RNase HI depletion strongly potentiates cell killing by rifampicin in mycobacteria.

Abeer Al-Zubaidi,[a,b] Chen-Yi Cheung,[c] Gregory M. Cook,[b,c] George Taiaroa,[d] Valerie Mizrahi,[e] J. Shaun Lott,[a,b#] and Stephanie S. Dawes.[a,b#]

[a] School of Biological Sciences and [b]Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Auckland 1142, New Zealand.

[c] Department of Microbiology and Immunology, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand

[d] Department of Microbiology and Immunology, The Peter Doherty Institute for Infection and Immunity at The University of Melbourne, Melbourne, Victoria, Australia

[e] MRC/NHLS/UCT Molecular Mycobacteriology Research Unit & DST/NRF Centre of Excellence for Biomedical TB Research & Wellcome Centre for Infectious Diseases Research in Africa, Institute of Infectious Disease & Molecular Medicine, & Department of Pathology, University of Cape Town, Anzio Road, Observatory 7925, Cape Town, South Africa.

Running Head: RNase HI depletion potentiates rifampicin killing.

Address correspondence to Stephanie S. Dawes (s.dawes@auckland.ac.nz) or J. Shaun Lott (s.lott@auckland.ac.nz)

Abstract

Multidrug resistant (MDR) tuberculosis (TB) is defined by the resistance of *Mycobacterium tuberculosis*, the causative organism, to the first-line antibiotics rifampicin and isoniazid. Mitigating or reversing resistance to these drugs offers a means of preserving and extending their use in TB treatment. R-loops are RNA/DNA hybrids that are formed in the genome during transcription, and can be lethal to the cell if not resolved. RNase HI is an enzyme that removes R-loops, and this activity is essential in *M. tuberculosis:* knockouts of *rnhC*, the gene encoding RNase HI, are non-viable. This essentiality supports it as a candidate target for the development of new antibiotics. In the model organism *Mycolicibacterium smegmatis*, RNase HI activity is provided by two RNase HI enzymes, RnhA and RnhC. We show that the partial depletion of RNase HI activity in *M. smegmatis*, by knocking out either of the genes encoding RnhA or RnhC, led to the accumulation of R-loops. The sensitivity of the knockout strains to the antibiotics moxifloxacin, streptomycin and rifampicin was increased, with sensitivity to the transcriptional inhibitor rifampicin strikingly increased by nearly 100-fold. We also show that R-loop accumulation accompanies partial transcriptional inhibition, suggesting a mechanistic basis for the synergy between RNase HI depletion and transcriptional inhibition. A model of how transcriptional inhibition can potentiate R-loop accumulation is presented. Finally, we identified four small molecules that inhibit recombinant RnhC activity and that also potentiated rifampicin activity in whole-cell assays against *M. tuberculosis*, supporting an on-target mode of action, and providing the first step in developing a new class of anti-mycobacterial drug.

Importance

This study validates mycobacterial RNase HI as a druggable, vulnerable candidate for a new therapeutic treatment of *M. tuberculosis* with a novel mode of action. RNase HI depletion shows synergistic bacterial killing with some current first- and second-line antibiotics, suggesting that RNase HI inhibitors would combine well with these regimens, and could potentially accelerate the clearance of drug-sensitive strains. RNase HI inhibitors also have the potential to reduce the effective dose of rifampicin, with the comcommitant reduction in side effects. The potentiation of rifampicin efficacy conferred by RNase HI deficiency suggests that RNase HI inhibitors may be able to mitigate against development of rifampicin resistance. The synergy may also be able to reverse rifampicin resistance, rescuing this antibiotic for therapy. The surprising finding that low levels of transcriptional inhibition potentiate R-loop formation provides a key new insight into R-loop metabolism.

Introduction

The emergence of drug-resistant bacteria has been recognised as a global health crisis (1). If left unchecked, antibiotic resistance threatens to overturn the substantial gains that have been made to human healthcare by the use of antibiotics to treat and prevent infectious disease, resulting in huge social and economic costs. Globally, the antibiotic resistance crisis has been spearheaded by the alarmingly rapid rise in the prevalence of *M. tuberculosis* strains that are resistant to the cornerstone antibiotics of first-line TB therapy, rifampicin (which inhibits transcription) and isoniazid (which inhibits cell wall biosynthesis). The situation has been exacerbated by the emergence and spread of extensively drug-resistant (XDR) *M. tuberculosis* strains that are also resistant to second-line TB therapeutics, such as the fluoroquinolones (which inhibit DNA topoisomerases) and linezolid (which inhibits the ribosome). It is estimated that there were at least 450,000 new cases of MDR-TB worldwide in 2019, and at least 180,000 deaths from MDR-TB were reported (2). The World Health Organisation (WHO) has set a goal to eliminate TB by 2035, but this clearly cannot be achieved without addressing antibiotic resistance. Recent striking progress has been made towards this: in 2019, the U.S. Food and Drug Administration (FDA) approved a new combination therapy composed of bedaquiline, pretomanid, and linezolid that able to treat both MDR- and XDR-TB disease effectively using three novel antibiotics, but with higher cost and greater side effects than standard first-line therapy (3). There have been no similar efforts towards improvements in first-line therapy, which is the primary line of defense. Clearly, retaining or improving the efficacy of the inexpensive first-line drugs for the treatment of drug-sensitive TB is imperative, as is enhancing their activity where possible to reduce side effects and thereby improve patient adherence.

Genome maintenance is essential for bacterial survival. Both transcription and replication processes inflict stresses that challenge genome topology and integrity. R-loops, which are RNA/DNA hybrids that form spontaneously in the genome during transcription, are a major threat to genome stability and can be lethal if not resolved (4, 5) (Fig 1A). Persistent R-loops can alter DNA topology, block both transcription and replication, and promote double-strand DNA breaks (DSBs) (6-9). Not all R-loops are pathological, since R-loop formation can be a necessary intermediate in plasmid replication (10) and in bacterial immune system surveillance via CRISPR (Clustered Regularly Interspersed Short Palindromic Repeats) (11). Multiple helicases and endonucleases can contribute to the resolution of these hybrids, but ribonuclease HI (RNase HI) activity provides the major dedicated function for this, by

targeted hydrolysis of the RNA strand in an RNA/DNA hybrid (12). Cells therefore require precisely controlled levels of RNase HI to prevent dysfunction. Recently it was demonstrated that loss of RNase HI function in *E. coli* drives the extinction of rifampicin and streptomycin-resistant strains (13), suggesting that RNase HI might be a promising new antimicrobial target in light of the current antimicrobial resistance crisis. RNase HI is already a validated target for HIV therapy: Much progress has been made over the last two decades in the discovery and optimisation of inhibitors with different binding modes that are specific and selective to the RNase H activity of HIV-1 RT and that have antiviral activity within cells (14-16).

The mycobacteria possess a bifunctional RNase HI enzyme called RnhC that is composed of an N-terminal domain with RNase HI activity and a C-terminal domain from the acid phosphatase family. This domain possesses cobalamin-P phosphatase activity (CobC) and is proposed to be involved in vitamin B_{12} biosynthesis (17-20). These domains can function separately, although the CobC domain confers increased activity on the RNase HI domain of *M. tuberculosis* RnhC (Rv2228c) (17). The mechanism for this is not understood. Although most mycobacteria have only one gene encoding RNase HI activity, the non-pathogenic saprophytic model organism, *M. smegmatis*, also possesses a single-domain RNase HI termed RnhA in addition to RnhC. Both RNases HI have been well characterised biochemically, as have single knock-out strains of both *rnhA* and *rnhC* (18, 19, 21), which both grow indistinguishably from wild type (18, 19). However, simultaneous deletion of both genes is lethal, indicating the essentiality of RNase HI activity in mycobacteria (18, 19). Transposon mutagenesis in *M. tuberculosis* located essentiality of function to the RNase HI domain of RnhC (22) (23). Moreover, *M. tuberculosis rnhC* can rescue *M. smegmatis* from the lethality of the double knockout of *rnhA* and *rnhC* (20).

In this study we used the dual RNase HI capacity of *M. smegmatis* provided by RnhA and RnhC to investigate RNase HI as a novel drug target in mycobacteria. We used single RNase HI knock-out strains to deplete RNase HI activity , and showed that deletion of RnhC in particular is sufficient to cause the accumulation of R-loops, and to induce markers of genome topological stress. Furthermore, deletion of RnhC enhanced the killing activity of various anti-tubercular drugs, with an especially marked effect on rifampicin activity. We also identified four small molecules, previously shown to inhibit HIV RNase HI, as micromolar inhibitors of recombinant *M. tuberculosis* RnhC *in vitro*, that also potentiate

rifampicin killing in whole *M. tuberculosis* cells. Together, these results validate RNase HI as a drug target in the mycobacteria, demonstrate the potential for enhanced therapy by combining rifampicin treatment with RNase HI inhibitors, and open avenues to rescue rifampicin-resistant strains for therapy with rifampicin through RNase HI inhibition.

Materials and methods

Strains, growth conditions and media

Strains and plasmids used in this study are listed in Tables S1 and S2. *E. coli* strains Top10 and BL21(DE3)pLysS, used for cloning and protein expression respectively, were grown in lysogeny broth (LB) and/or agar at 37°C or 28°C in the presence of the appropriate antibiotic(s) at the required concentrations *i.e.* 50 µg ml^{-1} kanamycin (Goldbio), 100 µg ml^{-1} ampicillin (Goldbio), 200 µg ml^{-1} hygromycin (Thermofisher), 5 µg ml^{-1} gentamicin (Goldbio) and 35 µg ml^{-1} chloramphenicol (Sigma-Aldrich). *M. smegmatis* mc^2155 and *M. tuberculosis* H37Rv were grown at 37°C in Middlebrook 7H9 broth (BD) or on Middlebrook 7H10 agar (BD) supplemented with 10% Albumin-Dextrose-Saline, 0.2% glycerol and 0.05% Tween 80 (Sigma) in the presence of appropriate antibiotics at the following concentrations: 25 µg ml^{-1} kanamycin, 50 µg ml^{-1} hygromycin and 5 µg ml^{-1} gentamicin.

