R-loop induced stress response by second (p)ppGpp synthetase in Mycobacterium smegmatis: Functional and domain interdependence.

Bhatt, Apoorva

DOI:
10.1111/mmi.13453

Document Version
Peer reviewed version

Citation for published version (Harvard):
Bhatt, A 2016, 'R-loop induced stress response by second (p)ppGpp synthetase in Mycobacterium smegmatis: Functional and domain interdependence.', Molecular Microbiology. https://doi.org/10.1111/mmi.13453

R-loop induced stress response by second (p)ppGpp synthetase in *Mycobacterium smegmatis*: Functional and domain interdependence

Sushma Krishnan[1], Anushya Petchiappan[1], Albel Singh[2], Apoorva Bhatt[2], Dipankar Chatterji[1] *

Running title: Role of MS_RHII-RSD in R-loop induced stress response

Key words: Replication stress, RNase H, bifunctional protein, RNA polymerase, R-loop

Author information

1. Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India. 560012.

2. School of Biosciences, University of Birmingham, Edgbaston, Birmingham. B152TT.UK

*Correspondence: Dipankar Chatterji, Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India. 560012.

Phone: 080-22932836; E-mail: dipankar@mbu.iisc.ernet.in

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/mmi.13453

Summary

Persistent R-loops lead to replicative stress due to RNA polymerase stalling and DNA damage. RNase H enzymes facilitate the organisms to survive in the hostile condition by removing these R-loops. MS_RHII-RSD was previously identified to be the second (p)ppGpp synthetase in *Mycobacterium smegmatis*. The unique presence of an additional RNase HII domain raises an important question regarding the significance of this bifunctional protein. In this report, we demonstrate its ability to hydrolyze R-loops in *E. coli* exposed to UV stress. MS_RHII-RSD gene expression was upregulated under UV stress and this gene deleted strain showed increased R-loop accumulation as compared to the wild type. The domains in isolation are known to be inactive and the full length protein is required for its function. Domain inter-dependence studies using active site mutants reveal the necessity of a hexamer form with high alpha helical content. In previous studies, bacterial RNase type HI has been mainly implicated in R-loop hydrolysis, but in this study the RNase HII domain containing protein showed the activity. The prospective of this differential RNase HII activity is discussed. This is the first report to implicate a (p)ppGpp synthetase protein in R-loop induced stress response.

Introduction

Bacteria encounter a variety of nutritional and environmental stresses, they invest their energy cautiously to survive in hostile condition and one of the responses is termed as stringent response. The stringent response is signaled by the synthesis of hyperphosphorylated guanosine nucleotides, guanosine penta phosphate (pppGpp) and guanosine tetra phosphate (ppGpp), collectively named as (p)ppGpp (Cashel & Gallant, 1969, Avarbock *et al.*, 2000, Chatterji & Ojha, 2001). These alarmones exhibit pleiotropic effect on many cellular processes such as

transcription, translation, replication, virulence, differentiation and persistence via diverse pathways (Srivatsan & Wang, 2008).

In Gram negative bacteria, (p)ppGpp is synthesized by monofunctional RelA and hydrolyzed by bifunctional SpoT, which also has weak (p)ppGpp synthetic activity (Xiao *et al.*, 1991). On the contrary, in Gram positive bacteria, (p)ppGpp is synthesized and hydrolyzed by a single bifunctional enzyme Rel (Avarbock *et al.*, 1999, Avarbock *et al.*, 2000, Hogg *et al.*, 2004). Apart from these classical long, multidomain RelA-SpoT (RSHs), some short RSH homolog have been identified in prokaryotes and eukaryotes. They are mostly single domain, monofunctional proteins either with Short Alarmone Synthetase (SAS) or Short Alarmone Hydrolase (SAH) activity (Sun *et al.*, 2010, Atkinson *et al.*, 2011). SAS enzymes have different functions in microorganisms like *Bacillus subtilis, Streptococcus mutans, etc.* (Hauryliuk *et al.*, 2015). Weiss & Stallings, (2013) observed that Rv 1366, a SAS in *M. tuberculosis,* does not affect its physiology and virulence. The role played by SAS proteins in mycobacteria remains a mystery.

An elevated level of (p)ppGpp in *M. smegmatis* under nutrient deprived conditions was first reported by Ojha et al (Ojha *et al.*, 2000). In *M. smegmatis,* (p)ppGpp synthesis is regulated by $rel_{Msm,}$ comprising of (p)ppGpp synthesis (RSD) and hydrolysis (HD) domains (Jain *et al.*, 2006). The Δrel_{Msm} of *M. smegmatis* was expected to have a (p)ppGpp null phenotype. However, we showed previously that the strain was still capable of synthesizing (p)ppGpp *in vivo,* and the second (p)ppGpp synthetase was named as MS_RHII-RSD (Murdeshwar & Chatterji, 2012). MSMEG_5849 codes for the bifunctional protein MS_RHII-RSD, which has a C-terminal RSD domain similar to other SAS but is different from them due to the presence of a N-terminal RNase HII (Ribonuclease HII) domain. We previously showed that it is capable of (p)ppGpp

synthesis as well as degrading RNA-DNA hybrids (Murdeshwar and Chatterji, 2012). No other such protein, which couples RNase H activity with stringent response, is reported in literature. Hence, this work was motivated to address the question - why is there a RNase HII domain coupled with (p)ppGpp synthetic machinery?

RNase Hs are endogenous enzymes which cleave the RNA moiety from RNA:DNA hybrids (Stein & Hausen, 1969). These widespread enzymes have profound role in many basic functions including replication, transcription, gene expression and DNA repair (Kogoma & Foster, 1983, Drolet, 2006). They are classified into two major groups, Type 1 and Type 2, on the basis of amino acid sequence conservation. Type 1 comprises of prokaryotic RNase HI, eukaryotic RNase HI and RNase H domain of reverse transcriptase. Type 2 contains prokaryotic RNase HII, eukaryotic H2 and prokaryotic RNase HIII (Ohtani *et al.*, 1999).

RNase H enzymes are involved in the removal of specific structures called R-loops. R-loops have triple stranded nucleic acid structure and are formed when the nascent RNA strand hybridizes with the complementary template strand leading to the displacement of the non-template DNA strand (Gowrishankar *et al.*, 2013). Existence of R-loops *in vivo* was first demonstrated by Crouch and colleagues (Drolet *et al.*, 1995). In bacteria, the replisome complex moves faster than the transcription complex along the same DNA template and they can collide with each other. These replication-transcription conflicts lead to the stalling of RNA polymerase and hence arrest the replication fork movement (Dutta *et al.*, 2011, Merrikh *et al.*, 2011, Alzu *et al.*, 2012). As a consequence, stable R-loops are formed behind the stalled RNA polymerase. Alternatively, the negative supercoiling of DNA leads to co-transcriptional R-loop formation in bacteria and hence replicative stress (Drolet, 2006, Poveda *et al.*, 2010, Stirling *et al.*, 2012). R-loops are considered as beneficial in regulating gene expression by governing transcription. They

are also deleterious to genome integrity as they impede replication fork progression and may cause DNA damage (Skourti-Stathaki & Proudfoot, 2014, Gan *et al.*, 2011). Therefore, the organisms need to maintain a physiological R-loop balance. RNA:DNA hybrids are stable and their removal is an energy - consuming mechanism (Roberts & Crothers, 1992). Hence, organisms possess various factors to remove persistent R-loops and one of the best characterized are RNase H enzymes (Aguilera & Garcia-Muse, 2012).

(p)ppGpp is known to modulate RNA polymerase by binding next to the active site to destabilize open transcription complexes in gram-negative bacteria (Cashel *et al.*, 1996, Paul *et al.*, 2004, Ross *et al*., 2013). Trautinger *et al.* showed that ppGpp decreases accumulation of stalled transcription complexes and thereby it resolves replication-transcription conflicts (Trautinger *et al.*, 2005). *E. coli* RNAP crystal structure studies, cross-linking and mass spectrometric analysis revealed multiple (p)ppGpp binding sites (Zuo *et al*. 2013, Syal and Chatterji, 2015, Rose *et al*., 2016). However, there is no evidence for the binding of ppGpp to RNAP in gram positive bacteria. *B. subtilis* RNAP has no ppGpp binding motifs but it uses an indirect mechanism by altering the GTP homeostasis to mount stringent response. GTP is one of the initiating nucleotides in *B. subtilis* and increase in (p)ppGpp synthesis reduces GTP pool thereby modulating the rRNA promoter activity (Krasny and Gourse 2004, Hauryliuk *et al*., 2015). Regulation of promoters by (p)ppGpp has been shown in Mycobacteria as well though the mechanism remains unknown (Tare *et al.,* 2013).

