacteria in Middlebrook 7H9 medium was monitored by measuring the attenuance at 600 nm. Bacteria carrying the pBEN vector served as the control. The growth patterns of both the bacteria were the same, and both reached an attenuance at 600 nm of 0.6 (mid-log phase) in approximately 21 h, and reached stationary phase at approximately 36 h (Fig. 10A). A Western blot using antibodies against Rv3423.1 confirmed the presence of the protein in the cell lysate (Fig. 10B). To analyze whether the protein provided any survival advantage to the intracellular bacteria, we infected THP-1-derived macrophages with *M. smegmatis* carrying Rv3423.1 in pBEN and *M. smegmatis* carrying the vector alone. We harvested and lysed the macrophages at regular intervals, and plated the lysate on Middlebrook 7H10 agar medium. During the initial phase, the number of recombinant *M. smegmatis* increased significantly, to reach a maximum (1.4×10^4) at 6 h (approximately two generations), while the number of non-recombinant bacteria increased only marginally (1.05×10^4, Fig. 10C). However, after 6 h, both the recombinant and non-recombinant bacteria started to die, and, in both cases, the number of viable bacteria decreased to almost zero at 48 h. Both recombinant and non-recombinant bacteria were completely cleared by the macrophages in 3 days.

Discussion

The primary objective of this study was to search for MTB effector proteins that enter the nucleus of the macrophage after infection and directly interact with the chromatin. We infected THP-1 human monocyte-derived macrophage cells with virulent MTB H37Rv, and immunoprecipitated the chromatin with antibody against histone H3. As we were interested in identifying proteins that interact with the chromatin rather than with the DNA, we subjected the chromatin-associated proteins to 2D gel electrophoresis, and looked for spots that were uniquely present in the infected samples. One of the prominent unique spots was subjected to MALDI-TOF analysis, and was identified as

Fig. 7. Rv3423.1 co-localizes with histones inside the macrophage nucleus and interacts with histone H3. (A) THP-1-derived macrophages were co-transfected with plasmids expressing eGFP-labeled histone H1 and DsRed-labeled Rv3423.1 by electroporation. The cells were visualized by confocal microscopy 12 h post-transfection. All images were acquired at 40× magnification. Images are representatives of at least three independent experiments. Scale bar = 1 μm. (B) Single cell showing GFP/DsRed co-localization. The arrow represents a 16 μm region of interest for which the intensities of red and green colors have been plotted in (C). (C) Fluorescence intensities show co-incident peaks from the red and green channels in (B), indicating co-localization. (D) Co-localization coeffiecients were calculated using Nis-Elements software; The relative numbers of co-localizing pixels in each channel are expressed compared with the total number of pixels above threshold. The value range is from 0 to 1, where 0 indicates no co-localization and 1 indicates that all pixels co-localize. All pixels above background are counted irrespective of their intensity. Results were obtained from images acquired from at least three separate co-transfection experiments.



MT3532.1, a hypothetical protein with a molecular mass of 8.4361 kDa in MTB CDC1551, whose counterpart in MTB H37Rv is Rv3423.1. Based on *in silico* analyses, three probable activities, namely acyl/acetyltransferase, oxidoreductase and nucleotide binding, were predicted for the protein. As the protein was

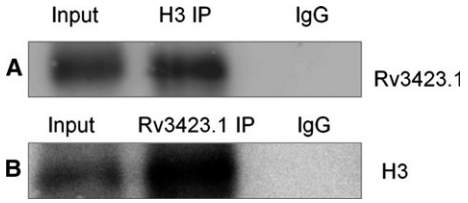


Fig. 8. Rv3423.1 associates with histone H3 *in vivo*. (A) THP-1-derived macrophages were infected with MTB H37Rv, and the cell extract was immunoprecipitated with histone H3 or rabbit non-specific IgG coupled to Protein A beads and immunoblotted with antibodies against Rv3423.1 (B) The cell extract was immunoprecipitated with antibodies against Rv3423.1 or rabbit non-specific IgG, and the immune complexes were subjected to Western blotting with antibody against histone H3.

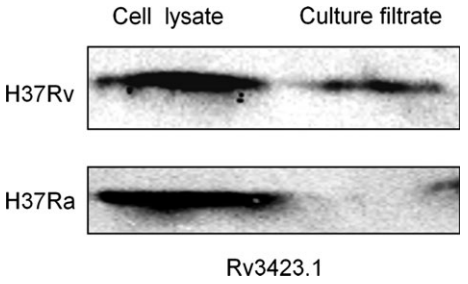


Fig. 9. Rv3423.1 is secreted by MTB H37Rv but not MTB H37Ra. *In vitro* expression and secretion of Rv3423.1 were probed using cell lysates and culture supernatants of MTB H37Rv and H37Ra. Total protein concentrations were determined using the Bradford assay, and 30 μg samples were subjected to SDS/PAGE and Western blot analysis. Detection was performed using antibodies against Rv3423.1.

originally isolated from the chromatin of infected cells, and as acetyltransferase activity was predicted as the most probable activity, we decided to test whether Rv3423.1 possesses HAT activity. Using four assays, i.e. a filter binding assay using radio-labeled acetyl CoA, ELISA, NADH release and Western blotting, we demonstrate that this protein is indeed a HAT. Using antibodies against variously acetylated lysines, we showed that Rv3423.1 acetylates histone H3 at K9 and K14. Post-translational modifications of histones such as acetylation of H3 and H4 are considered markers of gene expression, and are frequently associated with active transcription [21]. Acetylation of H3K9 and H3K14 occurs throughout the genome, and is shown to have high correlation with each other, and these modifications are referred to as 'transcription-linked' histone marks [22,23]. We showed that anacardic acid, a non-specific inhibitor of several known HATs due to its structural similarity to acetyl CoA, was able to inhibit the HAT activity of Rv3423.1, which further confirms the function of this protein.

