VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation
Winther, Kristoffer Skovbo; Tree, Jai J.; Tollervey, David; Gerdes, Kenn

Published in: Nucleic Acids Research

DOI: 10.1093/nar/gkw781

Publication date: 2016

Document version
Publisher's PDF, also known as Version of record

Document license: CC BY-NC

Citation for published version (APA):
VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation

Kristoffer Winther¹,²,* , Jai J. Tree³, David Tollervey⁴ and Kenn Gerdes¹,²,*

¹Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark, ²Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Richardson Road, NE2 4AX, Newcastle upon Tyne, UK, ³School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney 2033, Australia and ⁴Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, King’s Buildings, Edinburgh EH9 3BF, UK

Received April 15, 2016; Revised August 21, 2016; Accepted August 25, 2016

ABSTRACT

The major human pathogen *Mycobacterium tuberculosis* can survive in the host organism for decades without causing symptoms. A large cohort of Toxin–Antitoxin (TA) modules contribute to this persistence. Of these, 48 TA modules belong to the *vapBC* (virulence associated protein) gene family. VapC toxins are PIN domain endonucleases that, in enterobacteria, inhibit translation by site-specific cleavage of initiator tRNA. In contrast, VapC20 of *M. tuberculosis* inhibits translation by site-specific cleavage of the universally conserved Sarcin-Ricin loop (SRL) in 23S RNA. Here we identify the cellular targets of 12 VapCs from *M. tuberculosis* by applying UV-crosslinking and deep sequencing. Remarkably, these VapCs are all endoribonucleases that cleave RNAs essential for decoding at the ribosomal A-site. Eleven VapCs cleave specific tRNAs while one exhibits SRL cleavage activity. These findings suggest that multiple *vapBC* modules contribute to the survival of *M. tuberculosis* in its human host by reducing the level of translation.

INTRODUCTION

The major human pathogen *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), currently infects more than 8 million and kills >1 million people per year (1). The prevalence of multi-drug resistant TB is increasing and is a cause for significant concern (2). *M. tuberculosis* can persist latently without symptoms for many years in human carriers (3). The molecular mechanisms underlying latency and slowed bacterial growth are still poorly understood but may involve multiple regulatory pathways. One of these depends on the stringent response and Toxin–Antitoxin (TA) modules (4).

All bacteria produce slow growing ‘persister cells’ that are tolerant to a broad spectrum of antibiotics (5,6). Recent research on *Escherichia coli* and *Salmonella* shows that persistence is controlled by the stringent response and TA modules. Such TA modules are almost ubiquitous in bacteria and are often present in perplexingly high numbers (7). *Mycobacterium tuberculosis* has at least 88 type II TAs (8,9), raising important questions concerning their biological function(s).

Type II TA modules encode two genes in an operon, a protein ‘toxin’ that inhibits cell growth and a protein ‘antitoxin’ that counteracts the inhibitory effect of the toxin by direct protein contact (10). The antitoxins are usually metabolically unstable while the toxins are stable. Thus, regulated proteolysis of a given antitoxin determines the activity of the cognate toxin. Evidence from *E. coli* K-12 and *Salmonella* indicates that TA modules are effector genes that induce persistence when activated and step-wise deletion of 10 type II TAs progressively reduced persistence (11). Correspondingly, inactivation of Lon, the protease that degrades all known type II antitoxins of *E. coli* K-12, strongly reduced persistence. Remarkably, the TAs of *E. coli* are induced stochastically by a mechanism that depends on (pppGpp, polyphosphate and Lon. In a population of rapidly growing cells, approximately 10−3 have a high level of (pppGpp that leads to accumulation of polyphosphate, which activates Lon to degrade antitoxins. Toxin activation then leads to growth arrest, multidrug tolerance, and persistence (12). Independent support for this model has come from several comprehensive studies of *Salmonella enterica* serovar Typhimurium. S. Typhimurium has up to 20 type II TAs and deletions of single TA loci showed reduced survival of S. Typhimurium within fibroblasts (13). Furthermore, deletion of single TA genes reduced persister cell formation and survival when the bacterium was grown inside macrophages (14). Consistently, multiple type II TAs have been found to be activated in S. Typhimurium when residing...
The analysis of cDNA by high throughput sequencing growth-inhibition by selected VapCs and outline of the CRAC analysis

Figure 1. Known vapBC modules of Mycobacterium tuberculosis H37Rv, growth-inhibition by selected VapCs and outline of the CRAC analysis

inside macrophages (15). Studies with M. tuberculosis also support a role for TAs in persistence and virulence (8, 16, 17).

Mycobacterium tuberculosis has a highly expanded repertoire of TA modules, with 48 representatives of the vapBC family (7, 8) (Figure 1A). VapC toxins are PIN (pilT N-terminal) endoribonucleases containing three or four conserved acidic residues that coordinate Mg$^{2+}$ ion(s) in the active site (18). In Eukaryotes, PIN domains have been identified in multidomain endonucleases involved in RNA metabolism, RNA quality control and rRNA maturation (19). In Prokaryotes, however, most PIN domain proteins belong to the highly abundant VapC toxin family that are present in staggering numbers in certain prokaryotic genomes, including M. tuberculosis (7–9).