One may speculate that since R-loops are important in replication-transcription conflicts leading to replication stress, this stress can be efficiently managed by two mechanisms; R-loop removal by RNase H and destabilization of stalled RNA polymerase by (p)ppGpp synthesis.

MS_RHII-RSD possesses both these important activities (RNase HII and (p)ppGpp synthetase) in a single polypeptide. Therefore, we believe that the protein may be involved in withstanding replication stress by removing R-loops in *M. smegmatis*. As this protein also synthesizes (p)ppGpp, we speculate that it destabilizes the hindered RNA polymerase via an indirect mechanism. Such proposition gains further value due to the long term survival of mycobacteria under stringent conditions.

In this study, we have characterized MS_RHII-RSD for its function by various approaches. The results showed that MS_RHII-RSD efficiently hydrolyses preformed R-loops. Under UV stress, *ms_rhII-rsd* expression was found to be significantly upregulated and R-loop accumulation was high in the Δ*ms_rhII-rsd*. The expression of the RNase HII and (p)ppGpp synthetase domains *in tandem* on the same polypeptide chain is necessary for both activities. The domains in isolation are not functional *in vitro* or *in trans* complementation experiments (Murdeshwar & Chatterji, 2012). Therefore, the interdependence of the RHII and RSD domains was explored and the results showed that the hexameric form and high alpha helicity of the full length protein are essential for its function.

Results

MS_RHII-RSD hydrolyzes preformed R-loops

To determine if MS_RHII-RSD has R-loop hydrolysis activity, initially R-loop was synthesized *in vitro* based on a previous observation that they are generated when a G- rich RNA strand is transcribed from a C-rich template DNA strand (Yu *et al.*, 2003, Skourti-Stathaki & Proudfoot, 2014). Also, Toriumi *et al*. showed that an AGGAG repeat is adequate for R-loop formation (Toriumi *et al.*, 2013). Landgraf *et al*. reported that RNA as short as 50 nucleotide can form R-loops in 70% formamide (Landgraf *et al.*, 1995). In this study, a simple miniature model

of traditional *in vitro* R-loop formation was used (Fig. 1). Two single stranded DNA oligonucleotides (named as D1 and D2) containing 80 nucleotides each were custom synthesized such that they were complementary to each other except for a 30 nucleotide region in the middle. The non-complementary region of D1 (template strand) was designed with TCCTC repeats. These two DNA oligonucleotides D1 and D2 were annealed to form double stranded DNA (D1-D2) (Fig. 2A). Then, RNA oligo (R3) of 20 nucleotide length containing AGGAG repeats was added to the annealed ds DNA under suitable conditions to form R-loop *in vitro* as described previously (Landgraf *et al.*, 1995). The R-loop formation was favored due to the complementarity of R3 with D1 and the non-complementary region in the D1-D2 hybrid. Electrophoretic mobility shift is a convenient way of visualizing R-loop formation (Daniels & Lieber, 1995, Yu *et al.*, 2003). Here, R-loop formation was confirmed by a mobility shift of 20 bases from the 80 bp ds DNA in a 2% agarose gel (Fig. 2A) and also by immunoblotting with anti RNA:DNA hybrid antibody (S9.6) (Fig. 2B).

The R-loop removal activity of MS_RHII-RSD was assayed by incubating the preformed R-loop and the purified recombinant protein in a reaction mixture at 37°C for 1 hour. In agarose gel electrophoresis, a band was observed with a lower mobility shift of 20 bases compared to the R-loop substrate. This band corresponds to ds DNA product obtained after R-loop hydrolysis indicating that MS_RHII-RSD protein efficiently hydrolyzes R-loops *in vitro.* In the positive control reaction with *E. coli* RNase HI, some residual unhydrolyzed R-loops could be seen along with the hydrolyzed ds DNA product but not in reactions with MS_RHII-RSD (Fig. 2A). RNase HII inactivated mutant described in later sections was used as a negative control (last lane of Fig. 2A) to rule out the possibility of nuclease contamination during the protein preparation**.**

MS_RHII-RSD removes R-loops in RNase H-deficient E. coli

A possible role for MS_RHII-RSD could be to rescue *M. smegmatis* from replication stress caused due to R-loop formation. In an earlier study, UV irradiation was used to induce replication stress by double stranded break formation in DNA (Yajima *et al.*, 2009). In yet another study, an enhanced R-loop formation was shown at UV-C damage sites when RNase H was silenced in human dermal fibroblast culture (Tresini *et al.*, 2015). We used UV irradiation to induce R-loop formation in *E. coli* GJ12055 strain which is deficient in conventional RNase HI and RNase HII enzymes. This is a temperature-sensitive strain (*ts*), which can grow well at 30°C but not at 42°C unless it is provided a functional *rnhA* gene from RNase H-expressing plasmids such as pSK 760 (Kanaya & Crouch, 1983). Normalized mid- log phase cultures of *E. coli* GJ12055 parent strain as well as pSK760 and MS_RHII-RSD complemented strains were UV-irradiated for 30 minutes and the optical density at 600 nm (OD_{600}) was measured immediately after UV exposure and every hour thereafter till 10 hours. A growth defect was observed in all three strains till 2 hours post UV exposure. Subsequently, the complemented strains slowly started recovering whereas *E. coli* GJ12055 parent strain did not recover even after 6 hours (Fig. 3A). Genomic DNA samples isolated from the cells at three different time points (0, 1 and 4 hours) after UV treatment were subjected to immunoblotting using the anti RNA:DNA hybrid specific S9.6 primary antibody (Fig. 3B). R-loop intensity was found to be high in the DNA from cells harvested at 0^{th} and 1^{st} hour after UV exposure. The growth retardation combined with increased R-loop detection in the initial hours post UV treatment indicated that the cells were under UV-induced replication stress and accumulated R-loops *in vivo*. The decrease in the R-loop level, 4 hours after UV exposure and the growth recovery of the pSK 760 and MS_RHII-

RSD complemented strains suggested that one of the reasons for the rescue of growth defect may be the hydrolysis of R-loops by RnhA and MS_RHII-RSD.

The ability of MS_RHII-RSD to hydrolyze R-loops was tested both at 30°C (permissive temperature) and 42°C (non-permissive temperature). However, we did not observe any growth defect at 30°C after UV stress (Fig. S1) and also no signal for R-loops could be detected by immunoblotting except the positive control. This could be due to the lower sensitivity of the strain to UV stress at 30°C. This statement is supported by an earlier study in which the *ts E. coli* with mutated single-stranded DNA binding protein showed higher UV sensitivity at non-permissive temperature (Lieberman and Witkin, 1981). RNase H mutants are known to be thermosensitive (Itaya *et al*., 1999) with increased R-loop formation attributed as one of the reasons.

MS_RHII-RSD expression is elevated in M. smegmatis under UV stress likely to remove R-loops

Apart from MS_RHII-RSD, *M. smegmatis* genome contains three other RNase H enzymes, *rnhA* gene encoding RNase HI (Dawes *et al.*, 1995), *rnhB* encoding RNase HII (Minias *et al.*, 2015b), and *rnhC* encoding RNase HI-acid phosphatase fusion protein (Jacewicz & Shuman, 2015). *M. smegmatis* rnhA, rnhB, and rnhC are enzymatically active and cleave the RNA strand in RNA:DNA hybrids. Though there are three other active endogenous RNase H enzymes, the special requirement for the fourth RNase H in *M. smegmatis* was examined using quantitative real time-reverse transcriptase PCR (qRT-PCR) and also by comparing the UV-induced R-loop accumulation in wild type and Δ*ms_rhII-rsd* strains. Mid-log phase cultures of *M. smegmatis* wild type and Δ*ms_rhII-rsd* were UV-irradiated for 30 minutes and OD_{600} was

measured immediately and 1 hour after UV exposure, then at 4 hours intervals till 16 hours. The OD of the mid-log phase cultures of strains not exposed to UV was also measured as a control. A growth defect was observed in the UV-treated cultures till 4 hours post UV exposure after which they started recovering (Fig. S2). The expression levels of all four RNase H genes, encoding MSMEG_5562, MSMEG_2442, MSMEG_4305 and MSMEG_5849, 4 hours after UV exposure were analyzed using qRT-PCR. Specific set of primers used in this experiment are listed in Table. 1. It was observed that the expression levels of MSMEG_5562 (*rnhA)* and MSMEG_2442 *(rnhB)* were not affected much and MSMEG_4305 was downregulated significantly. However, MS_RHII-RSD (MSMEG_5849) expression was found to be significantly upregulated (Fig. 4A).