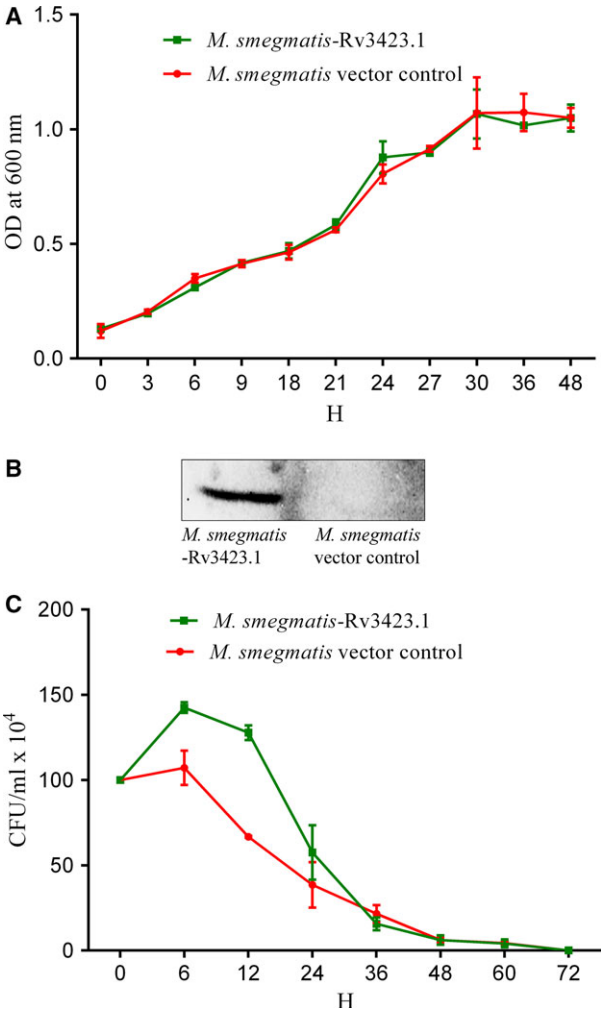


Fig. 10. Rv3423.1 confers distinct a intracellular survival advantage to recombinant *M. smegmatis* carrying the gene, but has no effect on growth under free-living conditions. (A) Recombinant and non-recombinant *M. smegmatis* were grown in Middlebrook 7H9 medium containing kanamycin (25 μg·mL^{-1}), and the attenuance at 600 nm was recorded every 3 h up to 48 h. The results for three independent experiments were similar. (B) Recombinant *M. smegmatis* expresses Rv3423.1. Recombinant and non-recombinant *M. smegmatis* were grown in Middlebrook 7H9 medium containing kanamycin (25 μg·mL^{-1}) for up to 24 h, and whole-cell lysates were probed for the presence of Rv3423.1 by Western blotting, with detection using antibodies against Rv3423.1. (C) THP-1-derived macrophages were infected with *M. smegmatis* mc^{2}155 transformed with recombinant pBEN vector carrying *Rv3423.1* or no insert as described in Experimental procedures. Aliquots of infected cells were lysed with water at the indicated time points up to 72 h. Serial dilutions from each time point were plated on Middlebrook 7H10 agar plates containing kanamycin (25 μg·mL^{-1}), and incubated for 3–4 days at 37 °C. The number of intracellular bacteria is shown as colony-forming units per ml (CFU·mL^{-1}) at various time points. Values are means ± SD from triplicate wells of a 96-well plate. Similar results were obtained for three independent experiments.

Until now, only very few MTB proteins have been shown to modulate host gene expression by chromatin remodeling/modification of the host genome. The 19 kDa lipoprotein LpqH and the 6 kDa early secreted antigen ESAT6 are known to inhibit interferon γ-induced chromatin remodeling of the class II transactivator CIITA that is involved in the class II MHC processing pathway [24–27]. However, this chromatin remodeling is an indirect process brought about through interferon γ-mediated signaling, and may be Toll like receptor-dependent or independent. Interestingly, two MTB proteins have been shown to be localized in the host nucleus and directly induce epigenetic modifications. Rv0256c enters the nucleus with the help of a bipartite nuclear localization signal, and inhibits NO production by specifically interacting with the inducible nitric oxide synthase (iNOS) promoter [28,29]. Rv2966c, a secreted DNA methyltransferase, has been shown to enter the nucleus of macrophages and bind specific DNA sequences, methylate cytosines in a non-CpG context, and interact with methylated histones H3 and H4 [30].

To demonstrate that Rv3423.1 enters the nucleus to bind to the chromatin, we tagged the MTB protein and histone H1 with fluorescent proteins, and analyzed the images using confocal microscopy. GFP-tagged histone H1 provided a uniform green background confined to the interior of the nucleus. However, RFP-tagged Rv3423.1 was found to be distributed throughout the inside of the cells. To confirm that the protein had entered the nucleus, we combined images from sequential focal planes (Z-stacking) (Movie S1) as well as co-localization analyses, which clearly showed that the protein did enter the nucleus.

Given the ability of Rv3423.1 to acetylate histone H3 *in vitro*, it is conceivable that the protein physically interacts with histone H3 inside the nucleus to exert its activity. Co-immunoprecipitation experiments clearly showed that Rv3423.1 interacts with histone H3. This interaction and subsequent transcriptional activation may modulate pathways that regulate host immune responses after MTB infection, and aid bacterial persistence/survival. To the best of our knowledge, this is the first evidence of a direct interaction between a bacterial HAT enzyme and the host chromatin.