The molecular targets of most prokaryotic VapCs are unknown; however, the VapCs of S. Typhimurium, Shigella flexneri and Leptospira interrogans cleave initiator tRNA site-specifically in the anticodon loop, thereby inhibiting global translation (20, 21). In contrast, VapC20 of M. tuberculosis inhibits translation via cleavage of the conserved Sarcin–Ricin loop (SRL) of 23S rRNA (22).

Here, we applied UV-induced RNA-protein crosslinking and analysis of cDNA by high throughput sequencing (CRAC) to identify transcriptome-wide targets of VapC paralogues in M. tuberculosis using M. smegmatis as a surrogate host organism. First, we identified the cellular targets of six different VapCs and showed that they all cleave tRNAs site-specifically within the anticodon loop. We then used phylogenetic analysis to identify the cellular targets of an additional six VapCs. Several of these VapCs were previously shown to have non-specific RNase activity in vitro (23–25). Strikingly, we show that all 12 VapCs catalyse site-specific cleavage of RNAs essential for protein synthesis.

**MATERIALS AND METHODS**

**Strains, plasmids and growth conditions**

*Mycobacterium smegmatis* strain MC²155 was routinely grown in LB medium (Difco) containing 0.1% Tween-80 (LBT) at 37°C. For UV-crosslinking experiments strain procedure. (A) Chromosomal location of 48 vapBC modules (7, 8). Genes shown in blue are analysed here, gene shown in red was analysed previously (22). (B) Growth-inhibition tests of selected vapCs of M. tuberculosis in M. smegmatis strain MC²155 containing plasmid pMEND carrying a tetracycline inducible promoter and encoding C-terminally His-TEV-FLAG tagged (HTF) vapC genes. Strain MC²155 carrying vapC4 (Rv0595c), vapC11 (Rv1561), vapC15 (Rv2010), vapC25 (Rv0277c), vapC26 (Rv0582), vapC28 (Rv0609), vapC29 (Rv0617), vapC30 (Rv0624), vapC32 (Rv1141), vapC33 (Rv1242), vapC37 (Rv2103c) or vapC39 (Rv2530c) were grown overnight in LBT at 37°C. The optical density (OD₆₀₀) was then adjusted to 0.5, and the bacterial cultures serially diluted (10-fold) and spotted (3 μl) on nutrient agar plates without or with inducer (20 ng/ml tetracycline). The plates were incubated 3 days at 37°C. (C) Outline of Crosslinking and Analysis of cDNA protocol (CRAC). (1) The HTF tagged VapC protein is UV cross-linked (UV-C) with target RNAs in vivo, the RNA–protein complexes are purified and the RNA trimmed by a cocktail of RNase A/T1 and barcoded linkers are ligated to the RNA. (2) cDNA is synthesized by reverse transcription using 3’-end specific primer and the library generated by PCR. (3) Finally, the DNA library is purified and deep sequenced using the Illumina platform. The sequencing output is aligned and analysed using the pyCRAC software package.
MC2-155 was grown in M9 medium (Difco) containing 0.1% Tween-80 (M9T) with 0.1% Casein hydrolysate, 1µg/ml thiamine and 0.2% glucose as carbon source at 37°C. When appropriate, 50µg/ml kanamycin was added to maintain the plasmid. Furthermore, when stated 20 ng/ml tetracycline was added to liquid or solid medium to induce transcription from tetracycline inducible promoters.

Plasmids

Construction of plasmids is described in Supplementary Information; Supplementary Table S1 contains a list of oligonucleotides used to construct plasmids and to detect RNAs in northern analysis.

Crosslinking and analysis of cDNA

Cultures (2 l) of M. smegmatis MC2-155 containing plasmids pMEND::HTF, pMEND::vapC4::HTF, pMEND::vapC11::HTF, pMEND::vapC28::HTF, pMEND::vapC29::HTF, pMEND::vapC30::HTF, pMEND::vapC37::HTF, were grown exponentially in M9T at 37°C under constant aeration. At OD<sub>600</sub> = 0.3–0.4 cultures were induced by addition of 20 ng/ml tetracycline. After 20 min of incubation cultures were irradiated with 1800 mJ of UV-C for 100 seconds (Van Remmen UV Techniek). The cells were subsequently harvested by centrifugation, washed in ice-cold phosphate buffered saline (PBS) containing 0.1% Tween-80 and snap frozen in liquid nitrogen. The pellets were stored at −80°C. The HTF tagged proteins were then purified, linkers ligated to crosslinked RNA, cDNA synthesized and DNA libraries generated as described in (26). The DNA libraries were then sequenced on the illumina MiSeq platform and the sequencing output analysed using the pyCRAC tool package (27).