UV-induced R-loop accumulation in wild type and Δ*ms_rhII-rsd* strain was compared using immunoblotting with anti RNA:DNA antibody (S9.6). Genomic DNA isolated from UV-exposed and untreated mid log phase cultures at three different time points (0, 1 and 4 hours) were used for this blot. R-loop intensity was found to be higher in the DNA samples of Δ*ms_rhII-rsd* harvested at 1st and 4th hour after UV exposure when compared to that of wild type (Fig. 4B). The significant upregulation of MS_RHII-RSD and the accumulation of R-loop in Δ*ms_rhII-rsd* strain under UV stress indicated that MS_RHII-RSD plays a major role in UV-induced replication stress.

Inactivation of RHII domain does not affect RSD domain activity in MS_RHII-RSD

Since the domains in separation are not functional (Murdeshwar & Chatterji, 2012), we designed experiments to analyse the level of their interdependence. For this purpose, we inactivated one domain and studied its effect on the other domain. Multiple sequence alignment of the N-terminal domain of MS_RHII-RSD, initially annotated as DUF 429 shows that it is homologous to RNase HII enzymes from other bacteria (Murdeshwar & Chatterji, 2012). It has

been shown earlier that D, E, D, D are the four acidic conserved residues necessary for RNase H function (Tadokoro & Kanaya, 2009) and the RHII domain of the MS_RHII-RSD has the corresponding conserved D, D, D, E (D11, D39, D60 and E86) amino acids which were mutated individually to alanine using site-directed mutagenesis. However, none of these four point mutations led to the complete inactivation of RNase H activity (Data not shown). Hence, towards obtaining a completely inactive RHII domain in the full length protein, all the four residues D11, D39, D60 and E86 were sequentially mutated to alanine in the same construct of pET21b-MS_RHII-RSD. The full length protein containing all these four mutated residues was named as RHII-mutant. The RHII-mutant was overexpressed in *E. coli* BL21(DE3) and purified to homogeneity using nickel affinity chromatography (Fig. S3).

The purified RHII-mutant was characterized for its activity *in vitro* as well as *in vivo* as described previously (Murdeshwar & Chatterji, 2012) . An *in vitro* fluorescence assay (Watkins & Baker, 2010) was performed in which a 12-mer of 5′-Fluorescein-labeled-poly ribooligonucleotide (rA) annealed to 3′-dabsyl-labeled-poly deoxyribooligonucleotide (dT) was used as a substrate. In the annealed hybrid (AH), the dabsyl moiety quenches the fluorescence. The increase in fluorescence due to hydrolysis of the RNA:DNA hybrid by RNase H was monitored. The increase in fluorescence intensity as compared to that of AH indicates the successful hydrolysis of the RNA:DNA hybrid by the wild type MS_RHII-RSD protein. No increase in the fluorescence intensity was observed with RHII-mutant (Fig. 5A) revealing the complete loss of RNase H activity.

In vivo RNase H activity of the RHII domain mutant was characterized using *E .coli* GJ12055 based complementation assay (Itaya & Kondo, 1991). It was observed that at 42°C, *E. coli* GJ12055 cultures complemented with RHII-mutant construct did not survive but those

complemented with wild type MS_RHII-RSD construct and plasmid pSK760 containing functional *rnhA* gene (Kanaya & Crouch, 1983) grew well (Fig. 5B). This proved that the above mutations inactivated the RHII domain function of the full length MS_RHII-RSD protein leading to the loss in RNase H activity.

To analyse the mutation effect of RHII domain on RSD activity in full length MS_RHII-RSD, (p)ppGpp synthesis was assayed. Purified MS_RHII-RSD and RHII-mutant were incubated with GTP in the presence of [Υ^{32}P]-ATP. It was observed from thin layer chromatography (TLC) analysis that the RHII-mutant has a (p)ppGpp synthesis activity similar to that of the wild type protein (Fig. 5C). The RNase H activity of the RHII-mutant, on the other hand was completely lost.

RSD domain inactivation does not affect RHII domain function

Similar to other bacterial Rel and SAS proteins, the RSD domain of MS_RHII-RSD has an acidic chord of conserved DDRD motif in the active site (Sajish *et al.*, 2009). Based on this, 4 different mutant constructs were made in which D356, D357, R358 and D359 were mutated individually to alanine in full length MS_RHII-RSD and termed as RSD-mutants. The mutated proteins were purified to homogeneity using nickel affinity chromatography (Fig. S3). It was found that even single site mutations in the DDRD region affected its (p)ppGpp synthesis activity (Fig. 6A). These RSD-mutants were further characterized for RNase H activity. However, the RNase H activity of RSD-mutants was not affected *in vitro* (Fig. 6B) and *in vivo* (Fig. 6C).

Therefore, one may conclude that the functional inactivation of one domain does not affect the activity of the other domain. However the domains are nonfunctional when separated and expressed independently. This led us to think that there could be a certain degree of

structural interdependence between the RHII domain and RSD domain. Such a proposition was explored further based on oligomerization studies with full length protein and its domain mutant variants.

Hexameric form of MS_RHII-RSD is essential for its function

In our previous study, it was shown that the N-terminal RHII domain is monomeric and C-terminal RSD domain is tetrameric in nature. The full length active protein is, however, hexamer in solution (Murdeshwar & Chatterji, 2012). We checked the oligomeric status of all mutant variants using gel filtration chromatography followed by native PAGE. The wild type protein and its mutants were purified by nickel affinity followed by gel filtration chromatography. The analysis showed that MS_RHII-RSD and all the mutant variants remain in hexameric form in solution. The gel filtration profile for MS_RHII-RSD wild type, RHII-mutant and one of the RSD- mutants, D356A is shown (Fig. S4A). This is further confirmed by native-PAGE analysis of the peak elution fractions of the proteins after gel filtration (Fig. S4B). Native-PAGE marker, Ferritin (440 kDa, pI 4.8) and bovine serum albumin (BSA, 66 kDa, pI 5.2) were used as the corresponding molecular mass standards. The band corresponding to MS_RHII-RSD and mutant proteins was observed between 440 kDa (Ferritin) and 242 kDa (marker) and this in accordance with the mass estimated from gel filtration (Fig. S4A) indicated the presence of a hexamer form for all proteins.

Since all the domain mutant variants retained one activity in hexameric form, a question comes to mind whether the oligomeric form is necessary for them to be functional. Conversion of hexamer to lower oligomeric form can be attempted after identifying the mode of interaction in the hexamer. To understand the interaction type, initially, native hexameric MS_RHII-RSD was treated with increasing concentrations of Sodium Dodecyl Sulphate (SDS) and Beta –

Mercaptoethanol (β-ME) separately, and analyzed using native PAGE for the disruption of hexameric form. SDS concentration even as low as 0.05 % showed the disruption of hexamer into lower oligomers (Fig. S4C). It was observed from native PAGE analysis that with 0.2% SDS treatment for 30 minutes, the hexamer was completely converted to dimer and monomer forms. However, β-ME-treated samples remained as hexamer, even at high β-ME concentrations (Fig. S4D). These results indicated that the hexamer is formed by non-covalent interactions as it is well known that β-ME treatment breaks the covalent interactions by reducing the disulfide bonds. Iodoacetamide labeling and mass spectral analysis further confirmed that both the cysteine residues in MS_RHII-RSD are free. (Table. S1).