It is interesting to note that Rv3423.1 lacks any signal peptide for secretion from the bacterium. Virulent MTB is known to secrete many small, highly immunogenic proteins that lack classical signal sequences, the most studied being ESAT6 [31]. This protein is a major substrate of the ESX-1 type VII secretion system, and is known to induce rupture of phagosomal membranes to allow transport of many mycobacterial effector proteins into the cytosol [31–34]. We tested the culture supernatants of virulent as well as avirulent strains of MTB for the presence of Rv3423.1, and our findings suggest that this protein is synthesized by both virulent and avirulent MTB but is secreted only by virulent MTB. It has been reported that the ESX-1 type VII secretion system is impaired in avirulent H37Ra [35]. Therefore, it is possible that Rv3423.1 reaches the host cytoplasm after phagosomal rupture, from where it is directed into the nucleus.

Even though the protein was found inside the nucleus, there is no evidence for any known nuclear localization signal in the sequence. However, this is not surprising as very few bacterial proteins possess a well-defined classical monopartite or bipartite nuclear localization signal [36,37]. There are also several classes of proteins with a non-classical nuclear localization signal that do not contain well-conserved sequences [38]. It is likely that Rv3423.1, being a very small protein (8.4 kDa), enters the nucleus on its own or utilizes host proteins to hitchhike into the nucleus. As Rv3423.1 is localized on the chromatin, we were expecting it to be a 'recruiter' of host chromatin remodeling complexes or chromatin-modifying enzymes. To our surprise, however, it turned out to be a histone-modifying enzyme in itself, rather than being a facilitator that recruits host machinery to the chromatin.

The role of nucleomodulins in infection, virulence and diseases caused by intracellular pathogens has gained attention recently. In this study, we have demonstrated that the MTB protein Rv3423.1 is a HAT, enters the nucleus of macrophages, binds to the chromatin, and acetylates H3 at the K9/14 positions. Therefore, in all likelihood, this MTB HAT may regulate transcription of host genes whose products may help the pathogen survive and multiply within the macrophage, and hence may play a role in the disease progression. At present, it is not clear whether this protein helps the intracellular MTB directly by playing a role in enhancing the survival/virulence, or whether it is indirectly involved in suppressing the immune response of the host. The role of this protein inside the bacterium is also not known. As it is a lysine acetyltransferase, it may also acetylate bacterial proteins that have yet to be identified. Although there is currently no evidence of any protein with significant homology to Rv3423.1, it would be interesting to look for proteins with similar functions in other intracellular pathogens. Because of its probable role in the regulation of host genes, MTB HAT may constitute a novel drug target or vaccine candidate. It is hoped that our

study will stimulate a search for novel nucleomodulins that form a new class of targets for therapeutic intervention.

Experimental procedures

Infection of THP-1-derived macrophage cells with MTB H37Rv

Human acute monocytic leukemia cell lines (THP-1) were obtained from the National Centre for Cell Sciences (Pune, India). Cells were routinely cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 25 mM HEPES, 1.5 $g \cdot L^{-1}$ sodium bicarbonate, by incubation at 37 °C in the presence of 5% CO_2. Phorbol 12-myristate 13-acetate (Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 20 $ng \cdot mL^{-1}$ was added to the cells (3×10^6 per mL) for differentiation into macrophages. The virulent *M. tuberculosis* strain H37Rv, avirulent strain H37Ra and non-pathogenic *M. smegmatis* mc^2155 were used in the study. All experiments with MTB were performed in a Biosafety Level 3 (Standard usage) facility at Rajiv Gandhi Centre for Biotechnology for safe handling of the bacilli. Bacteria were grown in Middlebrook 7H9 broth (Difco, Detroit, MI, USA) supplemented with 10% oleate-albumin-dextrose-catalase (Becton Dickinson, Franklin Lakes, NJ, USA), 0.4% glycerol and 0.05% Tween-80. The 1 mL culture was sonicated (20 pulses) in a wave sonicator (Elmasonic, Singen, Germany) to disrupt the clumps. The suspension was dispersed by aspirating five times through a 24-gauge needle, followed by ten times through a 30-gauge needle, and was allowed to stand for 5 min. The upper half of the suspension was then used for the experiments. Bacteria were counted by measuring the attenuance at a wavelength of 600 nm (0.2 units corresponds to approximately 300×10^6 bacteria per ml), and then harvested by centrifugation at 1000 ***g*** for 10 min, washed with RPMI −1640 for 5 min at 1000 *g* at 25 °C and resuspended in the same medium. Further dilutions were made to obtain a multiplicity of infection of 20 : 1 in RPMI medium. Bacterial suspension was added to the macrophages in fresh RPMI −1640 medium without fetal bovine serum. Internalization of the bacteria was allowed to proceed for 3 h at 37 °C and 5% CO_2 in the humidified incubator. Uninfected cells served as the negative control. After internalization, the macrophages were washed three times with PBS containing gentamycin ($2.5\ \mu g \cdot mL^{-1}$) to remove non-internalized bacteria. Fresh medium was added to the cells before incubation at 37 °C for 24 h.

Visualization of infected macrophages

THP-1-derived macrophages were grown on glass coverslips in 24-well plates (Tarsons, Kolkata, India) at a density of 2×10^5 cells per well. The cells were infected with mid-log phase *M. tuberculosis* H37Rv expressing GFP (pBEN-GFP plasmid) as described above. After incubation for 24 h, the cells were washed briefly in PBS and then fixed in PBS containing 4% paraformaldehyde for 15 min at 25 °C. The cells were then washed twice with ice-cold PBS. For counter-staining, the cells were incubated with 4′,6-diamidino-2 phenylindole (DAPI) ($0.1\ \mu g \cdot mL^{-1}$) in PBS/Tween for 5 min. The cells were rinsed with PBS and smeared onto glass slides and air-dried for 5 min. Coverslips were mounted on the smears using a drop of 50% glycerol. The coverslips were sealed with di-*n*-butyl phthalate in xylene (Sigma-Aldrich) mount to prevent drying and movement under the microscope. The cells were viewed under a confocal laser scanning microscope (Nikon Eclipse Ti-S, Tokyo, Japan) at 40× magnification. Images of bacteria were captured via the FITC channel (excitation 490 nm, emission 525 nm), and the images of macrophage nuclei were captured in the DAPI channel (excitation 350 nm, emission 470 nm).