Northern blotting analysis of tRNA and rRNA after VapC induction

M. smegmatis MC2-155 containing plasmid pMEND-HTF with VapC of interest were grown exponentially in LBT containing kanamycin (50µg/ml) at 37°C. At an OD<sub>600</sub> of 0.3–0.4, vapC transcription was induced by the addition of tetracycline (20 ng/ml). A sample was collected before and after 120 min of incubation with inducer. Total RNA was purified using the FastRNA Blue Kit (MPBiomedicals) according to manufacturers instructions. The cell samples were lysed by homogenization using the MagNA lyzer (Roche) twice at 6500 rpm for 40 s, separated by a 1 min rest on ice. After precipitation the RNA was dissolved in nuclease free water. Total RNA (2.5 µg) was denatured in Formamide loading buffer and separated on a 4.5% (rRNA) or 8% (tRNA) 8 M urea polyacrylamide gel in 1× TBE (100 mM Tris–borate and 2 mM ethylenediaminetetraacetic acid, EDTA). The RNA was then transferred to a Zeta-Probe membrane (Bio-Rad) by semi-dry electrophoretic transfer and the membrane pre-hybridized with SESI buffer (0.25 M NaPO<sub>4</sub> pH 7.2, 1 mM EDTA and 7% SDS) at 42°C for at least 30 min. Probe oligonucleotides (20 pmol) (see Supplementary Table S1) were labelled using 30µCi γ<sup>32</sup>P-ATP and T4 Polynucleotide kinase (Fermentas) according to manufacturers procedures and added to the membrane (due to the high primary sequence similarity between M. smegmatis and M. tuberculosis tRNAs, M. smegmatis tRNA probes could be used in most cases except for tRNA<sub>24Ser-GGA</sub> and tRNA<sub>28Ser-GGA</sub>). The tube was then incubated overnight at 29°C in a hybridization oven. After hybridization the membrane was washed 2–3 times in SESI buffer (0.25 M NaPO<sub>4</sub> pH 7.2, 1 mM EDTA and 5% SDS) at room temperature. The bands were visualized by phosphorimaging. The membrane could be stripped with repeated washes of 0.1% SDS at 85°C and subsequently re-probed.

Purification of VapC4-HTF, VapC11-HTF, VapC28-HTF and VapC37-HTF

VapC toxins were purified from Mycobacterium smegmatis using histidine affinity chromatography as we described previously (22). Mycobacterium smegmatis MC2-155 containing plasmids pMEND::vapC4::HTF, pMEND::vapC11::HTF, pMEND::vapC28::HTF and pMEND::vapC37::HTF were grown exponentially in 1-L LB at 37°C with shaking. At OD<sub>600</sub> = 0.4, the culture was induced with tetracycline (20 ng/ml) for 1 h before cells were pelleted by centrifugation. The cells were then washed in ice-cold PBS containing 0.1% Tween-80 and lysed in 1 ml (v/w) lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8, 300 mM NaCl, 10 mM Imidazole, 0.1% NP-40, 5 mM β-mercaptoethanol) by bead beating (2 × 6500 rpm for 40 s, MagNA lyzer cell homogenizer) using 0.5 ml 100 µM glass beads per 1.5 ml of cell suspension. The lysate was cleared by centrifugation at 14 k rpm for 20 min and incubated overnight at 4°C with 0.5 ml equilibrated Ni-NTA resin (Qiagen). The resin was then loaded onto a gravity-flow column and washed with 40 column volume of wash buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 20 mM Imidazole, 0.1% NP-40, 5 mM β-mercaptoethanol). The column was then washed with four column volume wash buffer containing 40 mM imidazole and four column volume wash buffer containing 50 mM imidazole. The protein was eluted by incubating the column twice with elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8, 300 mM NaCl, 500 mM Imidazole, 0.1% NP-40, 5 mM β-mercaptoethanol). Elution fraction containing VapC toxin was dialysed for 6 h against PBS containing 1 mM DTT and overnight against PBS containing 50% glycerol and 1 mM DTT at 4°C.

Detection of in vitro RNA cleavage by Northern blotting analysis

M. smegmatis total RNA was purified using the FastRNA blue kit (see previous section). To ensure high quality, RNA was purified by an additional step of phenol/chloroform extraction. Total RNA from M. tuberculosis H37Rv was a generous gift from Douglas Young and Rachel Lai (National Institute for Medical Research, London and Imperial College London, UK). For the cleavage reaction 1 µg or 2.5 µg of total RNA was mixed with 1 µg or 2 µg of purified VapC4-HTF VapC11-HTF, VapC28-HTF or VapC37-HTF in cleavage buffer (final: 10 mM HEPES pH 7.5, 15 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 25% glycerol).
and left to incubate at 37°C for 30 min. As a control VapC toxin was omitted in some reactions and Mg²⁺ chelated by the addition of 25 mM EDTA. The cleavage reactions were terminated by addition of FD-loading buffer and the RNA separated on an 8% Urea polyacrylamide gel buffered with 1× TBE. The RNA was transferred to a Zeta-Probe membrane and the RNA of interest detected using a radiolabelled probe as previously described.

Mapping of VapC4, VapC11, VapC28 and VapC37 cleavage site by primer extension analysis

The cleavage reactions were set up as previously described, using total RNA (4 μg) with or without 2 μg of purified VapC4-HTF, VapC11-HTF, VapC28-HTF or VapC37-HTF. The cleavage reactions were then incubated for 30 min at 37°C in cleavage buffer. The reaction was terminated by addition of 400 μl 100 mM Na-acetate pH 4.5 followed by phenol/chloroform extraction and ethanol precipitation. The precipitate was washed with 70% ethanol, air dried and dissolved in nuclease free dH₂O. The reverse transcription was performed by setting up a hybridization reaction containing 0.2 pmol radiolabelled KW-MS-PXT-CysGCA-rv, KW-MS-PXT-TrpCAG-rv, KW-MS-PXT-SerTGA-S1-rv or KW-MS-PXT-LeuCAG-S-rv oligonucleotide and 1 μg of purified cleaved RNA in nuclease free dH₂O. The oligonucleotide (4 pmol) was phosphorylated using 30 μCi γ³²P-ATP and T4 Polynucleotide kinase (Fermentas) and subsequently desalted using G-25 desalting columns (GE healthcare). The hybridization reactions were incubated at 80°C for 5 min and transferred to an icebath and left to incubate for 5 min. To the chilled reaction 1× FS buffer (Invitrogen), 10 mM DTT and 1 mM dNTP was added and the tube transferred to 54°C to incubate for 2 min. Then 20 U of Superscript III reverse transcriptase (Invitrogen) was added and the tubes left to incubate for 1 h. The reaction was terminated by addition of an equal volume of FD-loading buffer. The reactions were loaded onto a 10% polyacrylamide gel containing 8 M urea and 1× TBE. Along with the reverse transcription reactions a dyeoxy sequencing ladder was loaded which had been made from a PCR template generated using oligonucleotides KW-MS-PXT-CysGCA-f and KW-MS-PXT-CysGCA-rv, KW-MS-PXT-TrpCAG-rv, KW-MS-PXT-SerTGA-S1-rv, KW-MS-PXT-LeuCAG-S-rv and KW-MS-PXT-LeuCAG-f. After the cDNA had been separated, the gel was dried and the bands visualized by phosphorimaging.