Plasmid and Strain construction

Plasmids were constructed using standard cloning techniques. Inactivation of *rnh* alleles was carried out by homologous recombination with suicide plasmids (24) containing alleles inactivated by insertion of either a hygromycin antibiotic resistance cassette (Δ*rnhC*) or kanamycin resistance cassette (Δ*rnhA*). Allele replacement was confirmed initially by Southern blot and subsequently by whole-genome sequence analysis. Bioluminescent strains were constructed by transformation with the integrating plasmid pMV306G13+Lux which carries the *luxABCDE* operon (25). *M. smegmatis rnhA* or *M. tuberculosis rnhC* genes for complementation studies were cloned into the pAL5000-based replicating plasmids (26) pOLYG or p32GoriM, respectively, to make p*rnhA* and p*rnhC*. These vector backbones differ only in the antibiotic resistance cassette, hygromycin or gentamycin respectively (27). The *M. smegmatis rnhA* gene was cloned with 350 bp of upstream region containing its own promoter (28), and *M. tuberculosis rnhC* was expressed from the well-characterised leaderless Ag85 promoter (29).

158 **Immunodetection of RNA/DNA hybrid**

159 Bacterial strains were grown as above and harvested at mid-log phase (OD_{600} =1.5). Aliquots
160 (20 ml) of each culture were harvested and the nucleic acid was extracted using the phenol-
161 chloroform technique (30) with a slight modification: Following lysozyme treatment, RNase
162 A (10µg/ml) was added to degrade single-stranded RNA, and small RNAs were precipitated
163 by 30% (w/v) PEG 8000/30 mM $MgCl_2$. Nucleic acid was quantified using a
164 spectrofluorometric method using QuantiFluor® dsDNA (Thermofisher). Nucleic acid (400
165 ng) was spotted on to Amersham Hybond N+ nylon membrane (GE Healthcare Lifesciences),
166 air-dried and UV cross-linked for 3 min using a transilluminator. Control aliquots were pre-
167 incubated with 10 U of *E. coli* RNase H (NEB) for two hours at 37 °C before spotting. The
168 membrane was blocked in TBST (20mM Tris pH 7.4, 150mM NaCl and 0.1% (v/v) Tween
169 20) containing 5% (w/v) milk powder at room temperature for 1 hour. The blocking solution
170 was removed and replaced with fresh TBST containing 5% (w/v) milk and S9.6 anti-DNA-
171 RNA Hybrid Antibody (Kerafast) at 1µg/ml and incubated at 4°C overnight. The following
172 day, the membrane was washed (3 x 5 min) with fresh TBST, and finally incubated with
173 1:1000 dilution of HRP-conjugated anti-mouse secondary antibodies (Invitrogen) in TBST
174 for 1 hour at room temperature. Unconjugated secondary antibody was removed by washing
175 (3 x 5 min) with fresh TBST, and developed with West Pico PLUS Chemiluminescent
176 Substrate (Thermofisher) according to manufacturer instructions. Chemiluminescence was
177 detected using an Amersham Imager 600, and dot intensity was quantified using the Image
178 Lab software suite (BioRad). Student's unpaired *t*-test, calculated using GraphPad Prism was
179 used to test for significance.
180
181 **β-galactosidase reporter assay**
182 Cultures of *M. smegmatis* mc^2155 and Δ*rnhC* were grown as above and harvested at mid-log
183 phase. Optical density was used for normalisation of β-galactosidase activity. The assay was
184 carried out according to Miller with variations (31, 32). Briefly, aliquots (2 ml) from each
185 culture were harvested and resuspended in Z-buffer (60mM Na_2HPO_4, 40mM NaH_2PO_4,
186 10mM KCl, 1mM $MgSO_4$, 50 mM β-mercaptoethanol) to a final volume of 600µl. Cells were
187 permeabilized by adding 100µl of chloroform and 50 µl of 0. 1% SDS, and the mixture was
188 incubated at 28°C for 5 minutes. Then 200µl of ONPG (4 mg/ml) was added to the mixture,
189 which was incubated at 28°C until sufficient yellow colour developed, and time of incubation
190 was recorded. The reaction was stopped by the addition of 500µl of 1 M Na_2CO_3 then the
191 mixture was centrifuged at 10 000 g for 2 min. The OD_{420} of the supernatant was measured

and the β-galactosidase activity was expressed in Miller Units using the following equation: $OD_{420}(1000)/(tv)$ OD $_{600}$, where t = incubation time (min) and v = volume of culture (ml). The Student's unpaired t-test embedded in GraphPad Prism software was used to test for significance.

Growth inhibition assays

M. smegmatis

Two different methods were used to assess growth inhibition.

Agar proportion method

M. smegmatis strains were grown and harvested at mid-log phase (OD_{600} = 1 to 1.5). A single-cell suspension was prepared by centrifugation of the culture at 2,000 x g for 10 min to pellet clumps, the supernatant was collected, and 20% glycerol was added. Aliquots were frozen with liquid nitrogen and stored at -80 ºC. For killing assay, 150 CFU were plated in Middlebrook 7h10 agar plates containing two-fold serial dilutions of rifampicin and isoniazid and incubated at 37°C for 3 days. Bacterial survival at each concentration was assessed and MIC determination was calculated from non-linear regression analysis using GraphPad Prism. Significant differences were assessed using the Student's unpaired t-test embedded in GraphPad Prism software.

Bioluminescence reporter assay

Bacterial strains carrying pMV306G13+Lux were grown as above to mid-log phase (OD_{600} = 1.5), and diluted 30-fold to OD_{600} = 0.05. All assays were set up in 96 well plates (Greiner Bio-One) in a total volume of 100µL where 50µL of the diluted culture was added to 50µL of two-fold serially diluted antibiotics in Middlebrook 7H9 media. Two controls were added, bacteria only and media only. The plates were covered with lids, sealed with parafilm to prevent evaporation, and were incubated at 37°C for 18 hours. Bioluminescence was quantified using a SpectraMax iD3 Plate Reader. Bacterial survival at each concentration was calculated based on the ratio of the RLU in the presence or absence of the inhibitor. Dose response curves and the MIC_{50} for each drug was determined by fitting the data to log (inhibitor) vs. response – variable slope equation using GraphPad Prism 7. The Student's unpaired t-test embedded in GraphPad Prism software was used to test for significant differences in the killing between wild type and Δ*rnhC*.

225

226 ***M. tuberculosis***

227 *M. tuberculosis* $mc^2$6230 (Δ*RD1* Δ*panCD*) (33) was grown in Middlebrook 7H9 broth
228 supplemented with OADC (0.005 % oleic acid, 0.5 % bovine serum albumin, 0.2 % dextrose,
229 0.085 % catalase), 0.05 % tyloxapol and 25 μg mL^{-1} pantothenic acid (PAN) at 37 °C. The
230 assay was carried out in 96 well plates in a total volume of 100μL. The cultures were diluted
231 to OD_{600} = 0.05 and 50μL of the diluted culture was added to 50μL of two-fold serially
232 diluted inhibitor in Middlebrook 7H9 media. Plates were incubated without shaking at 37 °C
233 for 7 days before the minimal inhibitory concentrations that inhibited 90% of growth (MIC_{90})
234 were determined from the visual presence or absence of growth in biological triplicates. The
235 fractional inhibitory concentration index (FICI) was determined for RNase HI inhibitors in
236 combination with rifampicin by the addition of a sub-MIC concentration of rifampicin (5 nM)
237 to the MIC assay. FICI values were calculated using the equation $FICI = FIC_A + FIC_B =$
238 $(C_A/MIC_A) + (C_B/MIC_B)$, where MIC_A and MIC_B are the MICs of drugs A and B alone, and
239 C_A and C_B are the concentrations of the drugs in combination.

240

241 **Recombinant RnhC purification**

242 The *RnhC/Rv2228c* open reading frame (ORF) was amplified from *M. tuberculosis* DNA and
243 cloned into pMAL-C2 (34) as an N-terminal maltose-binding protein (MBP) fusion construct
244 which was further modified to include a 6xHis-tag following the MBP ORF and a 3C
245 protease recognition sequence (35) immediately upstream of *RnhC*. The plasmid was
246 transformed into *E. coli* BL21(DE3) pLysS (36). Cultures grown in ZYP-5052 medium (37)
247 were auto-induced by incubation for 7 hr at 37 °C followed by incubation at 28 °C for 20 hr.
248 The cells were harvested (5,300 x g for 30 min at 4 °C) and re-suspended in lysis buffer (20
249 mM HEPES and 150 mM NaCl, pH 7.5) containing 10μg/ml of DNase 1 and complete Mini
250 EDTA-free protease inhibitor tablets (Roche) (one tablet/10ml total volume). The cells were
251 lysed using a French press cell disruptor (ConstantSystem) at a pressure of 18 kpsi. The
252 lysate was centrifuged at 20,000 × g for 30 min at 4 °C. Recombinant protein was purified by
253 affinity chromatography using an amylose column, and cleaved from the tags (MBP and 6X-
254 His) using His-tagged recombinant 3C protease (38) overnight at 4 °C. The 3C protease and
255 the tags were removed by reverse affinity chromatography. Recombinant Mt RnhC was
256 concentrated and further purified by size exclusion chromatography using a

16/600Superdex® 200 pg columns (GE Healthcare). Mt RnhC eluted in a single peak and the fractions were pooled, concentrated to 10 mg.ml^{-1} and stored at -20°C.

HIV-RNase HI Inhibitors

HIV RNase H inhibitors were selected from the PubChem database and kindly provided by the National Institutes of Health (NIH) through the Developmental Therapeutics Programme (DTP). The compounds were resuspended in DMSO at a final concentration of 10 mM and stored at -20°C.

RNase HI inhibition assay

The RNase HI activity assay was performed using a Fluorescence Resonance Energy Assay (FRET) as previously described in (39) using an 18bp RNA/DNA , where the RNA strand was labelled with fluorescein on the 3′-end (5'-GAU CUG AGC CUG GGA GCU-fluorescein-3') and the complementary DNA strand labelled with Iowa black on the 5′-end (5' -Iowa black-AGC TCC CAG GCT CAG ATC-3'). Mt RnhC (4nM) was incubated with 25nM of substrate in FRET buffer (50 mM Tris-HCl, pH 8, 60 mM KCl, 5mM $MgCl_2$). The increase in fluorescein signal due to substrate hydrolysis was monitored at an excitation/emission of 490nM/528nM using a SpectraMax iD3 Plate Reader. Compounds active at an initial inhibitor concentration of 100 µM were subsequently assessed using a dose response assay. Drug dose-response curves and IC_{50} determinations were carried out by non-linear regression analysis in GraphPad Prism 7.