Viability of intracellular mycobacteria

Infection of macrophages with MTB was performed as described above. Heat-killed bacteria were prepared by heating the culture at 80 °C for 2 h in a heating block. Total RNA was isolated using an RNAspin mini kit (GE Healthcare, Piscataway, NJ, USA) according to the manufacturer's instructions. cDNA was prepared from 0.5 μg total RNA using a GoScript™ reverse transcription system (Promega, Madison, WI, USA). The relative expression of *CORO1A* was evaluated by quantitative real-time PCR using a SYBR Green Supermix kit (Bio-Rad, Hercules, CA, USA). The PCR conditions were as follows: 61 °C for 20 min, 95 °C for 4 min, 95 °C for 10 s, 63 °C for 30 s, 72 °C for 30 s (30 cycles), then 55–95 °C with a 0.5 °C increment (melting curve). Each sample was normalized to the level of *GAPDH* transcript.

Modified chromatin immunoprecipitation (mChIP)

The cells were treated with RPMI medium containing 1% formaldehyde for 15 min at 25 °C to crosslink DNA and proteins. Chromatin was prepared using a Bioruptor wave sonicator (Diagenode, Seraing, Belgium). The pulse conditions were 30 s on, 30 s off for 25 cycles at 4 °C. Whole-cell lysate was prepared using 1 mL lysis buffer comprising 50 mM HEPES/KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS and protease inhibitor cocktail (Sigma-Aldrich). The lysate was incubated with antibody against histone H3 (Abcam, Cambridge, UK) at 4 °C overnight on a rotating platform at 40 rpm. Protein A Sepharose beads (Sigma-Aldrich) were used for immunoprecipitation. Briefly, the overnight sample was mixed with pre-blocked protein A Sepharose beads, and the immunoprecipitate was pulled down by incubating on a platform rotating at 40 rpm for

2 h at 4 °C. The beads were then removed by centrifugation at 1500 ***g*** for 5 min, and then washed three times with 1.0 mL of ice-cold ChIP washing buffer (10 mM Tris/Cl pH 8.0, 0.25 M LiCl, 0.05% IGEPAL CA-630 (Sigma- Aldrich) and 0.5% sodium deoxycholate) at 1500 ***g*** at 4 °C. The beads were then resuspended in ice-cold 2D rehydration/lysis buffer, and incubated at 4 °C for 4 h on a rotating platform. The beads were spun down at 1000 ***g*** at 4 °C, and the supernatant was collected.

2D gel electrophoresis

The immunoprecipitates from infected and non-infected cells were subjected to iso-electric focusing in a Multiphor II electrophoresis system (GE Healthcare). The protein sample (100 μg) was mixed with rehydration buffer to make up the final volume to 125 μL, and applied to IPG strips (linear pH gradient 3-10; GE Healthcare). The iso-electric focusing conditions used were 500 V for 2 h, 1500 V for 1 h, and 3500 V for 5 h, successively. Once the run was completed, the strips were treated with equilibration solution (75 mM Tris/HCl pH 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS and 0.002% bromophenol containing 65 mM dithiothreitol), and then incubated for 15 min in equilibration solution containing 135 mM iodoacetamide. Finally, the strips were briefly rinsed with 5.0 mL SDS running buffer (25mM Tris/HCl pH 8.3, 250mM glycine, 0.1% w/v SDS) before proceeding to the second dimension of separation in an 10% SDS/PAGE gel, and subsequently stained with a mass spectrometry-compatible ProteoSilver staining kit (Sigma-Aldrich).

MALDI-TOF mass spectrometry

The proteins in the gels of infected and non-infected samples were compared manually, and the spots that appeared on the gels containing infected samples but did not appear on the gels containing non-infected samples were identified. One such spot (S1) was excised from the gel and stored in 1% acetic acid. The spot was digested using Tosyl phenylalanyl chloromethyl ketone-treated trypsin (Applied Biosystems, Foster City, CA, USA), purified using a trypsin profile IGD kit (Sigma-Aldrich), and then loaded onto an Applied Biosystems 4800 MALDI TOF-TOF system to generate the spectrum. Spectral acquisition was performed using 4000 SERIES DATA EXPLORER software (Applied Biosystems), and data analysis was performed using GPS EXPLORER (Applied Biosystems) and Mascot. The most relevant hit was identified and submitted to BLAST analysis, and the protein/gene was identified.

Function prediction through bioinformatics analysis

The query Rv3423.1 sequence (UNIPROT accession ID P9WKY5, 76 amino acids long) was searched against the Protein Data Bank using FASTA from the Sequence Annotated by Structure (SAS) tool [16], in order to find templates that were homologous to the query. Sequence-based function predictions were performed using the PFP (http://kiharalab.org/web/pfp.php) and ProtFun (http://www.cbs.dtu.dk/services/ProtFun/) tools. The query sequence was checked for short sequence motifs and patterns using the Prosite (prosite.expasy.org/) and Prints (http://www.bioinf.manchester.ac.uk/dbbrowser/PRINTS/index.php) databases. Further, fold recognition was performed to identify fold templates using the @TOME (version 2.1) database [17]. Modeller (version 9) (https://salilab.org/modeller/) was used to build 3D structure models for the query. Pymol (www.pymol.org) was used for structure visualization and generation of images.