The RNase H and (p)ppGpp synthesis activity of the lower oligomeric populations of full length MS_RHII-RSD were analyzed by *in vitro* fluorescence assay and (p)ppGpp synthetase assay. Expectedly, 0.2 % SDS treated MS_RHII-RSD was inactive in both the assays (Fig. 7). *E. coli* RNase HI (Cerritelli & Crouch, 2009) and N-terminal variant of Rel_{Msm} (Rel-NTD) (Jain *et al.*, 2006) are known to be active as monomers and served as the positive control for RNase H assay and (p)ppGpp synthesis assay, respectively. Both retained activity in the presence of 0.2 % SDS. To rule out the possibility of protein unfolding by SDS, we performed a CD analysis with MS_RHII-RSD treated with 0.2 % SDS and it showed a clear alpha helical pattern (Fig. S5). (p)ppGpp synthesis assay was also carried out with 0.05 % SDS treated MS-RHII-RSD which is the minimum SDS concentration tested for inhibiting oligomerization (Fig. S4C). Even at this SDS concentration, MS-RHII-RSD lost its (p)ppGpp synthetic activity (Fig.S6). According to the previous reports, the unfolding and aggregation of the protein generally starts at concentration near or above critical micellar concentration (CMC) and lower SDS concentration promotes helicity in several proteins such as BSA, HSA, Ubiquitin etc., (Moriyama & Takeda,

1999; Xu & Keiderling, 2004; Singh et al., 2008; Shaw et al., 2012; Qadeer et al., 2014). We have calculated the CMC for our reaction condition which is 0.07 % based on the thermodynamics of SDS reported by Makowska et al., (2015). Since 0.05% SDS is below the critical micellar concentration (CMC) in the reaction condition used in the assay, and the control proteins, *E. coli* RNase HII and Rel – NTD variant showed activity at 0.2 % and 0.05 % SDS, it is clear that the protein activity is not affected due to SDS. These results indicated that the hexamer form of the full length protein is essential for RNase HII as well as (p)ppGpp synthesis activity. In contrast to the individual domains in separation, the mutant variants retain the hexamer form and therefore the activity of non-mutated domain is preserved.

High alpha helical nature of full length MS_RHII-RSD favors its biological activity

Secondary structure of the MS_RHII-RSD, RHII-mutant and RSD-mutants was further analyzed by circular dichroism (CD) spectroscopy. These proteins were found to be alpha helical in nature as calculated by K2D2 server (Table. S2). The individual RHII and RSD domains, on the other hand, had reduced alpha helicity (Fig. S7).

Discussion

The discovery of MS_RHII-RSD in *M. smegmatis* raises an important question as to the requirement of a second (p)ppGpp synthetase in the microorganism, especially due to the unique presence of an RNase HII domain. In our previous study, we had hypothesized a possible role for the protein in R-loop removal. Here, we provide evidence for the involvement of MS_RHII-RSD in R-loop induced stress response in *M. smegmatis*. Taking hints from previous studies, we synthesized a miniature R-loop *in vitro* and observed that MS_RHII-RSD is indeed capable of hydrolyzing R-loops. By using a RNase H-deficient *ts E. coli* strain, we were able to observe that MS_RHII-RSD removes the R-loops formed after UV stress. Our argument is further

strengthened by the observation that Δ*ms_rhII-rsd* accumulates R-loops to a much larger extent after UV stress than the wild type.

Moreover, in qRT-PCR analysis, a significant upregulation of *ms_rhII-rsd* gene was observed when *M. smegmatis* wild type was exposed to UV for 30 minutes. However, this contradicts our earlier preliminary observation in which the western blot analysis showed a constitutive expression of this protein when the organism was exposed to UV for 20 minutes (Murdeshwar & Chatterji, 2012). This discrepancy is perhaps due to the increased UV exposure (30 min) in the present experiment and hence the increase in gene expression. Therefore, we performed a western blot analysis to check the protein expression level in the same UV stress condition followed in qRT-PCR analysis (i.e., 4 hours after 30 minutes UV exposure). The densitometry of the western blot band intensity showed 33 % increase in the protein level in the strains with UV stress than without UV stress (Fig. S8). Hence both the gene and protein expression levels are raised when the strain was exposed to UV for 30 min.

R-loops can be formed due to several reasons and have a major role in replication – transcription conflicts. These conflicts lead to stalled arrays of RNA polymerase and are detrimental to replication fork movement. Furthermore, the exposed single-stranded DNA is prone to mutagenesis. Due to deleterious effects on genome integrity, it becomes crucial for any organism to remove R-loops. (p)ppGpp mediates stringent response, plays a key role in replication arrest, transcription regulation and DNA supercoiling. It is known to resolve replication – transcription conflicts by destabilizing the stalled RNA polymerase. The presence of the RHII and RSD domains together in MS_RHII-RSD hints at a possible link between R-loop removal and stringent response in *M. smegmatis.*

Conflicts between replication and transcription arise more often in highly transcribed genes such as those of rRNA (Merrikh *et al.*, 2011). (p)ppGpp is a known regulator of rRNA genes expression and therefore the presence of a R-loop removal activity in the (p)ppGpp synthetase may be beneficial in such cases. Inhibition of translation due to nutrient starvation and antibiotic stress leads to increased RNA polymerase backtracking and therefore R-loop formation (Dutta *et al.*, 2011). Since stringent response is involved in both stress conditions, it is tempting to speculate the involvement of MS_RHII-RSD in both circumstances.

As per our knowledge, this is the first report in which a prokaryotic RNase HII domain containing bifunctional protein is shown to hydrolyze R-loops. So far Type 1 RNase H has been more commonly implicated than Type 2 RNase H to remove R-loops. There are many previous reports which discuss the role of RNase HI in R-loop removal. To list a few, Drolet et al. reported that over production of RNase HI using pSK 760 plasmid in Topoisomerase I deficient mutant partially complements the R-loop linked serious growth defect (Drolet *et al.*, 1995). Dutta et al. showed that R-loops are implicated in RNA polymerase backtracking and it is reversed upon RNase HI overexpression (Dutta *et al.*, 2011). However, there are a few contradicting reports where the only RNase HII/HIII can substitute the function of RNase HI (Zhang *et al.*, 1997, Itaya *et al.*, 1999, Ohtani *et al.*, 1999). Based on the defined substrate specificity of the RNase HII enzymes, they can only cleave single misincorporated ribonucleotide in ds DNA and remove Okazaki fragments from the lagging strand DNA (Murante *et al.*, 1998), whereas, RNases HI require minimum four consecutive ribonucleotide for substrate recognition and hence cannot remove single ribonucleotide. However, Eder et al. showed that human RNase H1 is able to act on substrates containing a single RNA residue, which is a proof for the differential substrate preference of the RNase H enzymes from different sources (Eder *et al.*, 1993).

R-loops are formed naturally during the initiation of DNA replication in mitochondrial DNA, bacterial plasmids, the bacteriophages ColE1 and T4, and in immunoglobulin (Ig) class-switch recombination (CSR) (reviewed by Santos-Pereira & Aguilera, 2015). Hence, it can be seen from extensive literature that even if the organism lacks RNase HI, the R-loops are expected to be present, where the other RNase HII and HIII tend to take care of the R-loop related defects. Prokaryotic RNase HII proteins are homolog of the eukaryotic RNase H2 catalytic subunit, which possesses two accessory domains. Since most RNase H2 proteins are implicated in R-loop removal (Chon *et al.*, 2013), there are possibilities for bacterial RNase HII also to recognize R-loop. MS_RHII-RSD is not a classical RNase HII but has an additional C-terminal RSD domain, which may help in R-loop recognition similar to the requirement of two accessory domains for eukaryotic RNase H2 activity.

Multiple types of RNase Hs may exist in a single organism to protect it from various stresses leading to a lethal mutation in RNase H gene (Ohtani *et al.*, 1999). In *M. smegmatis*, there are four functional RNase Hs. There are two RNase HI and two RNase HII type of proteins. No other organism with four active RNase Hs has been reported so far. The presence of three functional RNase Hs in *Streptomyces coelicolor is*, however, known (Ohtani *et al.*, 2005). *M. tuberculosis* has RNase H domain and (p)ppGpp synthetase domain homologs. It contains no classical *rnhA* gene. However, *rnhB* gene annotated as Rv2902c and another gene encoding Rv2228c as a bifunctional protein with N-terminal RNase HI domain and C-terminal α-ribazole phosphatase (CobC) domain have been identified (Watkins & Baker, 2010). Apart from these two RNase Hs, a homologue of RHII domain of MS_RHII-RSD, MT0800 has been found in *M. tuberculosis* but it remains to be characterized (Minias *et al.*, 2015a).