Results

Loss of RnhC induces markers of DNA topology stress in *M. smegmatis*

As a first indicator of whether depletion of RNase HI activity affected genome maintenance, we looked at the expression of topoisomerases responsible for topology control. In metabolically active cells, R-loops can stall RNA polymerase complexes and increase topological stress in the genome because persistent R-loops prevent the resolution of both the negative DNA supercoiling that is formed in the wake of an RNA polymerase, and the positive DNA supercoiling that is generated ahead of it (40, 41) (**Figure 1A**). Both the DNA gyrase and topoisomerase I promoters in *M. smegmatis* respond directly to the topological stresses experienced by the chromosome since RNA polymerase binding elements in each promoter adopt optimal conformations for transcription if the DNA is either overwound (gyrase)(42), or underwound (topoisomerase I) (43). An increase in the expression of both promoters simultaneously would therefore indicate topological stress in the chromosome consistent with the presence of persistent R-loops. We fused these promoters to a *lacZ* reporter gene and integrated these reporter cassettes into the *attB* site of the chromosomes of both wild-type and Δ*rnhC M. smegmatis* strains. Reporter activity from gyrase and topoisomerase promoters was increased in the Δ*rnhC* strain compared to wild-type, by 3-fold and 1.7-fold respectively ($p < 0.05$) (**Figure 1B**). Notably, these increases in activity were abolished in the complemented knockout strain, suggesting that chromosomal R-loops were indeed effectively processed by the complementing *rnhC* gene provided *in trans*.

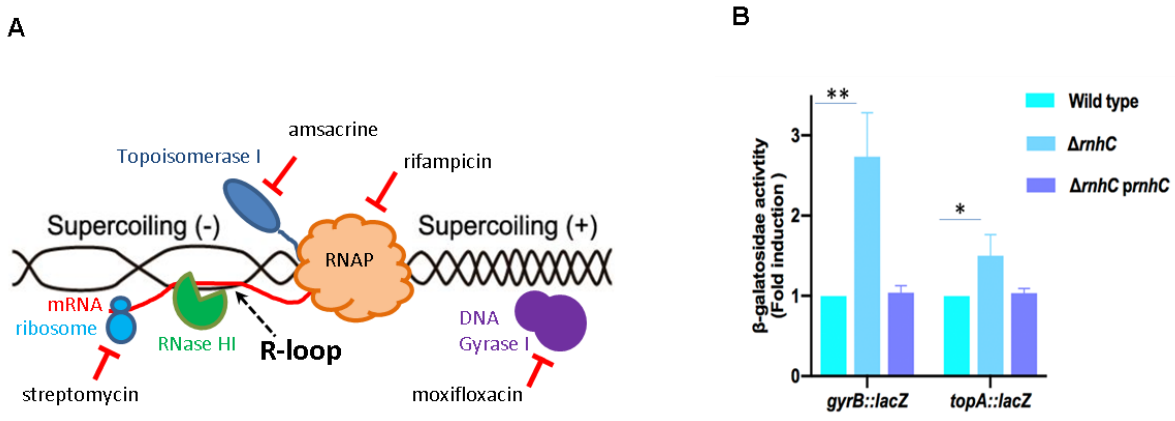

Figure 1. **A**. Twin domain model of transcription and location of enzymes (and their inhibitors) involved in topological modification or R-loop metabolism. RNAP: RNA polymerase I. **B**. Expression of gyrase (*gyrB*) and topoisomerase I (*topA*) promoters fused to *lacZ* in wild type, Δ*rnhC* and complemented strains of *M. smegmatis* mc^2155 . The relative activities of the β-galactosidase reporters are shown, normalized to the wild type. Data shown are representative of the average of three independent experiments with standard deviations indicated by error bars. The statistical significance of the differences assessed using Student's unpaired *t*-test calculated by GraphPad Prism 7. *, $P < 0.05$; **, $P < 0.01$.

Loss of either *rnhA* or *rnhC* results in R-loop accumulation in *M. smegmatis*.

To quantify R-loops in the RNase HI depletion strains, we isolated total nucleic acid from cultures of both Δ*rnhA* and Δ*rnhC* strains grown to mid-log phase and quantified the relative amounts of R-loops in each, using South-Western blots with a monoclonal antibody specific for DNA:RNA hybrids (44). The specificity of the antibody for RNA:DNA hybrids was confirmed by subjecting samples to RNase HI digestion prior to detection and noting the loss of signal (**Figure 2A**). R-loops were detected at low levels in the parental strain, and accumulated in the Δ*rnhA* mutant by a factor of ~2-fold, and in the Δ*rnhC* mutant by a factor of ~11-fold (**Figure 2B**). This indicated that both RnhA and RnhC act as part of a protective system to remove R-loops in replicating and actively transcribing cells. In both cases, the accumulation of R-loops also indicated that neither the remaining RNase HI isoform, nor other potential hybrid removal mechanisms such as the DNA helicase RecG (45), are able to compensate fully for the loss of either RNase HI enzyme.

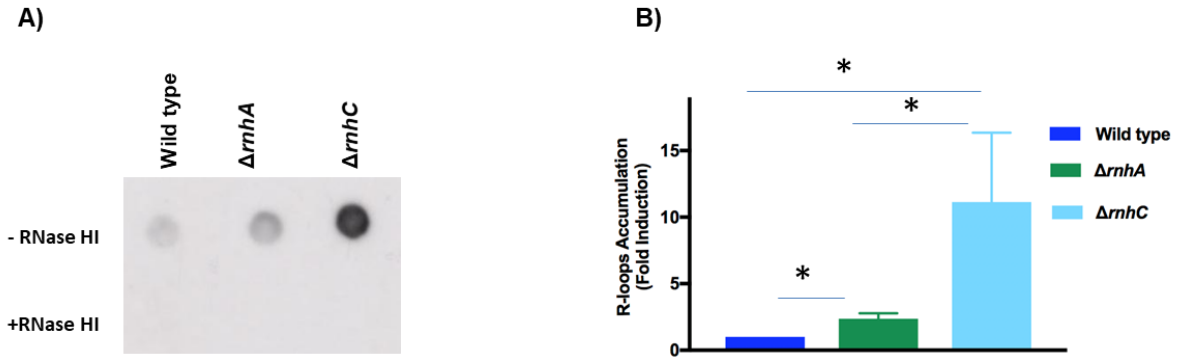

Figure 2. R-loop quantitation in wild type, Δ*rnhA* and Δ*rnhC* strains of *M. smegmatis* mc^2155. **A.** Dot-blot analysis of R-loop accumulation: total nucleic acid from each strain was spotted onto the membrane and detected using a RNA:DNA hybrid-specific monoclonal antibody. Controls (+RNase HI) were treated with *E. coli* RNase HI before spotting. **B.**The amounts of R-loops in the wild type, Δ*rnhA*, and Δ*rnhC* strains were quantitated using Image Lab (BioRad). Relative amounts are shown normalized to the wild type. Data shown are representative of the average of three independent experiments with standard deviations indicated by error bars. The statistical significance of the differences are assessed using Student's unpaired *t*-test calculated by GraphPad Prism 7. *, $P < 0.05$

Plasmid replication exacerbates the R-loop burden in the *M. smegmatis* Δ*rnhC* strain.

To confirm that loss of RNase HI activity was responsible for the observed increase in the number of R-loops in the Δ*rnhA* and Δ*rnhC* strains, we tested whether either *rnhA* or *M. tuberculosis rnhC* provided *in trans* could complement the Δ*rnhA* and the Δ*rnhC* mutants,

respectively (**Figure S1 A/B**). As expected, complementation with the plasmid-borne *rnhA* gene reduced levels of R-loops in Δ*rnhA* to a wild-type level. In contrast, introduction of a plasmid carrying *rnhC* into the Δ*rnhC* strain conferred a 3-fold *increase* of R-loops over Δ*rnhC* without plasmid. A similar effect was obtained by transforming the Δ*rnhC* strain with the plasmid containing *rnhA*, or with vector containing the cobC domain only (**Figure S1C**). This suggested that the increase in R-loops was a function of plasmid carriage in Δ*rnhC*, and possibly unrelated to plasmid cargo.

We carried out whole genome sequencing of the *ΔrnhC* strain complemented with *rnhC* which determined that this strain was isogenic to *ΔrnhC*, except for the plasmid, and that the plasmid was present at between 24-48 copies per cell. Typically, pAL5000-based plasmids are present at 3-5 copies per cell (46), but a high copy number variant has been isolated. No mutations were found in the plasmid compared to the published sequence that could account for the increased copy number (47). This suggested that loss of RnhC activity affected plasmid copy number in *M. smegmatis* through a physical rather than genetic mechanism. Since the chromosomally-located reporters *gyrB::lacZ* and *topA::lacZ*, showed wild-type levels of topoisomerase activity in Δ*rnhC* p*rnhC*, but only report on the topology status of the chromosome, we concluded that the increased number of R-loops detected in the complemented *ΔrnhC* strain must therefore be plasmid-borne. It is possible that increased hyper-negative supercoiling of the plasmid, either through localised transcription effects (48) and/or titration of RNase HI activity to the chromosome (49), may have affected binding efficiencies of some of the proteins involved in plasmid replication, resulting in increased initiation frequencies, as well as promotion of R-loop formation.

Regardless of the precise mechanism, since the chromosomal DNA topology stress resulting from loss of *rnhC* could be fully complemented, we concluded that the plasmid copy number was well suited to the cellular requirements for RNase HI activity, and the supercoiling status of the plasmid might even reflect a useful cellular feedback mechanism to titrate the level of RNase HI appropriately. For these reasons, subsequent experiments were carried out using these constructs and strains. The contrasting behaviour of the complemented Δ*rnhA* and Δ*rnhC* strains also underscored the different physiological consequences of the loss of *rnhA versus rnhC* which may be a function of the severity of R-loop accumulation.