Cloning, expression and purification of Rv3423.1

The *Rv3423.1* gene was amplified from MTB H37Rv genomic DNA using specific primers (Table 1). The PCR conditions comprised a DNA denaturation step at 95 °C for 5 min, followed by 25 cycles at 95 °C for 30 s, 65 °C for 30 s and 72 °C for 45 s, and a final elongation step at 72 °C for 5 min. The 231 bp PCR product was purified using an Illustra GFX PCR DNA and gel band purification kit (GE Healthcare). The PCR product was cloned into a linearized Champion pET SUMO expression vector (Invitrogen). To do this, the PCR product and vector were mixed in an equimolar ratio and ligated overnight using T4 DNA ligase at 20 °C. The resulting construct was then transformed into One Shot® Mach1™ T1 Phage-Resistant Chemically Competent *E. coli* cells provided in the Champion pET SUMO expression kit (Invitrogen). Positive clone selection was based on kanamycin resistance. The identity of the construct was confirmed by DNA sequencing. The Rv3423.1–SUMO fusion construct was then transformed into One Shot® BL21(DE3) chemically Competent *E. coli* cells supplied in the above-mentioned kit. Isopropyl-β-D-thiogalactopyranoside was then added to a final concentration of 1 mM to induce expression of the Rv3423.1–SUMO fusion. The

Table 1. List of primers used in the study.

Primer name	Sequence (5′→3′)
Human quantitative PCR primers	
GAPDH F	AATGAAGGGGTCATTGATGG
GAPDH R	AAGGTGAAGGTCGGAGTCAA
MTB *Rv3423.1* whole-gene primers	
Rv3423.1 F	GTGTCGTTTGGCTTTCCGA
Rv3423.1 R	CTACGCTCCCGTCAATTCATG
MTB quantitative PCR primers	
SigA F	CTCGGTTCGCGCCTACCTCA
SigA R	GCGCTCGCTAAGCTCGGTCA
Rv3423.1 qF	GCTTTCCGACATTCTCCC
Rv3423.1 qR	AACCACCGAGCCACAG

over-expressed fusion protein was purified under native conditions using Ni-NTA agarose resin (Thermo Scientific, Waltham, MA, USA). The purified protein was concentrated and buffer-exchanged using Amicon Ultra-15 centrifugal filter units (3 kDa molecular weight cut-off, Millipore, Billerica, MA, USA). To cleave the fusion tag, SUMO protease (Invitrogen) was added to the concentrated fusion protein to a final concentration of 0.5 $U \cdot \mu g^{-1}$. The cleavage reaction was incubated at 16 °C, 40 rpm overnight. The native protein was separated from the 6× His-tagged SUMO protein by applying the cleaved sample to an Ni-NTA spin column (Promega). The cleaved Rv3423.1 protein was recovered from the flow-through fraction of the Ni-NTA spin column, and was then concentrated and buffer exchanged using Amicon Ultra-15 centrifugal filter units, and its identity was confirmed using MALDI-TOF mass spectroscopy.

Histone acetyltransferase assays

Filter binding assay

A filter binding assay was performed using P81 phosphocellulose filter plates (Millipore). Microcentrifuge tubes containing 30 μL reaction mixture comprising 200 μg calf thymus histones (Sigma-Aldrich), HAT assay buffer (50 mM Tris/HCl pH 8.0, 50 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol, 5% v/v glycerol, 1 mM phenylmethanesulfonyl fluoride and 10 mM sodium butyrate), 2 μL of ^{3}H-acetyl CoA (0.1 $mCi \cdot mL^{-1}$; Perkin Elmer, Waltham, MA, USA) and purified Rv3423.1 protein (0–1000 ng) were incubated at 30 °C for 30 min. Rv2059, a recombinant MTB protein purified under similar conditions, was used as a mock control. Each sample was transferred into the P-81 cellulose filter well and incubated for 10 min. The wells were washed five times with 100 μL of 50 mM sodium bicarbonate (pH 9.0) for 5 min each. Chilled acetone (50 μL) was added to the wells and incubated overnight at 25 °C. Liquid scintillation cocktail (80 μL) was added, mixed and incubated at 25 °C for 4 h. The ^{3}H radiation was recorded as 1 min counts using a scintillation counter (Perkin Elmer).

Detection of NADH generation

A histone acetyl transferase activity assay kit (Abcam) was used according to the manufacturer's instructions. Increasing concentrations of purified recombinant Rv3423.1 protein in PBS were added for each assay in a 96-well plate at a final volume of 40 μL. To each well, 68 μL assay mix containing 50 μL of 2× HAT assay buffer, 5 μL HAT substrate I, 5 μL HAT substrate II and 8 μL NADH-generating enzyme were added. The plate was then incubated at 37 °C for 2 h, and samples were read in an iMark microplate absorbance reader (Bio-Rad) at 450 nm. A nuclear extract (50 μg) from HeLa cells provided in the HAT assay kit served as a positive control. The recombinant protein (Rv2059) from MTB was used as the mock control. The background reading for the negative control (buffer and substrate) was subtracted from all the readings.

Determination of HAT enzyme activity by ELISA

The HAT enzyme activity was determined using an ELISA HAT assay kit (Upstate Biotechnology, Lake Placid, NY, USA) according to the manufacturer's instructions. Briefly, streptavidin-coated strip plates were incubated with biotinylated histone H3 peptide. Rv3423.1 protein (250 $ng \cdot mL^{-1}$) was used for the assay. Acetyl histone H3 and HeLa nuclear extract were used as positive controls. Anacardic acid, a HAT inhibitor (10 μM), was used to further confirm the HAT activity. The HAT reaction was carried out in the HAT buffer provided in the assay kit, and incubated at 30 °C, and the strip plates were washed and probed for acetylation of the histone peptides using an antibody against α-acetyl lysine provided in the kit for 90 min. Then the wells were washed with HAT wash buffer, and tetramethyl benzidine substrate mixture was added to each well before incubation in the dark for 10 min. Finally, 0.1 M sulfuric acid was added to stop the reaction, and the enzyme activity was measured on a microplate reader (excitation 450 nm, emission 570 nm), and expressed as ng acetylated histone per mg enzyme. The 450 nm absorbance values were subtracted from the 570 nm values to remove any well-to-well variation.