RESULTS

Identification of cellular VapC targets using CRAC

We implemented CRAC to systematically identify the cellular targets of VapC RNases from *M. tuberculosis*, using *M. smegmatis* as a surrogate host organism. CRAC identifies RNAs that interact directly with a tagged bait protein in living cells and potentially detects both stable and relatively transient interactions (26,28). To apply the CRAC protocol, all 48 vapC genes of *M. tuberculosis* H37Rv (Figure 1A) were cloned into plasmid pMEND-HTF downstream of a tetracycline inducible promoter and in frame with a dual affinity tag that introduced a hexaHis, TEV protease cleavage site, and 3× FLAG tags at the C-termini (HTF-tag). Figure 1B shows the effect of transcriptional induction of selected vapC genes. Expression of VapC4, 11, 25, 26, 28, 29, 30, 33, 37 and 39 strongly inhibited cell growth, whereas expression of VapC15 and 32 had a more moderate effect (for the results of phenotypic testing of all 48 VapCs, see Supplementary Figure S1A). The pattern of growth inhibition was generally consistent with previous tests in which native VapCs were expressed in *M. smegmatis* (8,29). The only exception was VapC32 that did not significantly reduce cell growth in our experimental system (Figure 1B). Thus, we can conclude that the C-terminal HTF tag does not interfere with VapC-mediated inhibition of cell growth by the toxins selected in this study.

The majority of VapCs expressed in *M. smegmatis* showed weak or no inhibition of growth (Supplementary Figure S1A). Therefore, we determined the expression levels of a number of these VapCs (Supplementary Figure S1B). We found that in some cases, including VapC41, 43 and 44, no VapC expression was detectable 60 min after induction that could explain the lack of growth inhibition. However, for the majority of non-toxic VapCs that we tested, i.e. VapC6, 10, 12, 19, 31 and 47, protein expression was detectable, indicating that the proteins either are not functional (e.g. the cellular target could be missing in the heterologous host). Alternatively, their biological functions do not entail inhibition of cell growth.

 Cultures of *M. smegmatis* MC²¹⁵ were pulsed with HTF-tagged VapC and were UV-irradiated to covalently crosslink VapCs to their target RNAs in vivo (Figure 1C). The covalently bound VapC-RNA complexes were then purified using M2 anti-FLAG resin, eluted by TEV-cleavage, and RNAs were trimmed to allow ‘footprinting’ of the protein interaction site. Trimmed VapC-RNA complexes were bound to Ni-NTA resin and barcoded linkers were ligated to the 5’ and 3’ termini. Linker-ligated RNA–protein complexes were eluted and size-selected using SDS-PAGE. Protein-RNA complexes of the appropriate molecular weight were extracted and protease digested. cDNAs were generated by RT-PCR, subjected to deep sequencing and reads were mapped to the *M. smegmatis* genome. VapC binding sites were initially identified using the py-CRAC software package to identify transcripts bound by each VapC. We used duplicate controls (pMEND::HTF) to assess the background of the assay. CRAC data from samples expressing different VapCs were obtained from single experiments. The results are shown as reads per million allowing us to visually compare for significance across samples.

In the following section, we describe the identification of the cellular targets of previously uncharacterized VapCs. We will refer to RNA interactions that were confirmed to confer RNA cleavage as ‘productive interactions’, and other interactions as ‘unproductive interactions’.

All six VapCs interact with a subset of tRNAs, SRP RNA and 23S rRNA

Analysis of CRAC data for VapC4, 11, 28, 29, 30 and 37 each revealed enrichment for tRNA⁵⁷Thr-GGT.
VapC4 specifically cleaves tRNA44Cys-GCA in M. smegmatis

Analysis of specific RNA interaction with VapC4 using CRAC revealed strong enrichment for RNA fragments mapping to tRNA44Cys-GCA (Figure 2A). The interaction with tRNA44Cys-GCA was productive as the full-length tRNA decreased upon induction of vapC4 simultaneous with the accumulation of a smaller cleavage product (compare lanes 1 and 2 with 3 and 4 in Figure 2B). We also observed weak cleavage of tRNA44Cys-GCA before induction of vapC4. This observation can be explained due to ‘leaky’ transcription initiation from the tetracycline inducible promoter as no cleavage is observed in the control. Induction of the other vapC toxins did not affect tRNA44Cys-GCA stability (lanes 5–14). The tRNAs; tRNA23Pro-CGG and tRNA32Phe-GAA were also specifically enriched by VapC4 (Supplementary Figure S3A and S3B, respectively) but none of these tRNA species were cleaved by VapC4 (compare lanes 1 and 2 in Supplementary Figure S3D and S3E, respectively). The other cysteine-accepting tRNA of M. smegmatis, annotated as tRNA40Cys-GCA, was not enriched by VapC4 in the analysis, and was not investigated further (Supplementary Figure S3C).