Inhibition of transcription by low levels of rifampicin unexpectedly increases the accumulation of R-loops.

Since transcription is a prerequisite for R-loop formation, we hypothesized that transcriptional inhibition would reduce the number of R-loops in the cell. Rifampicin inhibits transcription at the initiation step and does not affect RNA polymerase complexes that have added more than three ribonucleotides from proceeding to active transcription (50).

We exposed mid-log phase cultures of wild-type and Δ*rnhC* strains of *M. smegmatis* to rifampicin for 1 hour before harvesting the cells and quantitating the R-loops present in them. Rifampicin at a concentration equivalent to the MIC_{90} value for *M. smegmatis* (4.2 µM or 3.5 $\mu g.ml^{-1}$) abolished R-loops, a result consistent with complete inhibition of transcription (**Figure 3**). Surprisingly, both the wild-type and Δ*rnhC* strains contained significantly increased amounts of R-loops when exposed to a concentration of rifampicin equivalent to the MIC_{50} value (1.2 µM or 1 $\mu g.ml^{-1}$). A pronounced response to rifampicin exposure was shown by the Δ*rnhC* strain, with the number of R-loops increasing by ~3-fold compared to the non-treated Δ*rnhC* strain. Strikingly, even at a rifampicin concentration equivalent to MIC_{25} (0.6 µM or 0.5 $\mu g.ml^{-1}$), more R-loops accumulated in Δ*rnhC* cells compared to untreated cells, indicating an overload of the cellular systems required for R-loop resolution. This also suggested that the low number of R-loops observed in wild-type cells at this rifampicin concentration was not due to lack of R-loop formation, but rather was reflective of efficient resolution by RNase HI. Notably, R-loop resolution is clearly impaired even in wild-type cells exposed to rifampicin at the MIC_{50}. These results indicate that under these conditions, sub-lethal concentrations of rifampicin cause R-loop accumulation, a phenomenon that has not previously been described, and is opposite to the effect of a lethal concentration of rifampicin.

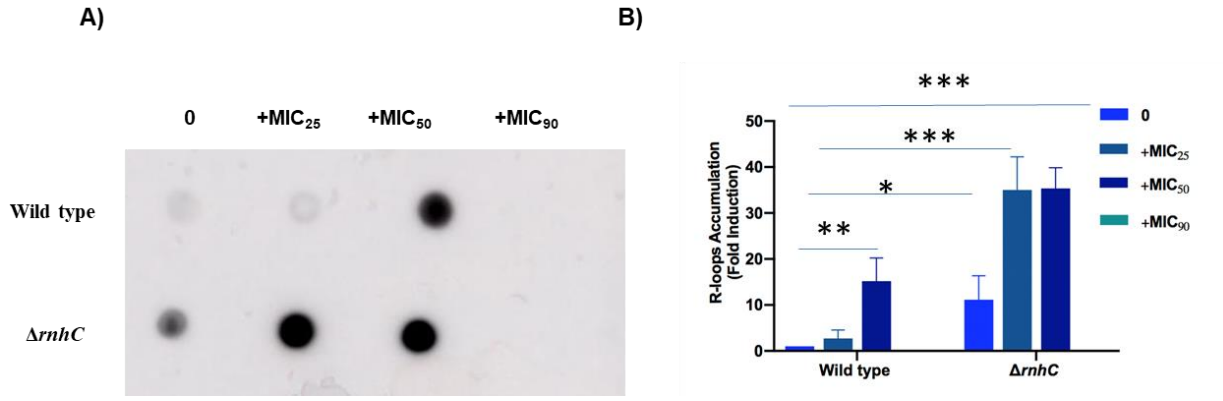

Figure 3. R-loop quantitation in wild-type and Δ*rnhC* strains of *M. smegmatis* mc^{2}155 after exposure to various concentrations of rifampicin. **A.** Dot blot analysis of R-loop formation. **B.** Quantification of R-loops observed in the dot blot, normalized to the wild type strain in the absence of rifampicin. Data shown are representative of the average of three independent experiments with standard deviations indicated by error bars. The statistical significance of the differences assessed using Student's unpaired *t*-test calculated by GraphPad Prism 7. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$

Loss of RNase HI is synergistic with transcriptional inhibition.

To better understand the cellular consequences of rifampicin-induced R-loop accumulation, we carried out a dose-response assay to assess the extent to which RNase HI contributes to the survival of cells exposed to rifampicin. We found that wild-type *M. smegmatis* is inhibited as expected by previously established concentrations of rifampicin (51, 52). In contrast, loss of *rnhC* strongly sensitised the knock-out strain to rifampicin, reducing the MIC_{50} almost 100-fold, from 1.2 μM to 18 nM (**Figure 4A**). This sensitivity was fully complemented by the provision of *rnhC in trans*. To confirm that this remarkable loss of viability under rifampicin stress is due to the loss of the RNase HI domain rather than the CobC domain of RnhC, which is involved in vitamin B_{12} biosynthesis (20), the assay was repeated in the presence of vitamin B_{12} in the growth medium. Supplementation with vitamin B_{12} altered the MIC_{50} slightly. but not significantly, from 18nM to 30nM, (**Figure 4A**), indicating that the rifampicin sensitisation was due to the depletion of RNase HI function.

As the Δ*rnhA* strain of *M. smegmatis* showed a smaller accumulation of R-loops than the Δ*rnhC* strain (**Figure 2**), we asked whether the loss of RnhA could also confer increased sensitivity to rifampicin. *M. smegmatis* Δ*rnhA* showed a ~3-fold increase in sensitivity to rifampicin compared to wild-type, which was fully complemented by the provision of *rnhA in trans* (**Figure 4B**). Together, these findings indicated that depletion of RNase HI activity was responsible for the increased sensitivity of *M. smegmatis* to transcriptional inhibition, and that loss of RnhA had a smaller contribution to this phenotype than the loss of RnhC. In fact, it

was remarkable that such an apparently small contribution to R-loop metabolism still conferred synergy with rifampicin, underscoring the vulnerability of mycobacterial cells to even small amounts of RNase HI depletion.

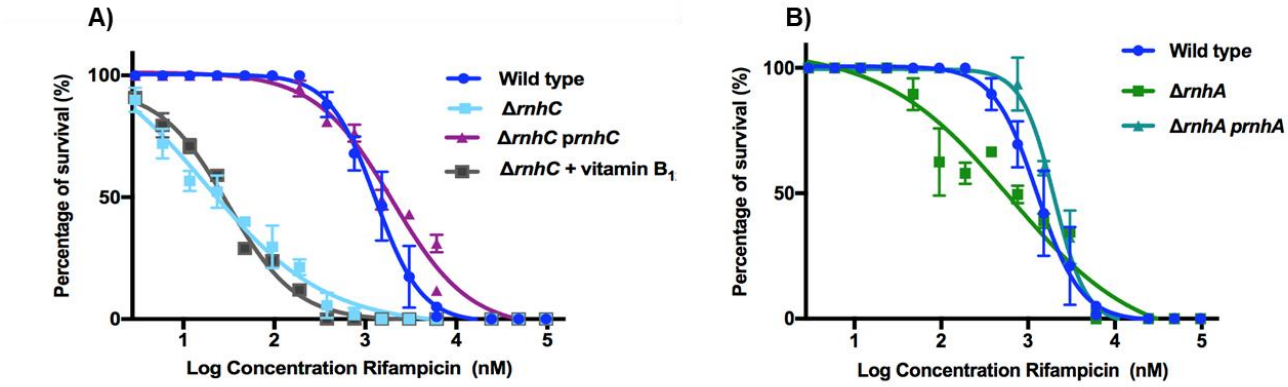

Figure 4: Dose-response curves of rifampicin killing in **A**) wild-type, Δ*rnhC* and **B**) Δ*rnhA* strains of *M. smegmatis* mc^{2}155 and the respective complemented strains. Data shown are representative of the average of three independent experiments with standard deviations indicated by error bars. Vitamin B_{12} was included in the growth medium at 10 µg.ml^{-1}.

Loss of RNase HI also sensitizes *M. smegmatis* to moxifloxacin and streptomycin.

R-loop formation can also be affected by uncoupling transcription and translation (53, 54), or by altering DNA supercoiling in the genome (5, 55). Therefore, we tested the sensitivity of the *M. smegmatis* Δ*rnhC* strain to streptomycin (an inhibitor of translation), amsacrine (a topoisomerase I inhibitor) and moxifloxacin (a DNA gyrase inhibitor) (**Figure 1A**). As a control, we investigated the sensitivity of the Δ*rnhC* strain to isoniazid, which disrupts mycobacterial cell wall synthesis by inhibiting mycolic acid synthesis (56).

The loss of the *rnhC* gene significantly reduced the MIC_{50} for streptomycin, from 31 nM in the wild-type to 17 nM in the Δ*rnhC* strain ($P < 0.05$), and to moxifloxacin, from 194 nM to 62 nM (**Figure 5A,B**). ($P < 0.05$). In contrast, the loss of *rnhC* did not result in a significant increase in the sensitivity of *M. smegmatis* to amsacrine (MIC_{50} = 160 µM for the wild-type and 140 µM for the Δ*rnhC* strain, $P > 0.05$) (**Figure 5C**). As expected, the loss of the *rnhC* gene did not significantly affect the MIC_{50} to isoniazid (85µM for wild-type and 63 µM for the knock-out strain, $P > 0.05$) (**Figure 5D**).

Taken together, these results suggest that RNase HI inhibition would improve the efficacy of rifampicin, streptomycin and moxifloxacin in a synergistic manner. In all cases, the increase in efficacy arising from RNase HI depletion was more pronounced at sub-MIC concentrations of antibiotic, which is consistent with a requirement for some transcription to occur for the effect to be seen.