Immuno dot blot

The HAT reaction cocktail was prepared as described in the previous experiment. The reaction was incubated for 2 h at 30 °C. The reaction mix (20 μL) was spotted onto activated poly(vinylidene difluoride) membrane, and then allowed to dry for 30 min. The dried membrane was incubated in blocking solution (5% BSA in Tris-buffered Saline with Tween 20) for 1 h. The membrane was then incubated with antibody against acetyl H3 K9/K14 (Abcam) for 2 h, followed by alkaline phosphatase-conjugated secondary antibody (Sigma-Aldrich) for 1 h, then incubated in nitro-blue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3′-indolyphosphate p-toluidine salt (BCIP) substrate solution until spots were visible, and an image was captured using VersaDoc imager (Bio-Rad, Hercules, CA, USA).

Nuclear localization of Rv3423.1

The *Rv3423.1* gene was cloned into mammalian expression vector pDsRed-Express-C1 (Invitrogen) as a C-terminal fusion protein. The histone H1–eGFP fusion construct was used as the nuclear marker. THP-1 monocytes (10^6 cells/100 μL) were mixed with 0.5 μg plasmid DNA from both constructs, and electroporated using Nucleofector (Lonza, Germany). After electroporation, 500 μL of pre-warmed

culture medium containing 20 ng·mL^{-1} phorbol 12-myristate 13-acetate was added to the cells, and the cell suspension was transferred into 96-well glass-bottomed plates (Perkin Elmer). The cells were incubated at 37 °C in a CO_2 incubator for 12 h post-transfection before assaying for gene expression. Co-localization of histone H1–eGFP and Rv3423.1–DsRed was analyzed by confocal fluorescence microscopy using a Nikon Eclipse Ti-S microscope. Image capture, Z-stacking and calculation of co-localization coefficients were performed using NIS ELEMENTS software (Nikon) under identical exposure and gain settings.

Generation of polyclonal antibodies and co-immunoprecipitation

Native Rv3423.1 protein was purified as described above, and used to generate polyclonal antibodies. Procedures involving animals were performed in accordance with the guidelines of the Institute Animal Ethics Committee, Rajiv Gandhi Centre for Biotechnology (Thiruvananthapuram, India). The primary dose (300 μg protein in Titermax Gold adjuvant; Sigma-Aldrich) was administered intradermally into the interscapular area at multiple sites (10–20 μL per site) of two New Zealand White female rabbits, followed by a booster injection on the 28th day. Serum was collected on the 60th day, and IgG was purified from the serum using AminoLink Plus Resin (Thermo Scientific) according to the manufacturer's instructions. The *in vivo* interaction between Rv3423.1 and histone H3 was analyzed by co-immunoprecipitation experiments. THP-1 macrophages infected with MTB were treated with lysis buffer (50 mM Tris pH 8.0, 100 mM NaCl, 20 mM NaF, 50 mM KH_2PO_4, 1% Triton X-100, 10% glycerol and 0.1 mM dithiothreitol) containing protease inhibitor cocktail (Sigma-Aldrich). Antibodies against histone H3 (Abcam) or Rv3423.1 (raised in-house) were captured using Protein A Sepharose beads. Lysates (500 μg/500 μL) were sonicated, clarified, and resuspended in binding buffer (50 mM Tris/HCl pH 8.0, 100 mM KCl, 0.1 mM EDTA, 0.2% NP-40, 2.5% glycerol and 1 mM dithiothreitol). The immune complexes bound to the beads were collected and washed using wash buffer (100 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.5% NP-40), and bound proteins on the beads were resuspended in SDS/PAGE sample buffer. The immunoprecipitates were resolved by SDS/PAGE and transferred to polyvinylidene difluoride membranes (GE Healthcare), which were probed with the purified antibodies against Rv3423.1 or antibody against histone H3 and developed using a chemiluminescence detection kit (GE Healthcare).

Quantitative PCR to detect intracellular Rv3423.1

Infection of macrophages with *M. tuberculosis*, total RNA extraction and cDNA preparation were performed as described above. The relative expression of *Rv3423.1* was evaluated by quantitative real-time PCR using SYBR Green Supermix kit (Bio-Rad). The PCR conditions were as follows: 60 °C for 20 min, 95 °C for 4 min, 95 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s (30 cycles), then 55–95 °C with a 0.5 °C increment (melting curve). Fold changes in expression levels at various time points were calculated relative to the endogenous gene *Sigma Factor A (SigA)* using the $2^{-\Delta\Delta CT}$ method [39].

Preparation of MTB cell lysates and culture filtrates for Western blot analysis

MTB H37Rv and H37Ra strains were grown in liquid medium for 7 days as described above, and the cells were harvested by centrifugation at 1000 ***g*** for 10 min at 25 °C. The culture supernatant was recovered after filtration through 0.22 μm pore-size filters (Millipore), and concentrated by trichloroacetic acid precipitation. The cell pellet was washed and resuspended in PBS followed by rupture using 0.1 mm glass beads (Sigma-Aldrich) for three 1 min pulses at 2500 rpm in a mini bead beater (Biospec, Bartlesville, OK, USA). The lysate was recovered after centrifugation at 13 000 ***g*** for 30 min. Total protein concentrations were determined by the Bradford assay (Sigma-Aldrich), and 30 μg samples were subjected to SDS/PAGE. Western blot analysis was performed using polyclonal antibodies against Rv3423.1 and developed using a chemiluminescence detection kit.