To confirm that the observed cleavage of tRNA44Cys-GCA was direct, cleavage was investigated in vitro by Northern analysis using purified VapC4-HTF in a reaction containing total RNA from M. smegmatis. Indeed we observed cleavage of tRNA44Cys-GCA only in the presence of VapC4 (Figure 2C). Consistent with the VapC PIN domain coordinating Mg2+ in the active site, addition of EDTA to the reaction abolished cleavage. Mapping of the cleavage site in tRNA44Cys-GCA by primer extension analysis revealed cleavage in the anticodon loop between bases C34AJ A35, A35J A36, A36J A37 and A37J G38 (arrows indicate cleavages between numbered nucleotides), with the A36J A37 site being the dominant cleavage site (Figure 2D and E). The reverse transcriptase weakly terminated at this site before addition of VapC4, probably due to modification of the corresponding base.

VapC4 has recently been suggested to cleave tRNA2Ala-TGC, tRNA24Ser-GGA, tRNA26Ser-GCT of M. smegmatis, with the A35↓G38 cleavage in the anticodon loop between bases C34AJ A35, A35J A36, A36J A37 and A37J G38 (arrows indicate cleavages between numbered nucleotides), with the A36J A37 site being the dominant cleavage site (Figure 2D and E). The reverse transcriptase weakly terminated at this site before addition of VapC4, probably due to modification of the corresponding base.

VapC4 cleaves tRNA44Cys-GCA in M. smegmatis

Analysis of specific RNA interaction with VapC4 using CRAC revealed strong enrichment for RNA fragments mapping to tRNA44Cys-GCA (Figure 2A). The interaction with tRNA44Cys-GCA was productive as the full-length tRNA decreased upon induction of vapC4 simultaneous with the accumulation of a smaller cleavage product (compare lanes 1 and 2 with 3 and 4 in Figure 2B). We also observed weak cleavage of tRNA44Cys-GCA before induction of vapC4. This observation can be explained due to ‘leaky’ transcription initiation from the tetracycline inducible promoter as no cleavage is observed in the control. Induction of the other vapC toxins did not affect tRNA44Cys-GCA stability (lanes 5–14). The tRNAs; tRNA23Pro-CGG and tRNA32Phe-GAA were also specifically enriched by VapC4 (Supplementary Figure S3A and S3B, respectively) but none of these tRNA species were cleaved by VapC4 (compare lanes 1 and 2 in Supplementary Figure S3D and S3E, respectively). The other cysteine-accepting tRNA of M. smegmatis, annotated as tRNA40Cys-GCA, was not enriched by VapC4 in the analysis, and was not investigated further (Supplementary Figure S3C).

To confirm that the observed cleavage of tRNA44Cys-GCA was direct, cleavage was investigated in vitro by Northern analysis using purified VapC4-HTF in a reaction containing total RNA from M. smegmatis. Indeed we observed cleavage of tRNA44Cys-GCA only in the presence of VapC4 (Figure 2C). Consistent with the VapC PIN domain coordinating Mg2+ in the active site, addition of EDTA to the reaction abolished cleavage. Mapping of the cleavage site in tRNA44Cys-GCA by primer extension analysis revealed cleavage in the anticodon loop between bases C34AJ A35, A35J A36, A36J A37 and A37J G38 (arrows indicate cleavages between numbered nucleotides), with the A36J A37 site being the dominant cleavage site (Figure 2D and E). The reverse transcriptase weakly terminated at this site before addition of VapC4, probably due to modification of the corresponding base.

VapC4 has recently been suggested to cleave tRNA2Ala-TGC, tRNA24Ser-GGA, tRNA26Ser-GCT of M. smegmatis.
VapC28 and VapC30 specifically enriched for RNA fragments derived from tRNA3Leu-CAG, tRNA13Leu-GAG and CRAC analysis showed enrichment for RNA frag-
ments derived from tRNA25Ser-TGA and tRNA28Ser-CGA. CRAC analysis with VapC28 and VapC30 did not affect the stabil-
ity of tRNA16His-GTG, tRNA24Ser-GGA, tRNA26Ser-GCT and tRNA22Thr-TGT by VapC30 alone (Supplementary Figure S4D–G). These enrichments were also categorized as unproductive as no changes in stabilities were observed after induction of VapC30 (compare lanes 1 and 2 in Supplementary Figure S4K–N).

The activity of VapC28 was confirmed by in vitro cleavage. Thus, VapC28 cleaved tRNA25Ser-TGA of M. smegmatis (Figure 4E) and a discrete cleavage product was detected, suggesting that the absence of clear cleavage product in vivo reflected their rapid degradation. Addition of EDTA abolished cleavage, consistent with the requirement for Mg2+ in the reaction. The cleavage site of VapC28 in tRNA25Ser-TGA was mapped to nucleotide pairs G36↓A37 in the anticodon loop (Figure 4F and G). At A38, we observed strong termination of reverse transcriptase before the addition of VapC28, probably reflecting base modification at this position.