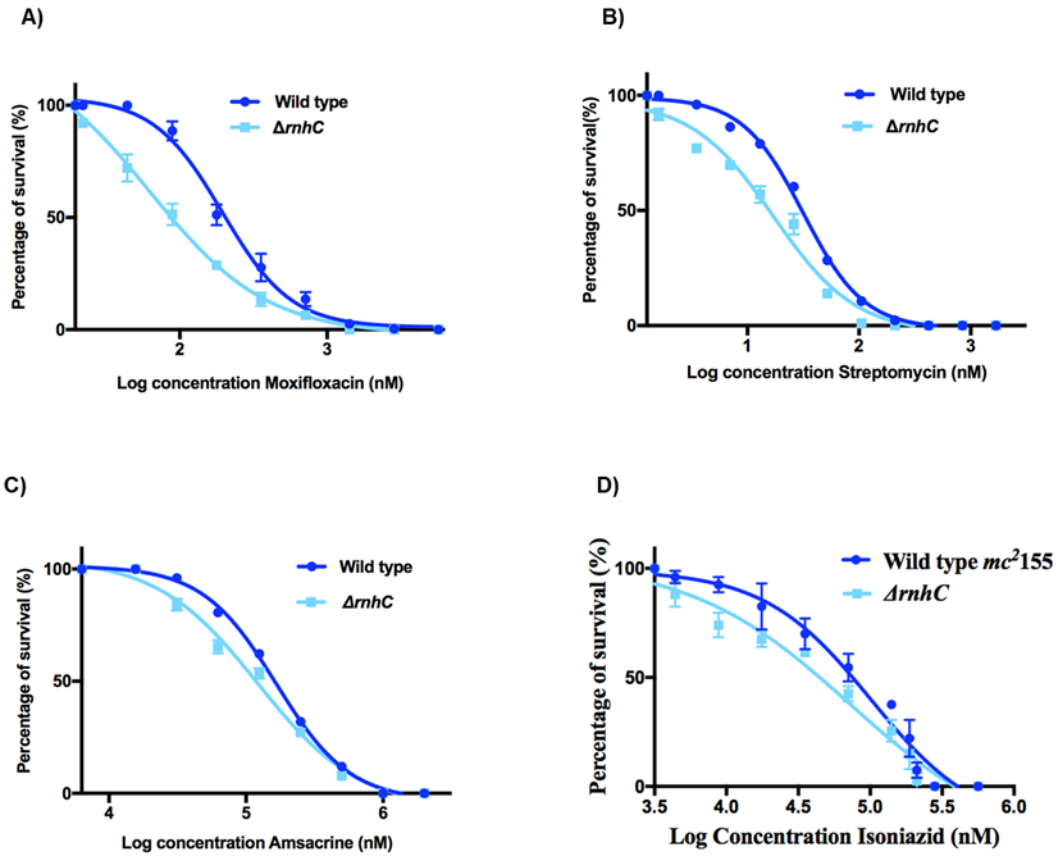

Figure 5. Dose-response curves of antibiotic killing for **A**) streptomycin, **B**) moxifloxacin,**C**) amsacrine and B) isoniazid in wild-type and ΔrnhC strains of *M. smegmatis* mc^2 155. Data shown are representative of the average of three independent experiments with standard deviations indicated by error bars.

Drug candidates with diverse scaffolds inhibit recombinant *M. tuberculosis* RNase HI.

The enhanced antibiotic activity observed when RNase HI activity is depleted in *M. smegmatis* indicated that RNase HI inhibition is likely to be compatible with existing antibiotics used for *M. tuberculosis* therapy. Effective RNase HI inhibitors would be of particular interest as synergistic partners for rifampicin, which is a key first-line antibiotic used to treat tuberculosis. Rifampicin is a powerful inducer of hepatic cytochrome P450 enzymes, which leads to antagonistic interactions with many other drugs (57). Hence, any drug that sensitizes *M. tuberculosis* to rifampicin also has the potential to improve therapeutic outcomes by decreasing the dose of rifampicin required.

Recently, specific inhibition of HIV RNase HI was reported (14-16), preventing viral replication in cells, and supporting the possibility that small molecule compounds could be similarly discovered for bacterial RNases HI. The drug discovery effort against HIV RNase HI has identified promiscuous chemical scaffolds that not only have activity on HIV RNase HI *in vitro*, but also have activity on *E. coli* RNase HI or human RNase HI. We therefore selected a small library of 33 inhibitors of the RNase H activity of HIV-1 RT (**Table S3**) which we screened for activity against recombinant *M. tuberculosis* RNase HI. Ten compounds were found to inhibit *M. tuberculosis* RNase HI at a concentration less than 100μM (**Table S3**).

Rifampicin synergy in whole-cell screens can eliminate off-target compounds.

The ten most effective inhibitors of *M. tuberculosis* RNase HI *in vitro* were then tested for their antimicrobial activity against *M. tuberculosis* in a growth inhibition assay. All showed weak or no growth inhibition of *M. tuberculosis* when tested alone, which could be consistent with their modest inhibition of RNase HI *in vitro,* or with off-target activity. However, we hypothesised that on-target inhibition of *M. tuberculosis* RNase HI *in vivo* would phenocopy the genetic depletion of RNase HI observed in *M. smegmatis,* by increasing the sensitivity of *M. tuberculosis* to sub-MIC_{50} concentrations of rifampicin. Four compounds (NSC353720, NSC600285, NSC18806, and NSC99726) significantly potentiated killing of *M. tuberculosis* by sub MIC rifampicin indicating that they are likely to be inhibiting RNase HI *in vivo* (**Figure 6A and Table 1**). *In vitro* dose-response curves against *M. tuberculosis* RNase HI show that NSC353720 (IC_{50} = 14 μM) was the most potent inhibitor, while NSC99726 was the weakest inhibitor (IC_{50} = 45 μM) of this compound set (**Figure 6B**).

Table 1. Inhibition of *M. tuberculosis* RNase HI recombinant protein and *M. tuberculosis* whole cells by *M. tuberculosis* RNase HI inhibitors and their interaction with rifampicin.

	M. tuberculosis RnhC IC_{50} (µM)	*M. tuberculosis* MIC_{90} (µM)		
		Compound alone	Compound with 5 nM Rifampicin	FICI#
NSC353720	12	>512	64	0.3125
NSC353681	17	256	256	2
NSC600285	17	512	64	0.375
NSC600286	32	64	64	2
NSC143101	0.4	256	128	0.75
NSC18806	32	>512	128	0.5
NSC99726	45	512	64	0.375
NSC80693	11	256	256	1
NSC51535	2.8	>512	512	1
NSC117949	29	>512	512	1

Interpretation of FICI: The cut off for synergistic effect is ≤0.5, cut off for additive or indifferent effect is FICI >0.5 and <2 and cut off for antagonistic effect is FICI >2.

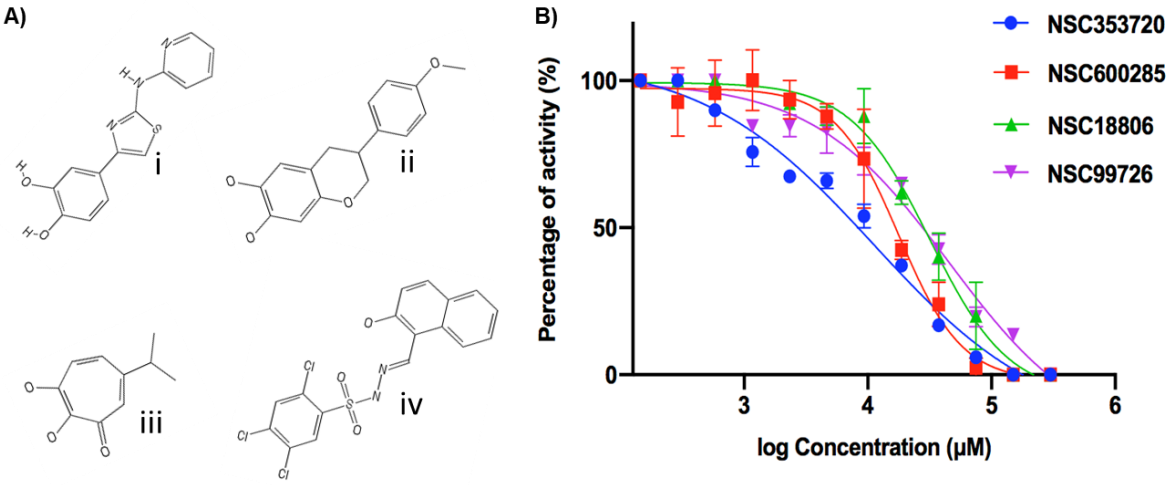

Figure 6. Chemical structures (**A**) and dose-response curves (**B**) for NSC353720 (i), NSC600285 (ii), NSC18806 (iii) and NSC99726 (iv) against recombinant *M. tuberculosis* RNase HI. Data shown are the average of two independent experiments with standard deviations indicated by error bars

Discussion

R-loops are genotoxic stresses for all cells and the resolution of these structures is critical for cell survival. RNase HI is the primary enzyme that resolves R-loops, although cells have evolved multiple mechanisms to reduce their occurrence or to repair them. In *M. tuberculosis*, RNase HI activity is essential for the growth of the bacterium *in vitro*, which makes it a potential drug target. Nevertheless, not much is known about the cellular consequences of RNase HI inhibition in the mycobacteria. In this study, we showed for the first time that depletion of RNase HI activity in *M. smegmatis* caused an accumulation of R-loops. Loss of *rnhC* had a greater effect on R-loop accumulation than loss of *rnhA*.

The involvement of RnhC, and to a lesser extent RnhA, in the resolution of R-loops in *M. smegmatis* indicates that both enzymes are part of a protective system to remove the RNA/DNA hybrids that form spontaneously in the genome during transcription, before they induce DNA damage. The ability of *M. tuberculosis rnhC* to fully complement the Δ*rnhC* phenotype in *M. smegmatis* strongly suggests that RnhC carries out the equivalent role in both bacteria. Our data is consistent with previous reports that showed loss of RNase HI activity in *Bacillus subtilis* (58), *Saccharomyces cerevisiae* (59, 60) or human cells (61) resulted in the accumulation of R-loops. R-loop formation is favoured by high G+C content, G/C skew and is promoted by G4 quadruplex structures. Although sites susceptible to R-loop formation have not yet been mapped in the mycobacteria, over 10,000 sites for G4-quadruplexes in the *M. tuberculosis* genome have been predicted (62). The high (>65%) GC content of the mycobacterial genomes is likely to favour both the formation and the persistence of R-loops. If mycobacteria are more prone than other bacteria to R-loop formation due to their high GC content, this could account for the essential nature of RNase HI in the mycobacteria since R-loops promote replication-transcription collisions (6, 7, 9, 58, 63), DNA recombination (6, 59, 64-66), DSBs (9, 66-68), and gene silencing.