In vitro growth and intracellular survival of recombinant *M. smegmatis* carrying the *Rv3423.1* gene

The *Rv3423.1* gene was amplified from MTB H37Rv genomic DNA as described above except that the forward and reverse primers contained *Bam*HI and *Hin*dIII restriction sites, respectively, and an N-terminal FLAG coding sequence to assist protein detection. The gene was cloned into mycobacterial constitutive expression vector pBEN using the above-mentioned restriction sites in the mutiple cloning site downstream of the mycobacterial *hsp60* promoter. Electroporation of *M. smegmatis* mc^{2}155 competent cells was performed using a Gene Pulser electroporation system (Bio-Rad) at a voltage of 1.5 kV, resistance of 200 Ω and capacitance of 25 μF. The recombinant cells were selected on Middlebrook 7H10 agar plates containing 10% oleate-albumin-dextrose-catalase with 25 μg·mL^{-1} kanamycin. *In vitro* growth of recombinant *M. smegmatis* and *M. smegmatis* carrying the control vector was determined by measuring the attenuance at 600 nm. To analyze the intracellular survival in macrophages, THP-1 cells were seeded in 90 mm tissue culture-grade Petri plates at a seeding density of 2 million cells per plate in antibiotic-free RPMI-1640 medium, and induced using 20 ng·mL^{-1} phorbol 12-myristate 13-acetate

24 h prior to infection. Bacterial suspension was added to the macrophage cells at a multiplicity of infection of 20 : 1. Internalization was allowed to proceed for 1 h at 37 °C and 5% CO_2. After phagocytosis, each well was washed three times with PBS containing gentamycin (10 $\mu g \cdot mL^{-1}$) to remove extracellular bacteria. Fresh medium was added to the wells and incubated at 37 °C for 72 h. Every 6 h, the medium was removed, macrophage cells were lysed with 100 μL distilled water, and the lysate was serially diluted and plated onto plates containing LB agar (HiMedia, Mumbai, India) with 25 $\mu g \cdot mL^{-1}$ kanamycin, and incubated at 37 °C for 3 days. The experiment was performed in triplicate, and a survival curve was plotted.

Acknowledgements

We thank M. Mayadevi and P.L. Reshma, Neurobiology Group, Rajiv Gandhi Centre for Biotechnology, for help with performing 2D gel electrophoresis. The pBEN vector was a kind gift from Lalitha Ramakrishnan, Department of Microbiology, University of Washington, Seattle, WA, USA. The histone–eGFP construct was a kind gift from Shiv Shankar, Department of Biochemistry, Biophysics and Bioinformatics, National Center for Biological Sciences, Bangalore, India. R.A.K. acknowledges financial support from the Department of Biotechnology, Government of India. L.J. and R.L.G. acknowledge receipt of a senior research fellowship from the Council for Scientific and Industrial Research, Government of India. R.R. acknowledges financial support from the University Grants Commission, Government of India. A.C. and S.R. are grateful for the financial support provided by an INSPIRE fellowship from the Department of Science and Technology, Government of India.

Author contributions

R.A.K. and L.J. designed the research; L.J., R.R., R.L.G., A.C. and S.R. performed the research; R.B., O.R.V and N.C. contributed new reagents or analytical tools; L.J., R.R., R.B., R.L.G., S.M. and R.A.K. analyzed the data; L.J. and R.A.K. wrote the paper.

References

1 World Health Organization (2015) Global Tuberculosis Report 2015. World Health Organization, Geneva, Switzerland.
2 Berger SL (2007) The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412.
3 Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**, 693–705.
4 Lebreton A, Lakisic G, Job V, Fritsch L, Tham TN, Camejo A, Matteï P-J, Regnault B, Nahori MA & Cabanes D (2011) A bacterial protein targets the BAHD1 chromatin complex to stimulate type III interferon response. *Science* **331**, 1319–1321.
5 Hamon MA, Batsché E, Régnault B, Tham TN, Seveau S, Muchardt C & Cossart P (2007) Histone modifications induced by a family of bacterial toxins. *Proc Natl Acad Sci USA* **104**, 13467–13472.
6 Pero R, Peluso S, Angrisano T, Tuccillo C, Sacchetti S, Keller S, Tomaiuolo R, Bruni CB, Lembo F & Chiariotti L (2011) Chromatin and DNA methylation dynamics of *Helicobacter pylori*-induced COX-2 activation. *Int J Med Microbiol* **301**, 140–149.
7 Garcia-Garcia JC, Rennoll-Bankert KE, Pelly S, Milstone AM & Dumler JS (2009) Silencing of host cell CYBB gene expression by the nuclear effector AnkA of the intracellular pathogen *Anaplasma phagocytophilum*. *Infect Immun* **77**, 2385–2391.
8 Bierne H & Cossart P (2012) When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14**, 622–633.
9 Liu PT & Modlin RL (2008) Human macrophage host defense against *Mycobacterium tuberculosis*. *Curr Opin Immunol* **20**, 371–376.
10 Pieters J (2008) *Mycobacterium tuberculosis* and the macrophage: maintaining a balance. *Cell Host Microbe* **3**, 399–407.
11 Gómez-Díaz E, Jordà M, Peinado MA & Rivero A (2012) Epigenetics of host–pathogen interactions: the road ahead and the road behind. *PLoS Pathog* **8**, e1003007.
12 Hamon MA & Cossart P (2008) Histone modifications and chromatin remodeling during bacterial infections. *Cell Host Microbe* **4**, 100–109.
13 Seto S, Tsujimura K & Koide Y (2012) Coronin–1a inhibits autophagosome formation around *Mycobacterium tuberculosis*-containing phagosomes and assists mycobacterial survival in macrophages. *Cell Microbiol* **14**, 710–727.
14 Ricke RM & Bielinsky A-K (2005) Easy detection of chromatin binding proteins by the histone association assay. *Biol Proced Online* **7**, 60–69.
15 Lambert J-P, Mitchell L, Rudner A, Baetz K & Figeys D (2009) A novel proteomics approach for the discovery of chromatin-associated protein networks. *Mol Cell Proteomics* **8**, 870–882.
16 Milburn D, Laskowski RA & Thornton JM (1998) Sequences annotated by structure: a tool to facilitate the use of structural information in sequence analysis. *Protein Eng* **11**, 855–859.
17 Pons J-L & Labesse G (2009) @TOME-2: a new pipeline for comparative modeling of protein–ligand complexes. *Nucleic Acids Res* **37**, W485–W491.
18 Podobnik M, Siddiqui N, Rebolj K, Nambi S, Merzel F & Visweswariah SS (2014) Allostery and