In our study, we observed that both RHII and RSD are structurally inter-dependent for their functions. In a previous study, it was observed that loss of the C-terminal domain of Rv2228c significantly reduces its RNase H activity due to the presence of a very short linker region of 10 amino acids between both domains (Watkins & Baker, 2010). Similarly, reason for this domain dependence of MS_RHII-RSD appears to be due to the presence of short linker region of 16 residues. On the other hand, the bifunctional protein, SCO2299 from *S. coelicolor* has more than 120 linker residues and due to this the N-terminal RNase H and C-terminal acid phosphatase activities are independent of each other. Finally, this remarkable presence of R-loop removing RNase HII and (p)ppGpp synthetase activities in a single protein could be an energy conserving mechanism in *M. smegmatis* to endure replication stress.

Our study provides novel insights about the significance of the second (p)ppGpp synthetase of *M. smegmatis* in R-loop induced stress response. Based on the obtained results, we propose a model to explain the function of MS_RHII-RSD (Fig. 8). Upon UV stress, the levels of MS_RHII-RSD increase within the cell. Any R-loops formed are removed by the RNase H domain and (p)ppGpp helps to destabilize the RNAP via an unknown mechanism. Recently, Kamarthapu et al., have shown direct evidence for role of (p)ppGpp in DNA repair in *E. coli* and we are excited to speculate the same in Mycobacteria although evidence for (p)ppGpp binding to RNAP is lacking (Kamarthapu et al., 2016). The Δ*ms_rhII-rsd* strain is yet to be completely characterized and further experiments may lead to more novel observations about this unique protein.

Experimental Procedures

Strains, growth conditions and media

A list of strains and plasmids used in this study is mentioned in Table S3. *E. coli* strains DH5α and BL21 (DE3) were grown in Luria-Bertani (LB) media (HiMedia Laboratories) at 37°C at 180 rpm. MS_RHII-RSD and Rel_{Msm}-NTD were overexpressed using pET-21b plasmids. For complementation studies, an *E. coli* GJ 12055 was used. *E. coli* GJ 12055 and complemented strains were grown at 30°C or 42°C in LB media containing 50 µg ml^{-1} kanamycin (Sigma). All *E. coli* strains harboring pET-21b or pSK760 plasmids were grown in the presence of 100 µg ml^{-1} ampicillin (Sigma). All strains were grown on plates containing LB media and 1.5% (w/v) agar (HiMedia Laboratories) in the presence of required antibiotics. *M. smegmatis* wild type mc2155 (WT) and Δ*ms_rhII-rsd* strains were grown in Middlebrook 7H9 broth (MB7H9, Difco) containing 2% (w/v) glucose (HiMedia Laboratories) and 0.05 % (v/v) Tween-80 (Sigma), at 37°C with agitation at 180 rpm. MB7H9 plates contained 1.5% agar and lacked Tween-80. Apramycin (Sigma) was used at a concentration of 50 µg ml^{-1} for Δ*ms_rhII-rsd.*

Site-directed mutagenesis

Site-directed mutagenesis was performed using pET-MS_RHII-RSD as the template to generate single amino acid mutants of the MS_RHII-RSD protein. Non-overlapping primers incorporating the desired mutations were designed to amplify the sequence (listed in Table S4). The products were phosphorylated using T4 polynucleotide kinase (Thermo Scientific), ligated using T4 DNA ligase (NEB) and transformed into *E. coli* DH5α. The incorporation of the mutation was confirmed by sequencing the plasmids. For generation of the RHII-mutant containing four active site mutations, sequential mutagenesis was performed where each mutation was individually incorporated using the mutant plasmid generated in the previous round as a template.

Protein expression and purification

Rel-NTD, full length MS_RHII-RSD, RNaseHII (RHII) and (p)ppGpp synthesis domain variant (RSD) were cloned in pET 21b and purified as carboxy-terminal hexahistidine-tagged proteins from *E. coli* BL21 (DE3) strain using nickel affinity chromatography as described before (Murdeshwar & Chatterji, 2012). RHII-mutant and all RSD-mutants were purified using the same protocol as that of MS_RHII-RSD protein. The purity of all the proteins was analyzed by SDS-PAGE and the concentration estimated spectrophotometrically using their extinction coefficients at 280 nm.

In vitro R-loop synthesis

To synthesize the R-loop, the first step was the mixing of equimolar concentration of two custom synthesized deoxyribonuleotides D1 and D2 having 80 bases each (sequences listed in Table. 2) in an annealing buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA). The mixture was kept at 93˚C for 5 minutes followed by gradual cooling to room temperature (RT) in 2 hours. The annealing of the oligonucleotides was confirmed by gel electrophoresis using 2% agarose in Tris Borate EDTA buffer (TBE). This annealed double-stranded DNA (D1-D2) and the 20 base ribonucleotide R3 (Table. 2) were then mixed in 1:3 molar ratio in R-loop synthesis buffer containing 83 mM PIPES, pH 7.8, 33 mM NaCl, 10 mM EDTA, and 70 % formamide (Landgraf *et al*., 1995). The mixture was incubated at 55˚C for 30 minutes, gradually cooled to 42˚C in 4 hours and then cooled to RT. R-loops were purified from the reaction mixture using Qiagen PCR clean-up kit. The R-loop formation was confirmed by running the sample in 2% agarose in TBE. All oligonucleotides and reaction components were obtained from Sigma-Aldrich. R-loop concentration was estimated based on the D1-D2 concentration used in the R-loop synthesis reaction, assuming 100 percent conversion of D1-D2 to R-loop.

Dot-blot analysis for R-loop detection

An anti-DNA-RNA hybrid mouse monoclonal antibody S9.6 (Kerafast, Inc., USA) was used for R-loop detection (Bhatia *et al.*, 2014). For dot blot, 30 and 60 ng of R-loop was separately spotted onto a Hybond-XL nylon membrane (Amersham Pharmacia Biotech), air-dried and UV cross-linked for 10 minutes in the transilluminator. The membrane was blocked in TBST (10 mM Tris-Cl, pH 7.4, 150 mM NaCl, 0.1% Tween-20) containing 5% (w/v) Bovine serum albumin and incubated overnight at 4°C with 5 µg S9.6 antibody. After washing the membrane was probed with anti-mouse HRP-conjugated secondary antibody (Sigma Aldrich) for 1 hour at RT and developed by chemiluminescence using Luminata HRP substrate (Millipore). The annealed double-stranded DNA (D1-D2) (60 ng) was used as a negative control.

R-loop hydrolysis assay

100 nM of the purified protein was taken and incubated with 400 nM of the R-loop in a buffer containing 50 mM Tris-Cl, pH 7.9, 60 mM KCl, 5 mM $MnCl_2$ at 37°C for 1 hour. *E. coli* RNase HI (NEB) was used as a positive control. The products were analyzed by (2%) agarose gel electrophoresis using TBE buffer.

In vivo R-loop detection

E. coli GJ 12055 parent strain and complemented strains (with plasmids pET-MS_RHII-RSD and pSK760) were grown in LB broth at 42°C and normalized to OD_{600} of 0.6. The cultures were split into duplicates in which one set was allowed to grow normally and the other irradiated with UV light in the laminar hood for 30 minutes and allowed to grow at 42°C. OD_{600} was measured immediately after UV exposure and subsequently at 1 hour intervals. Aliquots were taken immediately after UV exposure (0^{th} hour) and at two time intervals (1 and 4 hours after UV exposure). The genomic DNA was isolated using a kit (Favorgen Biotech Corp.) as per

manufacturer's instructions. 300 ng of genomic DNA was spotted onto a nylon membrane and dot-blotted as mentioned before. The experiments were repeated in triplicates and performed atleast twice. The same protocol was followed for *M. smegmatis* wild type and Δ*ms_rhII-rsd* except that the strains were grown at 37°C.