The striking finding that low levels of transcriptional inhibition promote R-loop accumulation gives new insight into R-loop metabolism in general. It is well established that polymerases which stall or backtrack during transcription are susceptible to forming R-loops (40, 41, 69-71), and that R-loops formation is able to stall a translocating polymerases. The “twin domain” model of transcription shows that DNA in the wake of RNA polymerase becomes negatively supercoiled (*i.e.* underwound) and that DNA ahead of it becomes positively supercoiled (*i.e.* overwound) (72-74). Hence, in actively transcribed genes which contain

multiple polymerase complexes, the wake of unwound DNA from one polymerase is rewound by the following enzyme, thus reducing the need for topology-modifying enzymes and protecting this inter-polymerase region of DNA from R-loop formation (**Figure 7A**). Thus a topological balance exists in the DNA between actively transcribing RNA polymerases, which is disrupted if a polymerases is stalled.

Our results indicate that stalled RNA polymerases in the promoter region promote R-loop formation. It is clear that an R-loop cannot arise directly from the rifampicin-bound polymerase, as no more than 3 nucleotides have been added to the transcript. We hypothesize that the last active polymerase will generate an increasing amount of underwound DNA, which promotes R-loop formation, which stalls this polymerase. A domino effect may thus be created over the length of the gene (**Figure 7B**). The magnitude of this effect depends on the initial stalling event to be close to the beginning of the gene. In contrast, stalling a polymerase near to the end of the gene is unlikely to promote cascading R-loop formation since trailing polymerases will merely stack up neatly behind it until the initial blockage is cleared. Halting all transcription evenly will also have a lesser impact on R-loop formation. It is possible that pause sites in promoter regions might facilitate optimal spacing between polymerase molecules to promote productive transcription by preventing R-loop production. This domino model of formation of R-loops under rifampicin stress is a general, biologically plausible model that nonetheless will be affected by the propensity of the DNA sequence to form R-loops, and the availability of the complementary RNA to bind to DNA.

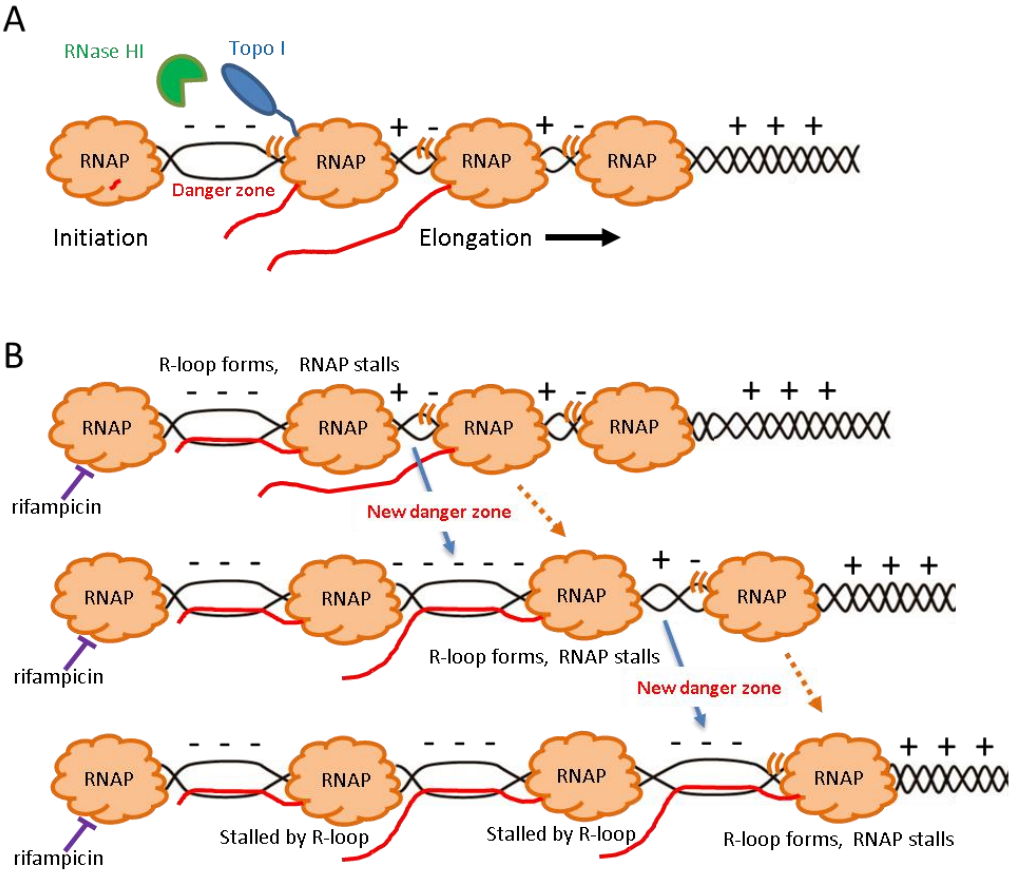

Figure 7 "Domino effect" model of R-loop formation under rifampicin stress. **A.** A negatively supercoiled DNA stretch (- - -) opens between a stationary RNAP (such as one in the promoter region, but can be paused anywhere) and a mobile RNAP (motion indicated by '((‘). This danger zone for R-loop formation is policed by topoisomerase I (blue, shown only once for clarity), which competes with mRNA (red line; only shown twice for clarity) for underwound DNA, and by RNase HI (green). Zones between mobile RNAP are more likely to be neutral (+ -). **B.** Under rifampicin stress, the promoter-bound RNAP cannot start elongating, the danger zone becomes critically underwound, an R-loop forms and stalls its cognate RNAP. A new danger zone opens up between this stalled RNAP and its adjacent mobile RNAP. A new R-loop forms, stalling its RNAP, and the cycle continues until all mobile RNAP molecules are stalled by R-loops. Ribosomes are not shown.

Our observation that transcriptional inhibition leads to the accumulation of R-loops may also open a window into the poorly understood mechanism by which rifampicin causes cell death. Transcriptional inhibition has often been equated to translational inhibition in terms of cellular consequences, as cell death is expected to arise primarily from the lack of protein product (75) and not from lack of transcription *per se*. Transcriptomics is a widely used approach to gain insight into the mode of action of an antibiotic, but as rifampicin inhibits transcription globally, this approach has only been used on rifampicin-resistant strains (76) that may have altered and variable gene expression as a result. As one of the effects of R-loop formation is gene silencing, this may be more profound in some genes than others, and thus RNA sequencing together with R-loop mapping may give alternative insights. Our model

suggests that an important aspect of rifampicin's bactericidal activity may be the genotoxic stresses inflicted by R-loop formation.

We showed that even partial loss of RNase HI activity through the deletion of *rnhC* or *rnhA* synergised extraordinarily well with transcriptional inhibition by rifampicin at concentrations below the MIC_{90} (Fig. 4). This reveals an unappreciated role for RNase HI in promoting survival and possibly persistence under rifampicin stress. Rifampicin is one of the front line drugs for TB therapy (2), and a vital component in resolving bacterial persistence *in vivo* (77). Notably, exposure to sub-MIC concentrations of rifampicin is thought to be instrumental in the emergence of antibiotic resistance in clinical settings. Our results suggest that inhibiting both RNase HI and transcription might act against the emergence of drug resistance, even in sub-MIC concentrations of inhibitors of both RNase HI and RNA polymerase. It also raises the possibility that RNase HI inhibition might rescue rifampicin as a treatment option in rifampicin-resistant bacteria, by increasing their sensitivity to its cellular effects.

Partial loss of RNase HI activity also enhanced killing with moxifloxacin and streptomycin (Fig. 5A&B), and was compatible with isoniazid (Fig. 5D), indicating that in addition to providing a new target for drug development and enhancing first-line therapy, RNase HI inhibition would be a useful adjunct to second-line therapy as well. Uncoupling of transcription and translation is known to enhance R-loop production (54), and ribosomes can actively reverse stalled polymerases that might be more prone to R-loop formation (78). It is noteworthy however, that the combination of translational inhibition and RNase HI depletion does not affect the cell as potently as the combination of transcriptional inhibition and RNase HI depletion. We speculate that in the context of actively translocating RNA polymerases, the opportunity for R-loop formation due to translational inhibition is much reduced compared to the propensity for R-loop formation arising from stalled RNA polymerases, resulting in less synergy.

Many antibiotics, including rifampicin, are antagonistic with moxifloxacin (79) so the synergy of the latter with RNase HI depletion is both intriguing and valuable. We hypothesize that the synergy of moxifloxacin with RNase HI depletion might result from an increased number of gyrase/DNA complexes due to the 3-fold increased expression of DNA gyrase (Fig. 1B), resulting in a higher DSB burden (80).

Our finding that RNase HI depletion does not potentiate amsacrine lethality (Fig. 5C) contrasts with the synthetic lethal effect in *E. coli* of the loss of RNase HI and topoisomerase I (5). This apparent anomaly could be due to various factors. First, the expression of topoisomerase I is less effectively induced in the Δ*rnhC* strain than DNA gyrase I, which would be consistent with a topology-ameliorating effect of R-loops that has recently been proposed (81). Secondly, amsacrine is an uncompetitive inhibitor of topoisomerase I which binds to the topoisomerase I:DNA complex (82). Since topoisomerase I and mRNA are both localised by RNA polymerase, and are both competing for the same underwound DNA substrate, any R-loop formation would effectively remove underwound DNA as a substrate for topoisomerase I, and proportionally reduce the binding of amsacrine to a DNA:topoisomerase I complex, an opposite effect to that seen with moxifloxacin. A final possibility is that, although amsacrine is on-target in eukaryotic cells, and has been shown to inhibit *M. tuberculosis* topoisomerase I *in vitro* (82), it has not formally been shown that it does the same in *mycobacterial* cells. Topoisomerase I is an emerging drug target for development of anti-TB therapeutics. We predict that competitive inhibitors of topoisomerase I, which would more accurately mimic loss of the *topA* gene, might still be effective synergistic partners for RNase HI inhibitors.