conformational dynamics in cAMP-binding acyltransferases. *J Biol Chem* **289**, 16588–16600.
19 Vetting MW, Magnet S, Nieves E, Roderick SL & Blanchard JS (2004) A bacterial acetyltransferase capable of regioselective N-acetylation of antibiotics and histones. *Chem Biol* **11**, 565–573.
20 Vetting MW, de Carvalho LPS, Yu M, Hegde SS, Magnet S, Roderick SL & Blanchard JS (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* **433**, 212–226.
21 Roh T-Y, Cuddapah S & Zhao K (2005) Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev* **19**, 542–552.
22 Gupta AP, Chin WH, Zhu L, Mok S, Luah Y-H, Lim E-H & Bozdech Z (2013) Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathog* **9**, e1003170.
23 Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H & Tora L (2012) H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genom* **13**, 424.
24 Pai RK, Pennini ME, Tobian AA, Canaday DH, Boom WH & Harding CV (2004) Prolonged Toll-like receptor signaling by *Mycobacterium tuberculosis* and its 19-kilodalton lipoprotein inhibits γ interferon-induced regulation of selected genes in macrophages. *Infect Immun* **72**, 6603–6614.
25 Pennini ME, Liu Y, Yang J, Croniger CM, Boom WH & Harding CV (2007) CCAAT/enhancer-binding protein β and δ binding to CIITA promoters is associated with the inhibition of CIITA expression in response to *Mycobacterium tuberculosis* 19-kDa lipoprotein. *J Immunol* **179**, 6910–6918.
26 Pennini ME, Pai RK, Schultz DC, Boom WH & Harding CV (2006) *Mycobacterium tuberculosis* 19-kDa lipoprotein inhibits IFN-γ-induced chromatin remodeling of MHC2TA by TLR2 and MAPK signaling. *J Immunol* **176**, 4323–4330.
27 Kumar P, Agarwal R, Siddiqui I, Vora H, Das G & Sharma P (2011) ESAT6 differentially inhibits IFN-γ-inducible class II transactivator isoforms in both a TLR2-dependent and -independent manner. *Immunol Cell Biol* **90**, 411–420.
28 Bhat KH, Das A, Srikantam A & Mukhopadhyay S (2013) PPE2 protein of *Mycobacterium tuberculosis* may inhibit nitric oxide in activated macrophages. *Ann N Y Acad Sci* **1283**, 97–101.
29 Mukhopadhyay S, Bhat KH & Khan N (2013) Inhibitors of Rv0256c [WWW document]. URL http://www.google.com/patents/US20100129809 [accesssed 20 October 2015].
30 Sharma G, Upadhyay S, Srilalitha M, Nandicoori VK & Khosla S (2015) The interaction of mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG methylation and histone H3/H4 binding. *Nucleic Acids Res* **43**, 3922–3937.
31 Simeone R, Bottai D & Brosch R (2009) ESX/type VII secretion systems and their role in host–pathogen interaction. *Curr Opin Microbiol* **12**, 4–10.
32 Brodin P, Majlessi L, Marsollier L, de Jonge MI, Bottai D, Demangel C, Hinds J, Neyrolles O, Butcher PD & Leclerc C (2006) Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. *Infect Immun* **74**, 88–98.
33 Simeone R, Bobard A, Lippmann J, Bitter W, Majlessi L, Brosch R & Enninga J (2012) Phagosomal rupture by *Mycobacterium tuberculosis* results in toxicity and host cell death. *PLoS Pathog* **8**, e1002507.
34 Simeone R, Sayes F, Song O, Gröschel MI, Brodin P, Brosch R & Majlessi L (2015) Cytosolic access of *Mycobacterium tuberculosis*: critical impact of phagosomal acidification control and demonstration of occurrence *in vivo*. *PLoS Pathog* **11**, e1004650.
35 Solans L, Aguiló N, Samper S, Pawlik A, Frigui W, Martín C, Brosch R & Gonzalo-Asensio J (2014) A specific polymorphism in *Mycobacterium tuberculosis* H37Rv causes differential ESAT-6 expression and identifies WhiB6 as a novel ESX-1 component. *Infect Immun* **82**, 3446–3456.
36 Kalderon D, Roberts BL, Richardson WD & Smith AE (1984) A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499–509.
37 Robbins J, Dilworth SM, Laskey RA & Dingwall C (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615–623.
38 Izaurralde E, Kann M, Pante N, Sodeik B & Hohn T (1999) EMBO Workshop Report. Viruses, microorganisms and scientists meet the nuclear pore. Leysin, VD, Switzerland, February 26-March 1, 1998. *EMBO J* **18**, 289–296.
39 Livak KJ & Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:
Movie S1. 3D representation of a deconvolved Z-stack of the high-magnification images in Fig. 7.