RNase H fluorescence assay

RNase H activity was determined by measuring the increase in fluorescence intensity of an artificially synthesized DNA:RNA hybrid (AH) following hydrolysis by RNase H as described previously (Murdeshwar & Chatterji, 2012). 100 μM of custom synthesized (Bioneer, Korea) 12-mers, 5'-fluorescein-polyA ribo-oligomucleotide and 3'-dabsyl-poly dT-deoxyribonucleotide were hybridized in annealing buffer (50 mM Tris-Cl. pH 7.5 at RT, 60 mM KCl) by incubating the mixture at 95°C for 5 minutes followed by slow cooling to RT. The hydrolysis reaction mixture contained 100 nM of the purified protein along with 500 nM AH in a reaction buffer (50 mM Tris-Cl. pH 7.5 at RT, 60 mM KCl, 5 mM $MnCl_2$). A control mixture lacking any protein was used as the negative control and *E. coli* RNase HI (NEB) was used as the positive control. The excitation and emission wavelength used were 480 nm and 520 nm respectively. To analyse the RHII activity of various oligomeric forms of the MS_RHII-RSD protein, the assay was performed using the same reaction mixture as described before but with the addition of 0.2% (w/v) SDS.

In vivo complementation assay

E. coli GJ 12055 cells were electrotransformed with the plasmids (expressing MS-RHII-RSD or its active site mutants). *E. coli* GJ 12055 cells complemented with pSK760 and pET21b were used as positive and negative control respectively. Serial dilutions of the log phase cultures

were prepared and 5 µl spotted on duplicate LB plates containing ampicillin. One set of plates was incubated at 30°C and the other at 42°C. The growth was analysed after 12 hours.

In vitro (p)ppGpp synthesis assay

The (p)ppGpp synthesis activity of the various proteins was assayed as described before (15) using 10 µl reaction mixtures containing 1 mM ATP, 10 µCi ml^{-1} [ϒ-^{32}P] ATP (3,000 Ci $mmol^{-1}$, BRIT, Hyderabad, India), 1 mM of either GTP or GDP, 50 mM HEPES (pH 7.5), 1 mM DTT, 10 mM $MgCl_2$ in the presence of 1 µM purified protein (Murdeshwar & Chatterji, 2012). To check the activity of various oligomeric forms of the MS_RHII-RSD protein, the (p)ppGpp synthesis assay was performed using the above mentioned reaction mixture but with the addition of 0.2% (w/v) SDS. Purified N-terminal domain of Rel_{Msm} was used as a positive control.

Gel filtration chromatography

Superdex 200 10/300 GL column (GE Health care) was preequilibrated with equilibration buffer (50 mM Tris-Cl. pH 7.9 at RT, 150 mM NaCl, 50 mM imidazole) and run at a flow rate of 0.2 ml min^{-1}. The detection wavelength used was 280 nm. The column was calibrated using molecular weight standards such as Thyroglobulin (670 kDa), Ferritin (440 kDa), ϒ Globulin (158 kDa), Ovalbumin (44 kDa), Myoglobin (17 kDa), Vitamin B12 (1.35 kDa) and a calibration curve was plotted using known standard protein mass and their elution volumes.

Native-PAGE analysis

Native PAGE was done in a manner similar to that of SDS-PAGE except for the absence of SDS in any of the gel components and electrophoresis buffer. The gel loading dye did not contain any SDS or β-ME and the samples were not boiled prior to loading. 20 µg of each protein taken from their peak fractions after gel filtration was loaded onto the gel. Electrophoresis was carried out at 20 mA at 20°C.

Circular dichroism analysis (CD)

The purified proteins were used at a final concentration of 2 μM in 10 mM Tris-Cl (pH 7.9 at 4°C). CD spectra were recorded from 190 nm to 250 nm on a Jasco J-715 spectropolarimeter using a quartz cuvette of 2 mm path length. The instrument parameters used were: 2 nm bandwidth, 2 seconds response time, 0.2 nm data pitch and 50 nm min^{-1} scanning speed. The CD spectra were averaged over 3 scans and buffer subtracted. The helical content of the proteins were estimated using the K2D2 software (http://www.ogic.ca/projects/k2d2/)

Quantitative Reverse Transcriptase assay (qRT-PCR)

qRT-PCR was done to determine the expression levels of MS_RHII-RSD (MSMEG_5849), *rnhA* (MSMEG_5562), *rnhB* (MSMEG_2442), and *rnh-cobC* fusion (MSMEG_4305) using the specific set of primers (Table. 1). *M. smegmatis* wild type was grown to an OD_{600} of 0.6. The culture was split into two and one set was exposed to UV radiation for 30 minutes in the laminar hood where as the other set was not given any UV exposure. The cultures were grown further for 4 hours. The cells were harvested; disrupted using BeadBeater (BioSpec) and the total RNA extracted using RNA isolation kit (Qiagen) as per manufacturer's instructions. RNA concentration was determined by measuring absorbance at 260 nm and the A_{260}/A_{280} confirmed to be around 1.9. cDNA was synthesized using random primers with a kit from Applied Biosystems as per the manufacturers protocol. SYBR green was used as the indicator dye. The RNA levels were normalized with respect to *rpoC* gene. The threshold value for significant change in expression level was calculated by students 't' test. Four independent experiments were done with each reaction being performed in triplicates and the graph plotted using R studio.

Generation of Δms_rhII-rsd strain

The *ms_rhII-rsd* gene deletion from *M. smegmatis* was carried out using phage-mediated specialized transduction method (Bardarov *et al.*, 2002). Approximately 1 kbp sequences upstream and downstream of the gene were amplified and cloned into the plasmid p0004S on either side of an apramycin resistance cassette. This generated the allele exchange substrate plasmid which was subsequently packaged *in vitro* (MaxPlax packaging kit) into a temperature-sensitive mycobacteriophage phAE159. *M. smegmatis* culture was transduced at 37°C and the colonies screened by PCR. The positive colonies were sequenced to confirm the gene deletion and one such colony was used for further experiments. All primers used are listed in supplementary table. S5.

Acknowledgments

DC thanks the Department of Biotechnology, Government of India, for funding this work. SK thanks the Department of Biotechnology, Government of India for the postdoctoral fellowship. AP acknowledges Indian Institute of Science, Bangalore for the fellowship. Financial support for the UK-India Research Initiative (UKIERI) enabling collaboration between IISc and the School of Biosciences, University of Birmingham is acknowledged. Robert Crouch, NICHD, National Institute of Health, Bethesda, MD, USA is acknowledged for the helpful discussions about R-loop synthesis. Farha Khan is acknowledged for the help in gel filtration analysis.

Author Contributions

SK, AP and DC were contributed for conception and design of the study. SK and AP performed and analysed all the experiments. The generation of MS_RHII-RSD knock out strain was contributed by AS and AB. SK, AP and DC participated in writing of the manuscript.

References

Aguilera, A. & T. Garcia-Muse, (2012) R loops: from transcription byproducts to threats to genome stability. *Molecular cell* **46**: 115-124.

Alzu, A., R. Bermejo, M. Begnis, C. Lucca, D. Piccini, W. Carotenuto, M. Saponaro, A. Brambati, A. Cocito, M. Foiani & G. Liberi, (2012) Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* **151**: 835-846.

Atkinson, G.C., T. Tenson & V. Hauryliuk, (2011) The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PloS one* **6**: e23479.

Avarbock, D., J. Salem, L.S. Li, Z.M. Wang & H. Rubin, (1999) Cloning and characterization of a bifunctional RelA/SpoT homologue from *Mycobacterium tuberculosis*. *Gene* **233**: 261-269.

Avarbock, D., A. Avarbock & H. Rubin, (2000) Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA. RelMtb complex. *Biochemistry* **39**: 11640-11648.

Bardarov, S., S. Bardarov Jr, Jr., M.S. Pavelka Jr, Jr., V. Sambandamurthy, M. Larsen, J. Tufariello, J. Chan, G. Hatfull & W.R. Jacobs Jr, Jr., (2002) Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology (Reading, England)* **148**: 3007-3017.

Cashel, M. & J. Gallant, (1969) Two compounds implicated in the function of the RC gene of *Escherichia coli*. *Nature* **221**: 838-841.

Cashel, M., D.M. Gentry, V.J. Hernandez & D. Vinella, (1996) The stringent response. In: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. . F.C. Neidhardt (ed). Washington D.C.: ASM Press, pp. 1458-1496.