The housekeeping nature of the enzymes involved in these synergistic combinations suggests that a wider application for combination therapeutics including RNase HI inhibitors also exists for other organisms. The loss of the *rnhC* gene in *Listeria monocytogenes* correlates with a loss of virulence in mice (58), and in *E. coli*, both streptomycin-resistant strains and rifampicin-resistant strains were competitively disadvantaged in Δ*rnhA* backgrounds (13). This indicates the potential for therapeutic aspects of RNase HI inhibition in bacteria even where RNase HI activity is not essential under laboratory conditions, or where rifampicin tolerance is high.

This study identified four HIV RNase HI inhibitors that have activity against *M. tuberculosis* RNase HI *in vitro* and show synergy with rifampicin in a whole cell assay (Fig. 6). The compound NSC600285 in particular is known to affect human RNase HI *in vitro* (bioassay ID: 5185715), but is not toxic to eukaryotic cells, unless the cells are deficient in DNA replication or repair, such as the Δ*rad50* and Δ*rad18* yeast strains (bioassay ID: 3254803). This supports an on-target effect in eukaryotic cells, but indicates that modifications are

necessary to develop these chemical scaffolds for higher affinity and specificity for *M. tuberculosis* RNase HI. However, the inhibition we observe indicates that these scaffolds can penetrate both eukaryotic and prokaryotic cell walls, a key consideration for the delivery of antibiotics to intracellular pathogens such as *M. tuberculosis* and *L. monocytogenes*. In combination, these lines of evidence strongly support an on-target effect on RNase HI *in vivo* for these compounds, but the generation of resistant mutants to formally prove on-target *in vivo* activity will require the development of compounds with higher affinity. Crystal structures of the compounds in complex with *M. tuberculosis* RNase HI will allow more extensive structure-activity relationships to be established, and we are actively pursuing this avenue. The use of rifampicin as a sensitizing compound will be advantageous in high-throughput screening, both for isolating compounds that might otherwise inhibit RNase HI too weakly to be identified in a normal growth inhibition screen, and for prioritizing hits for further investigation and development.

As rifampicin is a naturally occurring antibiotic, produced by the actinobacterium *Amycolatopsis rifamycinica* (83), it is not likely to be present at lethal concentrations in its native environment. The model of rifampicin action proposed in this paper suggests that the maximum cellular impact on the target organism is achieved by stochastically halting transcription at its earliest point on the gene. Coupled with the physical protection of the initiation region from subsequent initiation events, this produces a formidable and subtle weapon since it uses the target cell's transcriptional capacity to amplify the cytotoxic impact. Many further questions arise from this model. For example, does transcription arising from co-directional or divergent genes ameliorate or amplify this effect? Are terminal genes in operons disproportionally silenced? Does R-loop formation occur whenever there is a change in gene expression? Future work to answer these questions will help to further understand the intricacies of R-loop metabolism that this model implies.

In summary, this study validates RNase HI as a vulnerable, druggable target in the mycobacteria. It provides insight into R-loop metabolism in general, and specifically highlights the contribution that low-level transcriptional inhibition makes to R-loop formation. It also demonstrates the proof of principle that this is a novel cellular susceptibility, which can be utilised as an anti-bacterial strategy. Finally, it provides a starting

692 point for the development of new anti-TB therapeutics based on re-purposing and expanding
693 existing chemical scaffolds with favourable pharmaceutical properties.
694
695 **Acknowledgments and funding**:
696 We thank the Health Research Council of New Zealand (Grant 20/798), the Maurice Wilkins
697 Centre for Molecular Biodiscovery, the South African National Research Foundation, the
698 South African Medical Research Council, and the South African National Health Laboratory
699 Service Research Trust for financial support. We thank Prof Deborah Williamson for helpful
700 discussions.

References

1. O'Nell J. 2016. Tackling drug-resistant infections globally: final report and recommendations. Wellcome Trust 1-84.
2. WHO. 2020. Treatment of tuberculosis guideline
3. Conradie F, Diacon AH, Ngubane N, Howell P, Everitt D, Crook AM, Mendel CM, Egizi E, Moreira J, Timm J, McHugh TD, Wills GH, Bateson A, Hunt R, Van Niekerk C, Li M, Olugbosi M, Spigelman M, Nix TBTT. 2020. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. N Engl J Med 382:893-902.
4. Thomas M, White RL, Davis RW. 1976. Hybridization of RNA to Double-Stranded DNA: Formation of R-Loops. PNAS 73:2294-2298.
5. Drolet M, Phoenix P, Menzel R, Massé E, Liu LF, Crouch RJ. 1995. Overexpression of RNase H partially complements the growth defect of an *Escherichia coli* delta *topA* mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. Proc Natl Acad Sci U S A 92:3526-3530.
6. Alzu A, Bermejo R, Begnis M, Lucca C, Piccini D, Carotenuto W, Saponaro M, Brambati A, Cocito A, Foiani M, Liberi G. 2012. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell 151:835-846.
7. Santos-Pereira JM, Herrero AB, Garcia-Rubio ML, Marin A, Moreno S, Aguilera A. 2013. The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. Genes Dev 27:2445-58.
8. Dominguez-Sanchez MS, Barroso S, Gomez-Gonzalez B, Luna R, Aguilera A. 2011. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. PLoS Genet 7:e1002386.
9. Tuduri S, Crabbe L, Conti C, Tourriere H, Holtgreve-Grez H, Jauch A, Pantesco V, De Vos J, Thomas A, Theillet C, Pommier Y, Tazi J, Coquelle A, Pasero P. 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat Cell Biol 11:1315-24.
10. Dasgupta S, Masukata H, Tomizawa J. 1987. Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. Cell 51:1113-22.
11. Ivančić-Baće I, Al Howard J, Bolt EL. 2012. Tuning in to interference: R-loops and cascade complexes in CRISPR immunity. J Mol Biol 422:607-616.
12. Stein H, Hausen P. 1969. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. Science 166:393-5.
13. Balbontín R, Frazão N, Gordo I. 2020. DNA breaks-mediated cost reveals RNase HI as a new target for selectively eliminating antibiotic resistance. bioRxiv doi:10.1101/756767:756767.
14. Boyer PL, Smith SJ, Zhao XZ, Das K, Gruber K, Arnold E, Burke TR, Hughes SH. 2018. Developing and Evaluating Inhibitors against the RNase H Active Site of HIV-1 Reverse Transcriptase. J Virol 92:e02203-17.
15. Poongavanam V, Corona A, Steinmann C, Scipione L, Grandi N, Pandolfi F, Di Santo R, Costi R, Esposito F, Tramontano E, Kongsted J. 2018. Structure-guided approach identifies a novel class of HIV-1 ribonuclease H inhibitors: binding mode insights through magnesium complexation and site-directed mutagenesis studies. Medchemcomm 9:562-575.
16. Cao L, Song W, De Clercq E, Zhan P, Liu X. 2014. Recent progress in the research of small molecule HIV-1 RNase H inhibitors. Curr Med Chem 21:1956-67.
17. Watkins HA, Baker EN. 2010. Structural and functional characterization of an RNase HI domain from the bifunctional protein Rv2228c from *Mycobacterium tuberculosis*. J Bacteriol 192:2878-86.
18. Minias AE, Brzostek AM, Korycka- Machala M, Dziadek B, Minias P, Rajagopalan M, Madiraju M, Dziadek J. 2015. RNase HI Is Essential for Survival of *Mycobacterium smegmatis*. PLOS One 10:e0126260.
19. Gupta R, Chatterjee D, Glickman MS, Shuman S. 2017. Division of labor among Mycobacterium smegmatis RNase H enzymes: RNase H1 activity of RnhA or RnhC is

essential for growth whereas RnhB and RnhA guard against killing by hydrogen peroxide in stationary phase. Nucleic Acids Res 45:1-14.

20. Czubat B, Minias A, Brzostek A, Żaczek A, Struś K, Zakrzewska-Czerwińska J, Dziadek J. 2020. Functional Disassociation Between the Protein Domains of MSMEG_4305 of *Mycolicibacterium smegmatis* (*Mycobacterium smegmatis*) in vivo. Front Microbiol 11.
21. Jacewicz A, Shuman S. 2015. Biochemical Characterization of *Mycobacterium smegmatis* RnhC (MSMEG_4305), a Bifunctional Enzyme Composed of Autonomous N-Terminal Type I RNase H and C-Terminal Acid Phosphatase Domains. J Bacteriol 197:2489-98.
22. Lamichhane G, Zignol M, Blades NJ, Geiman DE, Dougherty A, Grosset J, Broman KW, Bishai WR. 2003. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 100:7213-8.
23. Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM. 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog 7:e1002251.
24. Parish T, Stoker NG. 2000. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. Microbiology 146 (Pt 8):1969-1975.
25. Andreu N, Zelmer A, Fletcher T, Elkington PT, Ward TH, Ripoll J, Parish T, Bancroft GJ, Schaible U, Robertson BD, Wiles S. 2010. Optimisation of bioluminescent reporters for use with mycobacteria. PLoS One 5:e10777.
26. Rauzier J, Moniz-Pereira J, Gicquel-Sanzey B. 1988. Complete nucleotide sequence of pAL5OOO, a plasmid from *Mycobacterium fortuitum*. Gene 71:315-321.
27. Machowski EE, Dawes S, Mizrahi V. 2005. TB tools to tell the tale-molecular genetic methods for mycobacterial research. Int J Biochem Cell Biol 37:54-68.
28. Dawes SS, Crouch RJ, Morris SL, Mizrahi V. 1995. Cloning, sequence analysis, overproduction in *Escherichia coli* and enzymatic characterization of the RNase HI from *Mycobacterium smegmatis*. Gene 165:71-5.
29. Downing KJ, McAdam RA, Mizrahi V. 1999. *Staphylococcus aureus* nuclease is a useful secretion reporter for mycobacteria. Gene 239:293-9.
30. Belisle JT, Mahaffey SB, Hill PJ. 2009. Isolation of mycobacterium species genomic DNA. Methods Mol Biol 465:1-12.
31. Miller J. 1972. Assay of B-galactosidase In: Experiments in molecular genetics. CSH Laboratory Press, Cold

Spring Harbor, New York.