VapC29 and VapC37 cleave tRNA7Trp-CCA

CRAC with VapC29 and VapC37 as baits identified interactions with RNA fragments derived from tRNA7Trp-CCA (Figure 5A). In addition, VapC29 also specifically enriched for tRNA4Tyr-GTA-derived fragments while VapC37 specifically enriched for tRNA42Leu-CAA-derived fragments (Supplementary Figure S5A and S5B). Northern analysis revealed productive interactions between tRNA7Trp-CCA and VapC29 and VapC37 in M. smegmatis cells and two discrete cleavage products accumulated upon induction (compare lanes 1 and 2 with lanes 9 and 10 and 13 and 14 of Figure 5B). None of the other VapCs analysed by CRAC was observed to affect the stability of this tRNA (lanes 3–8 and 11–12). Neither tRNA4Tyr-GTA nor tRNA42Leu-CAA were cleaved by VapC29 and VapC37, respectively, and these interactions were therefore categorized as unproductive (compare lanes 1 and 2 in Supplementary Figure S5C and S5D). Incubation of purified VapC37 with total RNA of M. smegmatis confirmed direct cleavage of tRNA7Trp-CCA that was Mg2+ dependent (Figure 5C). We only observe one cleavage product in vitro, indicating that the two cleavage products observed in vivo is a result of cellular RNAses similar to that observed for tRNA25Ser-TGA and tRNA28Ser-GCT.

The VapC37 cleavage sites in tRNA7Trp-CCA were mapped to bases A36↓A37, A37↓A38 and A38↓A39, with the A37↓A38 cleavage site being the most prominent (Figure 5D and E). Similar to the other tRNAs mapped by primer extension analysis we observe termination of reverse transcription at these sites before addition of VapC37 indicating that the sites are putative sites of modification.

VapC28 and VapC30 both cleave tRNA25Ser-TGA and tRNA28Ser-GGA

VapC28 and VapC30 specifically enriched for RNA fragments derived from tRNA25Ser-TGA, tRNA28Ser-GGA (compare lanes 1 and 2 with 7 and 8 and 11 and 12 of Figure 4A and B). Induction of vapC28 and vapC30 showed reduced amounts of full-length tRNA25Ser-TGA and tRNA28Ser-GGA. However, stable, discrete cleavage products were not detected (compare lane 1–2 with 7–8 and 11–12 Figure 4C and D). Expression of the other VapCs analysed by CRAC did not affect the stability of tRNA25Ser-TGA and tRNA28Ser-GGA. CRAC analysis also revealed tRNA47Gly-GCC, tRNA17Lys-CTT and tRNA2Ala-TGC to be enriched by VapC28 and VapC30 (Supplementary Figure S4A–C). However, none of these tRNAs were detectably affected by induction of vapC28 and vapC30 and these interactions were categorized as unproductive (Supplementary Figure S4H–J, compare lane 1 and 2). Furthermore, CRAC analysis revealed enrichment of tRNA16His-GTG, tRNA24Ser-GGA, tRNA26Ser-GCT and tRNA22Thr-TGT by VapC30 alone (Supplementary Figure S4D–G). These enrichments were also categorized as unproductive as no changes in stabilities were observed after induction of VapC30 (compare lanes 1 and 2 in Supplementary Figure S4K–N).

VapC29 and VapC37 cleave orthologous tRNAs in M. tuberculosis

The above-described in vitro cleavage reactions were performed using total RNA from M. smegmatis as the substrate. Even though the tRNAs of M. smegmatis are almost identical to the orthologous tRNAs of M. tuberculosis there are nucleotide differences that potentially could affect VapC recognition and cleavage (Supplementary Figure S6A–G). The VapC cleavage reactions were therefore also analysed...
Figure 3. VapC11 cleaves tRNA^{3Leu-CAG}, tRNA{10Leu-GAG} and tRNA^{13Gln-CTG} in M. smegmatis. Enrichment of (A) tRNA^{3Leu-CAG}, (B) tRNA{10Leu-GAG} and (C) tRNA^{13Gln-CTG} by six different VapCs analysed by CRAC as in Figure 2 (A). In vivo cleavage by six VapCs was analysed as in Figure 2(B) and the tRNAs of interest detected using radiolabelled oligonucleotides complementary to (D) tRNA^{3Leu-CAG}. (E) tRNA^{13Gln-CTG}. (F) tRNA^{10Gln-CTG}. The tRNAs were analysed before (indicated by 0) and after 120 min of induction (indicated by 120) in M. smegmatis cells. Cleavage products are indicated with arrows. (G) in vitro cleavage assay of tRNA^{3Leu-CAG} without (indicated by –) or with (indicated by +) purified VapC11-HTF detected by Northern analysis as in Figure 3(C). A reaction with EDTA was included as a control for dependency on Mg^{2+}. (H) Cleavage site mapping by primer extension analysis was performed as in Figure 2(D), using the 32P-labeled oligonucleotide MS-PXT-LeuCAG-S-rv. An asterisk indicates the position of a putative modification in the tRNA. (I) Visualization of the VapC11-mediated in vitro cleavage sites on the secondary structure of tRNA^{3Leu-CAG}. The base positions of the cleavage sites are indicated in the anticodon loop.