Cerritelli, S.M. & R.J. Crouch, (2009) Ribonuclease H: the enzymes in eukaryotes. *The FEBS journal* **276**: 1494-1505.

Chatterji, D. & A.K. Ojha, (2001) Revisiting the stringent response, ppGpp and starvation signaling. *Current opinion in microbiology* **4**: 160-165.

Chon, H., J.L. Sparks, M. Rychlik, M. Nowotny, P.M. Burgers, R.J. Crouch & S.M. Cerritelli, (2013) RNase H2 roles in genome integrity revealed by unlinking its activities. *Nucleic acids research* **41**: 3130-3143.

Daniels, G.A. & M.R. Lieber, (1995) RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. *Nucleic acids research* **23**: 5006-5011.

Dawes, S.S., R.J. Crouch, S.L. Morris & V. Mizrahi, (1995) Cloning, sequence analysis, overproduction in *Escherichia coli* and enzymatic characterization of the RNase HI from *Mycobacterium smegmatis*. *Gene* **165**: 71-75.

Drolet, M., (2006) Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Molecular microbiology* **59**: 723-730.

Drolet, M., P. Phoenix, R. Menzel, E. Masse, L.F. Liu & R.J. Crouch, (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.

Dutta, D., K. Shatalin, V. Epshtein, M.E. Gottesman & E. Nudler, (2011) Linking RNA Polymerase Backtracking to Genome Instability. *Cell* **146**: 533-543.

Eder, P.S., R.Y. Walder & J.A. Walder, (1993) Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochimie* **75**: 123-126.

Gan, W., Z. Guan, J. Liu, T. Gui, K. Shen, J.L. Manley & X. Li, (2011) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes & development* **25**: 2041-2056.

Gowrishankar, J., J.K. Leela & K. Anupama, (2013) R-loops in bacterial transcription: their causes and consequences. *Transcription* **4**: 153-157.

Hauryliuk, V., G.C. Atkinson, K.S. Murakami, T. Tenson & K. Gerdes, (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nature reviews. Microbiology* **13**: 298-309.

Hogg, T., U. Mechold, H. Malke, M. Cashel & R. Hilgenfeld, (2004) Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response. *Cell* **117**: 57-68.

Itaya, M. & K. Kondo, (1991) Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile *Thermus thermophilus* HB8: a thermostable RNase H can functionally replace the *Escherichia coli* enzyme in vivo. *Nucleic acids research* **19**: 4443-4449.

Itaya, M., A. Omori, S. Kanaya, R.J. Crouch, T. Tanaka & K. Kondo, (1999) Isolation of RNase H genes that are essential for growth of *Bacillus subtilis* 168. *Journal of bacteriology* **181**: 2118-2123.

Jacewicz, A. & S. Shuman, (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. *Journal of bacteriology* **197**: 2489-2498.

Jain, V., R. Saleem-Batcha, A. China & D. Chatterji, (2006) Molecular dissection of the mycobacterial stringent response protein Rel. *Protein science : a publication of the Protein Society* **15**: 1449-1464.

Kamarthapu, V., V. Epshtein, B. Benjamin, S. Proshkin, A. Mironov, M. Cashel & E. Nudler, (2016) ppGpp couples transcription to DNA repair in *E. coli*. *Science* **352**: 993-996.

Kanaya, S. & R.J. Crouch, (1983) DNA sequence of the gene coding for *Escherichia coli* ribonuclease H. *The Journal of biological chemistry* **258**: 1276-1281.

Kogoma, T. & P.L. Foster, (1983) Physiological functions of *Escherichia coli* RNase HI. In: Ribonucleases H R.J. Crouch & J.J. Toulmé (eds). Paris: Les Editions INSERM, pp. 39–66.

Krásný, L. & R.L. Gourse, (2004) An alternate strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. *The EMBO Journal* **23**: 4473-4483.

Landgraf, R., C.H. Chen & D.S. Sigman, (1995) R-loop stability as a function of RNA structure and size. *Nucleic acids research* **23**: 3516-3523.

Landgraf, R., C.H. Chen & D.S. Sigman, (1995) Double stranded scission of DNA directed through sequence-specific R-loop formation. *Nucleic acids research* **23**: 3524-3530.

Lieberman, H.B. & E.M. Witkin, (1981) Variable expression of the ssb-1 allele in different strains of Escherichia coli K12 and B: differential suppression of its effects on DNA

replication, DNA repair and ultraviolet mutagenesis. *Molecular and General Genetics* **183**: 348-355.

Makowska, J., D. Wyrzykowski, B. Pilarski & L. Chmurzyński, (2015) Thermodynamics of sodium dodecyl sulphate (SDS) micellization in the presence of some biologically relevant pH buffers. *Journal of Thermal Analysis and Calorimetry* **121**: 257-261.

Merrikh, H., C. Machon, W.H. Grainger, A.D. Grossman & P. Soultanas, (2011) Co-directional replication-transcription conflicts lead to replication restart. *Nature* **470**: 554-557.

Minias, A.E., A.M. Brzostek, M. Korycka-Machala, B. Dziadek, P. Minias, M. Rajagopalan, M. Madiraju & J. Dziadek, (2015a) RNase HI is essential for survival of *Mycobacterium smegmatis*. *PloS one* **10**: e0126260.

Minias, A.E., A.M. Brzostek, P. Minias & J. Dziadek, (2015b) The deletion of rnhB in *Mycobacterium smegmatis* does not affect the level of RNase HII substrates or influence genome stability. *PloS one* **10**: e0115521.

Moriyama, Y. & K. Takeda, (1999) Re-formation of the helical structure of human serum albumin by the addition of small amounts of sodium dodecyl sulphate after the disruption of the structure by urea. A comparison with bovine serum albumin. *Langmuir* **15**: 2003-2008.

Mukherjee, R. & D. Chatterji, (2005) Evaluation of the role of sigma B in *Mycobacterium smegmatis*. *Biochemical and biophysical research communications* **338**: 964-972.

Murante, R.S., L.A. Henricksen & R.A. Bambara, (1998) Junction ribonuclease: An activity in Okazaki fragment processing. *Proc Natl Acad Sci U S A* **95**: 2244-2249.

Murdeshwar, M.S. & D. Chatterji, (2012) MS_RHII-RSD, a dual-function RNase HII-(p)ppGpp synthetase from *Mycobacterium smegmatis*. *Journal of bacteriology* **194**: 4003-4014.

Ohtani, N., M. Haruki, M. Morikawa, R.J. Crouch, M. Itaya & S. Kanaya, (1999) Identification of the genes encoding Mn2+-dependent RNase HII and Mg2+-dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families. *Biochemistry* **38**: 605-618.

Ohtani, N., N. Saito, M. Tomita, M. Itaya & A. Itoh, (2005) The SCO2299 gene from *Streptomyces coelicolor* A3(2) encodes a bifunctional enzyme consisting of an RNase H domain and an acid phosphatase domain. *The FEBS journal* **272**: 2828-2837.

Ojha, A.K., T.K. Mukherjee & D. Chatterji, (2000) High intracellular level of guanosine tetraphosphate in *Mycobacterium smegmatis* changes the morphology of the bacterium. *Infection and immunity* **68**: 4084-4091.

Paul, B.J., M.M. Barker, W. Ross, D.A. Schneider, C. Webb, J.W. Foster & R.L. Gourse, (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**: 311-322.

Poveda, A.M., M. Le Clech & P. Pasero, (2010) Transcription and replication: breaking the rules of the road causes genomic instability. *Transcription* **1**: 99-102.

Qadeer, A., M. Zaman & R.H. Khan, (2014) Inhibitory effect of post-micellar SDS concentration on thermal aggregation and activity of papain. *Biochemistry (Moscow)* **79**: 785-796.

Roberts, R.W. & D.M. Crothers, (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* **258**: 1463-1466.

Ross, W., C.E. Vrentas, P. Sanchz-Vasquez, T. Gaal & R.L. Gourse, (2013) The magic spot: A ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. *Molecular Cell* **50**: 420-429.