32. Griffith KL, Wolf RE, Jr. 2002. Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. Biochem Biophys Res Commun 290:397-402.
33. Sambandamurthy VK, Derrick SC, Hsu T, Chen B, Larsen MH, Jalapathy KV, Chen M, Kim J, Porcelli SA, Chan J, Morris SL, Jacobs WR, Jr. 2006. *Mycobacterium tuberculosis* DeltaRD1 DeltapanCD: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. Vaccine 24:6309-20.
34. Walker IH, Hsieh PC, Riggs PD. 2010. Mutations in maltose-binding protein that alter affinity and solubility properties. Appl Microbiol Biotechnol 88:187-97.
35. Cordingley MG, Callahan PL, Sardana VV, Garsky VM, Colonno RJ. 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. J Biol Chem 265:9062-5.
36. Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189:113-30.
37. Studier FW. 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207-34.

38. Cordingley MG, Callahan PL, Sardana VV, Garsky VM, Colonno RJ. 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. J Biol Chem 265:9062-9065.
39. Parniak MA, Min K-L, Budihas SR, Le Grice SFJ, Beutler JA. 2003. A fluorescence-based high-throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity. Anal Biochem 322:33-39.
40. Belotserkovskii BP, Soo Shin JH, Hanawalt PC. 2017. Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter. Nucleic Acids Res 45:6589-6599.
41. Tous C, Aguilera A. 2007. Impairment of transcription elongation by R-loops in vitro. Biochem Biophys Res Commun 360:428-32.
42. Unniraman S, Nagaraja V. 1999. Regulation of DNA gyrase operon in *Mycobacterium smegmatis*: a distinct mechanism of relaxation stimulated transcription. Genes Cells 4:697-706.
43. Ahmed W, Menon S, D. N. B. Karthik PV, Nagaraja V. 2015. Autoregulation of topoisomerase I expression by supercoiling sensitive transcription. Nucleic Acids Research 44:1541-1552.
44. Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, Carrico RJ. 1986. Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. J Immunol Methods 89:123-30.
45. Hong X, Cadwell GW, Kogoma T. 1995. Escherichia coli RecG and RecA proteins in R-loop formation. EMBO J 14:2385-2392.
46. Ranes MG, Rauzier J, Lagranderie M, Gheorghiu M, Gicquel B. 1990. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a "mini" mycobacterium-*Escherichia coli* shuttle vector. J Bacteriol 172:2793.
47. Bourn WR, Jansen Y, Stutz H, Warren RM, Williamson AL, van Helden PD. 2007. Creation and characterisation of a high-copy-number version of the pAL5000 mycobacterial replicon. Tuberculosis (Edinb) 87:481-8.
48. Samul R, Leng F. 2007. Transcription-coupled hypernegative supercoiling of plasmid DNA by T7 RNA polymerase in *Escherichia coli* topoisomerase I-deficient strains. J Mol Biol 374:925-935.
49. Wolak C, Ma HJ, Soubry N, Sandler SJ, Reyes-Lamothe R, Keck JL. 2020. Interaction with single-stranded DNA-binding protein localizes ribonuclease HI to DNA replication forks and facilitates R-loop removal. Mol Microbiol 114:495-509.
50. McClure WR, Cech CL. 1978. On the mechanism of rifampicin inhibition of RNA synthesis. J Biol Chem 253:8949-8956.
51. Agrawal P, Miryala S, Varshney U. 2015. Use of *Mycobacterium smegmatis* deficient in ADP-ribosyltransferase as surrogate for *Mycobacterium tuberculosis* in drug testing and mutation analysis. PLoS One 10:e0122076.
52. Li X-Z, Zhang L, Nikaido H. 2004. Efflux pump-mediated intrinsic drug resistance in *Mycobacterium smegmatis*. Antimicrob Agents Chemother 48:2415-2423.
53. Gowrishankar J, Harinarayanan R. 2004. Why is transcription coupled to translation in bacteria? Mol Microbiol 54:598-603.
54. Masse E, Drolet M. 1999. R-loop-dependent hypernegative supercoiling in *Escherichia coli topA* mutants preferentially occurs at low temperatures and correlates with growth inhibition. J Mol Biol 294:321-32.
55. Yang Z, Hou Q, Cheng L, Xu W, Hong Y, Li S, Sun Q. 2017. RNase H1 Cooperates with DNA Gyrases to Restrict R-Loops and Maintain Genome Integrity in Arabidopsis Chloroplasts. Plant Cell 29:2478.
56. Winder FG, Collins PB. 1970. Inhibition by Isoniazid of Synthesis of Mycolic Acids in Mycobacterium tuberculosis. J Gen Microbiol 63:41-48.
57. Chen J, Raymond K. 2006. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. Ann Clin Microbiol Antimicrob 5:3-3.

58. Lang KS, Hall AN, Merrikh CN, Ragheb M, Tabakh H, Pollock AJ, Woodward JJ, Dreifus JE, Merrikh H. 2017. Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. Cell 170:787-799.e18.
59. Wahba L, Amon JD, Koshland D, Vuica-Ross M. 2011. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol Cell 44:978-88.
60. Amon JD, Koshland D. 2016. RNase H enables efficient repair of R-loop induced DNA damage. Elife 5.
61. Parajuli S, Teasley DC, Murali B, Jackson J, Vindigni A, Stewart SA. 2017. Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. J Biol Chem 292:15216-15224.
62. Rawal P, Kummarasetti VB, Ravindran J, Kumar N, Halder K, Sharma R, Mukerji M, Das SK, Chowdhury S. 2006. Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. Genome Res 16:644-55.
63. Gomez-Gonzalez B, Garcia-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marin A, Foiani M, Aguilera A. 2011. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. Embo j 30:3106-19.
64. Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL, Li X. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes Dev 25:2041-56.
65. O'Connell K, Jinks-Robertson S, Petes TD. 2015. Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity. Genetics 201:963-75.
66. Li X, Manley JL. 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122:365-78.
67. Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L, Yen A, Kamieniarz-Gdula K, Dimitrov S, Pathania S, McKinney KM, Eaton ML, Kellis M, Hill SJ, Parmigiani G, Proudfoot NJ, Livingston DM. 2015. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol Cell 57:636-647.
68. Sordet O, Redon CE, Guirouilh-Barbat J, Smith S, Solier S, Douarre C, Conti C, Nakamura AJ, Das BB, Nicolas E, Kohn KW, Bonner WM, Pommier Y. 2009. Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks. EMBO Rep 10:887-93.
69. Dutta D, Shatalin K, Epshtein V, Gottesman ME, Nudler E. 2011. Linking RNA polymerase backtracking to genome instability in *E. coli*. Cell 146:533-43.
70. Zatreanu D, Han Z, Mitter R, Tumini E, Williams H, Gregersen L, Dirac-Svejstrup AB, Roma S, Stewart A, Aguilera A, Svejstrup JQ. 2019. Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. Mol Cell 76:57-69.e9.
71. Zhang X, Chiang HC, Wang Y, Zhang C, Smith S, Zhao X, Nair SJ, Michalek J, Jatoi I, Lautner M, Oliver B, Wang H, Petit A, Soler T, Brunet J, Mateo F, Angel Pujana M, Poggi E, Chaldekas K, Isaacs C, Peshkin BN, Ochoa O, Chedin F, Theoharis C, Sun LZ, Curiel TJ, Elledge R, Jin VX, Hu Y, Li R. 2017. Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. Nat Commun 8:15908.
72. Wu HY, Shyy SH, Wang JC, Liu LF. 1988. Transcription generates positively and negatively supercoiled domains in the template. Cell 53:433-40.
73. Tsao Y-P, Wu H-Y, Liu LF. 1989. Transcription-driven supercoiling of DNA: Direct biochemical evidence from in vitro studies. Cell 56:111-118.
74. Ahmed W, Sala C, Hegde SR, Jha RK, Cole ST, Nagaraja V. 2017. Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase. PLoS Genet 13:e1006754.
75. Baquero F, Levin BR. 2021. Proximate and ultimate causes of the bactericidal action of antibiotics. Nat Rev Microbiol 19:123-132.
76. de Knegt GJ, Bruning O, ten Kate MT, de Jong M, van Belkum A, Endtz HP, Breit TM, Bakker-Woudenberg IA, de Steenwinkel JE. 2013. Rifampicin-induced transcriptome response in rifampicin-resistant *Mycobacterium tuberculosis*. Tuberculosis (Edinb) 93:96-101.

914 77. Hu Y, Mangan JA, Dhillon J, Sole KM, Mitchison DA, Butcher PD, Coates AR. 2000.
915 Detection of mRNA transcripts and active transcription in persistent *Mycobacterium*
916 *tuberculosis* induced by exposure to rifampin or pyrazinamide. J Bacteriol 182:6358-6365.
917 78. Stevenson-Jones F, Woodgate J, Castro-Roa D, Zenkin N. 2020. Ribosome reactivates
918 transcription by physically pushing RNA polymerase out of transcription arrest. Proc Natl
919 Acad Sci U S A 117:8462.
920 79. Cokol M, Kuru N, Bicak E, Larkins-Ford J, Aldridge BB. 2017. Efficient measurement and
921 factorization of high-order drug interactions in *Mycobacterium tuberculosis*. Sci Adv
922 3:e1701881.
923 80. Malik M, Zhao X, Drlica K. 2006. Lethal fragmentation of bacterial chromosomes mediated
924 by DNA gyrase and quinolones. Mol Microbiol 61:810-825.
925 81. Stolz R, Sulthana S, Hartono SR, Malig M, Benham CJ, Chedin F. 2019. Interplay between
926 DNA sequence and negative superhelicity drives R-loop structures. Proc Natl Acad Sci U S A
927 116:6260-6269.
928 82. Godbole AA, Ahmed W, Bhat RS, Bradley EK, Ekins S, Nagaraja V. 2014. Inhibition of
929 *Mycobacterium tuberculosis* topoisomerase I by m-AMSA, a eukaryotic type II
930 topoisomerase poison. Biochem Biophys Res Commun 446:916-920.
931 83. Bala S, Khanna R, Dadhwal M, Prabagaran SR, Shivaji S, Cullum J, Lal R. 2004.
932 Reclassification of *Amycolatopsis mediterranei* DSM 46095 as Amycolatopsis rifamycinica
933 sp. nov. Int J Syst Evol Microbiol 54:1145-1149.

934

935