Using total RNA of M. tuberculosis H37Rv, consistent with our results in M. smegmatis, VapC4 cleaved tRNA^{21Cys-GCA} (Figure 6A, lanes 2 and 3), VapC11 specifically cleaved tRNA^{3Leu-CAG} (Figure 6B, lane 2), VapC28 specifically cleaved tRNA^{25Ser-TGA} and tRNA^{28Ser-CCG} (Figure 6C and D, lane 5) and VapC37 specifically cleaved tRNA^{7Trp-CCA} (Figure 6E, lane 8). The cleavages were tRNA-specific and all cleavages were also inhibited by the EDTA addition, indicating Mg^{2+} dependence (Figure 6B-E). These results confirmed that heterologously expressed VapCs have identical cleavage specificities in M. smegmatis, validating this organism as a useful surrogate host for the analysis of VapC toxins from M. tuberculosis.

Phylogeny can predict novel VapC targets

The identification of novel VapC targets is challenging and in the past required the application of laborious ‘trial-and-error’ approaches. Since M. tuberculosis has at least 48 different VapCs (Figure 1A), we tested the possibility that the targets we had identified could provide insights into the targets of related VapC toxins using a phylogenetic approach.
From the phylogenetic tree shown in Supplementary Figure S1C, we selected six VapCs that group together with one of the VapCs with a known target, identified here or in our previous analysis (22), and tested whether these VapCs cleaved the same targets. The following examples showed that this approach was effective.

(i) VapC15 and VapC32 cleave tRNA3Leu-CAG. VapC15 and VapC32 group together with VapC11 in the phylogram (Figure 7E and Supplementary Figure S1C). Therefore, we asked if these two VapCs also cleave tRNA3Leu-CAG, the tRNA that was cleaved most efficiently by VapC11. Indeed tRNA3Leu-CAG was cleaved upon induction of vapC15 and vapC32 (Figure 7A, lanes 3 - 6). These cleavages were relatively weak compared to those of VapC11, consistent with the lower toxicity of these VapCs in M. smegmatis (Figure 1B). VapC11 was also found to have productive interactions with tRNA13Leu-GAG and tRNA10Gln-CTG, but these tRNA species were not cleaved by VapC15 or VapC32 (Supplementary Figure S7A and S7B).

(ii) VapC25, 33 and 39 cleave tRNA7Trip-CCA. VapC25, 33 and 39 group together with VapC29 and 37 in the phylogram (Figure 7E and Supplementary Figure S1C) and induction of these toxins in M. smegmatis inhibited cell growth (Figure 1B). Notably, induction of vapC25, 33 or 39 resulted in cleavage products identical to those of VapC29 and 37 (Figure 7B lanes 3 - 8).

(iii) VapC26 cleaves 23S rRNA in the Sarcin–Ricin Loop (SRL). VapC26 groups with VapC20 in the phylogram (Figure 7E and Supplementary Figure S1C) and inhibited growth of M. smegmatis very efficiently (Figure 1B). Since VapC20 cleaves 23S rRNA in the SRL loop (22), we tested if VapC26 also cleaves 23S rRNA. Indeed, induction of vapC26 resulted in a 23S rRNA cleavage pattern identical to that of VapC20 showing that the two VapCs have identical targets (Figure 7C, lanes 3–6).

VapCs are highly target-specific

The above results indicated that VapCs grouping together phylogenetically exhibit identical RNA cleavage specificities (Figure 7E and Supplementary Figure S1C). Therefore, we analysed whether VapCs from one phylogenetic subgroup would cleave RNAs targeted by a different VapC subgroup.

```
Figure 4. VapC28 and VapC30 cleave tRNA25Ser-TGA and tRNA28Ser-CGA of M. smegmatis. Enrichment of tRNA fragments from (A)
```
DISCUSSION

Using CRAC in combination with phylogenetic analysis and in vivo and in vitro RNA cleavage assays, we identified the cellular targets of 12 novel VapCs from *M. tuberculosis* H37Rv. UV-crosslinking and high throughput sequencing revealed binding sites that were both produc-
Phylogeny can predict cellular VapC targets. (A) Detection of \textit{in vivo} cleavage of tRNA\textsubscript{3Leu-CAG} by VapC15 and VapC32 by Northern analysis was performed as in Figure 3 (D); VapC expression was induced from plasmid pMEND::\textit{vapC15::HTF} or pMEND::\textit{vapC32::HTF}. (B) \textit{In vivo} cleavage of tRNA\textsubscript{2Ala-TGC} by VapC25, C33 and C39 detected by Northern analysis performed as in Figure 5 (B). Toxin expression was induced from plasmid pMEND::\textit{vapC25::HTF}, pMEND::\textit{vapC33::HTF} and pMEND::\textit{vapC39::HTF}. (C) Cleavage of 23S RNA detected by Northern analysis was performed on RNA purified from \textit{M. smegmatis} strain MC\textsubscript{2}155 containing plasmids pMEND::\textit{vapC20::HTF} and pMEND::\textit{vapC26::HTF}. Cell samples were collected before (indicated by 0) and after 120 min of induction (indicated by 120) and analysed as described in Materials and Methods. The arrow indicates the 276 nucleotide cleavage product of 23S rRNA generated by VapC20 and VapC26, as mapped previously (22). The full-length 23S rRNA (3164 nt) is also indicated with an arrow. (D) Region of the decoding center of 70S ribosome of \textit{Thermus thermophilus} in complex with mRNA and the ternary complex of EF-Tu•Thr•tRNA\textsubscript{Thr}•GDP (pdb entry: 4V5G from (31)). The Figure highlights that VapCs of \textit{M. tuberculosis} cleave RNAs essential for decoding during translation. The Sarcin - Ricin Loop (SRL), which is cleaved by VapC20 and 26, is shown in orange. The tRNA anticodon loop, which is cleaved by VapC4, 11, 15, 25, 28, 29, 30, 33, 37 and 39, is also marked with a box. Bases cleaved are coloured red. Major parts of the ribosome structures were omitted for clarity. (E) Phylogenetic tree of VapCs of \textit{M. tuberculosis} strain H37Rv with known targets was based on a multiple sequence alignment made using the ClustalX v2.0 software; amino acid sequences were retrieved from www.ncbi.nlm.nih.gov. VapCMYSH: VapC from \textit{Shigella flexneri} 2a plasmid pMYSH6000 (21). VapCLEPTO: VapC of \textit{Leptospira interrogans} (20). The tRNAs in the green, magenta, blue and light blue boxes have identical names in \textit{M. smegmatis} strain MC\textsubscript{2}155 and \textit{M. tuberculosis} H37Rv except tRNA\textsubscript{13Leu-GAG} and tRNA\textsubscript{10Gln-CTG} of \textit{M. smegmatis} that are denoted tRNA\textsubscript{A13Leu-GAG} and tRNA\textsubscript{A10Gln-CTG} in \textit{M. tuberculosis} H37Rv. tRNA\textsubscript{2Ala-TGC} in \textit{M. smegmatis} is named tRNA\textsubscript{A2Ala-TGC} in \textit{M. tuberculosis} H37Rv. The sequences of the orthologous tRNAs studied here are shown in Supplementary Figure S6.
only seen after highly extended incubation times only (3-
12hr; (30)), raising the possibility that these tRNAs repre-
sent low-affinity targets of VapC4. Using UV-crosslinking
and Northern analysis we found that VapC4 actually in-
teracts with, and cleaves tRNA44Cys-GCA in vivo. We used a
much shorter incubation time (30 min) and exclusively ob-
served cleavage of tRNA44Cys-GCA of M. smegmatis and the
orthologous tRNA21Cys-GCA of M. tuberculosis H37Rv.

In enterobacteria, TA gene modules encoding trans-
lational inhibitors promote increased survival following
exposure to antibiotics or stressful conditions within
macrophages (11–15,32). The biological function(s) of the
multitude of vapBC modules encoded by M. tuberculosis
has not yet been established. However, it is possible that
cleavage of RNAs essential for translation similarly confers
protection by inhibiting cell growth, as proposed previously
by other groups (8,16,17,33). Consistent with this proposal, E.
coli rapidly responds to oxidative stress by downregulat-
ing translation via tRNA degradation (34). It is also possible
that the vapBC genes can prevent bacteriophages from
spreading in cell populations by inducing so-called ‘abortive
infection’ that has been described for type III TA modules
also encoding inhibitors of translation (35,36).

The M. tuberculosis genome contains an unusually large
expansion of TA modules that present a unique challenge
for understanding their contribution to bacterial persis-
tence in this significant human pathogen. Here we de-
scribed a new methodology to identify the cellular targets of
PIN-domain RNases in vivo and demonstrated that related
clades of VapC toxins have similar substrate specificities. To
our knowledge this study represents the most comprehen-
sive analysis of VapC targets to date and greatly expands
our understanding of VapC targets in M. tuberculosis.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank Douglas Young and Rachel Lai (National Insti-
tute for Medical Research, London and Imperial College
London, UK) for the donation of purified M. tuberculosis
RNA and Ditlev E. Brodersen for critical comments on
the manuscript.

FUNDING
Danish Natural Research Foundation [DNRF120; to
K.G.]; Novo Nordisk Foundation (to K.G.); European
Research Council Advanced Investigator Grant PERSIST
[294517 to K.G.]. Funding for open access charge: Dan-
marks grundforskingsfond.

Conflict of interest statement. None declared.

REFERENCES
development of new tuberculosis drugs and treatment regimens. Nat.
3. Lilbeck,T., Dirksen,A., Baess,J., Strunge,B., Thomsen,V.O. and
of Mycobacterium tuberculosis after 33 years of latent infection. J.
Infect. Dis., 185, 401–404.
abundant in free-living but lost from host-associated prokaryotes.
functional analysis of Mycobacterium tuberculosis toxin-antitoxin
systems: implications for pathogenesis, stress responses, and
HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea.
3, 371–382.
Controls Bacterial Persistence by Stochastic Induction of
13. Lobato-Marquez,D., Moreno-Cordoba,I., Figueroa,V.,
and type II toxin-antitoxin modules control Salmonella lifestyle inside
Physiologic stresses reveal a Salmonella persister state and TA
family toxins modulate tolerance to these stresses. PLoS One, 10,
e0141343.
MazF ribonucleases promote Mycobacterium tuberculosis
and transcriptome analysis of Mycobacterium tuberculosis persisters.
MBio, 2, e00100–00111.
is required for D-site cleavage in 20S pre-rRNA. RNA, 10, 1698–1701.
(2010) Alternative RNA editing in Mycobacterium tuberculosis
inhibits translation and reveals a potential novel drug target.
inhibits translation and reveals a potential novel drug target.
inhibits translation and reveals a potential novel drug target.