Ross, W., P. Sanchez-Vazquez, A.Y. Chen, J.H. Lee, H.L. Burgos & R.L. Gourse, (2016) ppGpp binding to a site at the RNAP-DksA interface accounts for its dramatic effects on transcription initiation during the stringent response. *Molecular Cell* doi: 10.1016/j.molcel.2016.04.029

Sajish, M., S. Kalayil, S.K. Verma, V.K. Nandicoori & B. Prakash, (2009) The significance of EXDD and RXKD motif conservation in Rel proteins. *The Journal of biological chemistry* **284**: 9115-9123.

Santos-Pereira, J.M. & A. Aguilera, (2015) R loops: new modulators of genome dynamics and function. *Nature reviews. Genetics* **16**: 583-597.

Shaw, B.F., G.F. Schneider & G.M. Whitesides, (2012) Effect of surfactant hydrophobicity on the pathway for unfolding of ubiquitin. *Journal of the American Chemical Society* **134**: 18739-18745.

Singh, R.B., S. Mahanta & N. Guchhait, (2008) Destructive and protective action of sodium dodecyl sulphate micelles on the native conformation of bovine serum albumin: A study by extrinsic fluorescence probe 1-hydroxy-2-naphthaldehyde. *Chemical Physics Letters* **463**: 183-188.

Skourti-Stathaki, K. & N.J. Proudfoot, (2014) A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes & development* **28**: 1384-1396.

Srivatsan, A. & J.D. Wang, (2008) Control of bacterial transcription, translation and replication by (p)ppGpp. *Current opinion in microbiology* **11**: 100-105.

Stein, H. & P. Hausen, (1969) Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. *Science* **166**: 393-395.

Stirling, P.C., Y.A. Chan, S.W. Minaker, M.J. Aristizabal, I. Barrett, P. Sipahimalani, M.S. Kobor & P. Hieter, (2012) R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes & development* **26**: 163-175.

Sun, D., G. Lee, J.H. Lee, H.Y. Kim, H.W. Rhee, S.Y. Park, K.J. Kim, Y. Kim, B.Y. Kim, J.I. Hong, C. Park, H.E. Choy, J.H. Kim, Y.H. Jeon & J. Chung, (2010) A metazoan ortholog of SpoT hydrolyzes ppGpp and functions in starvation responses. *Nature structural & molecular biology* **17**: 1188-1194.

Syal, K. & D. Chatterji, (2015) Differential binding of ppGpp and pppGpp to E. coli RNA polymerase: photo-labeling and mass spectral studies. *Genes to Cells* **20**: 1006-1016.

Tare, P., B. Mallick & V. Nagaraja, (2013) Co-evolution of specific amino acid in sigma 1.2 region and nucleotide base in the discriminator to act as sensors of small molecule effectors of transcription initiation in mycobacteria. *Molecular Microbiology* **90**: 569-583.

Tadokoro, T. & S. Kanaya, (2009) Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *The FEBS journal* **276**: 1482-1493.

Toriumi, K., T. Tsukahara & R. Hanai, (2013) R-Loop formation in trans at an AGGAG repeat. *Journal of Nucleic Acids* **2013**: 7.

Trautinger, B.W., R.P. Jaktaji, E. Rusakova & R.G. Lloyd, (2005) RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Molecular cell* **19**: 247-258.

Tresini, M., D.O. Warmerdam, P. Kolovos, L. Snijder, M.G. Vrouwe, J.A. Demmers, I.W.F. van, F.G. Grosveld, R.H. Medema, J.H. Hoeijmakers, L.H. Mullenders, W. Vermeulen & J.A.

Marteijn, (2015) The core spliceosome as target and effector of non-canonical ATM signalling. *Nature* **523**: 53-58.

Watkins, H.A. & E.N. Baker, (2010) Structural and functional characterization of an RNase HI domain from the bifunctional protein Rv2228c from *Mycobacterium tuberculosis*. *Journal of bacteriology* **192**: 2878-2886.

Weiss, L.A. & C.L. Stallings, (2013) Essential roles for *Mycobacterium tuberculosis* Rel beyond the production of (p)ppGpp. *Journal of bacteriology* **195**: 5629-5638.

Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser & M. Cashel, (1991) Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. *The Journal of biological chemistry* **266**: 5980-5990.

Xu, Q. & T.A. Keiderling, (2004) Effect of sodium dodecyl sulphate on folding and thermal stability of acid-denatured cytochrome c: A spectroscopic approach. *Protein Science* **13**: 2949-2959.

Yajima, H., K.J. Lee, S. Zhang, J. Kobayashi & B.P. Chen, (2009) DNA double-strand break formation upon UV-induced replication stress activates ATM and DNA-PKcs kinases. *Journal of molecular biology* **385**: 800-810.

Yu, K., F. Chedin, C.L. Hsieh, T.E. Wilson & M.R. Lieber, (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature immunology* **4**: 442-451.

Zhang, Y.B., S. Ayalew & S.A. Lacks, (1997) The rnhB gene encoding RNase HII of *Streptococcus pneumoniae* and evidence of conserved motifs in eucaryotic genes. *Journal of bacteriology* **179**: 3828-3836.

Zuo, Y., Y. Wang, T.A. Steitz, (2013) The mechanism of E.coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex. *Mol Cell* **50:** 430-436.

Table 1: List of qRT-PCR primers used in this study

Primer name	Gene	Sequence (5' to 3')
5849_FP	*ms_rhII-rsd*	CACTCGACCTCGATCTGGAC
5849_RP	*ms_rhII-rsd*	GCTTGGCCTTGTACTTCAGC
5562_FP	*rnhA*	CACCTGTACACCGACAGCAC
5562_RP	*rnhA*	CCACAGGTCGACATTCTTCA
2442_FP	*rnhB*	GCACCGGTACTCGTATGTGA
2442_RP	*rnhB*	CGTCGTACCTCTTCGGTCTC
4305_FP	*rnh_cobC*	TGTCGATGGACTCCAAACTG
4305_RP	*rnh_cobC*	CCACGTGTAGTCGACATGGT

Table 2: Sequences of oligonucleotides used for R-loop synthesis

Oligo-nucleotide Name	Length (bases)	Type	Sequence (5' to 3')
D1	80	Deoxy ribonucleotide	CTGGTTCACCACGCGGGAAACGGTCTGA TACTCCTCTCCTCTCCTCTCCTCGAGACA CCGGCATACTCTGCGACATCGTA
D2	80	Deoxy ribonucleotide	TACGATGTCGCAGAGTATGCCGGTGCAG TACTAGTCATGTCTACGCTAATGCCAGG ACCGTTTCCCGCGTGGTGAACCAG
R3	20	Ribonucleotide	AGG AGA GGA GAG GAG AGG AG

Figure Legends

Figure. 1. Pictorial representation of R-loop formation. Two single-stranded DNA of 80 bases (D1 and D2) with 30 non-complementary bases in the middle were annealed to form double-stranded DNA. 20 bases containing RNA (R3) complementary to template DNA (D1) was then added to form R-loop.

Figure. 2. R-loop hydrolysis assay and dot blot analysis of synthesized R-loops. (A) R-loop was synthesized by annealing two single-stranded DNA (D1 and D2) to form a double stranded DNA (D1-D2) followed by mixing with complementary RNA (R3) in R-loop synthesis buffer. This was confirmed by band mobility shift in 2 % Agarose gel electrophoresis. R-loop (D1-D2-R3) was incubated with MS_RHII-RSD protein at 37°C for 1 hour in Tris buffer, pH 7.4 containing

KCl and $MnCl_2$ and analysed for R-loop hydrolysis. *E. coli* RNase HI (100 nM) was used as a positive control and the reaction mix lacking any protein and RNase HII activity dead mutant (RHII mutant) were used as a negative control. Indicates R-loop; indicates ds DNA product after R-loop hydrolysis. (B) Dot blotting to confirm R-loop (D1-D2-R3) synthesis. The synthesized R-loop was spotted onto nylon membrane and UV cross-linked. Anti DNA-RNA hybrid antibody S9.6 was used as primary antibody and HRP signal was developed using chemiluminescence. ds DNA (D1-D2) was used as a negative control.

Figure. 3. Growth profile and *in vivo* R-loop analysis of UV-treated *E. coli* strains. (A) Growth curve of UV-treated and untreated *E. coli* GJ 12055, GJ complemented with MS_RHII-RSD and pSK 760 plasmids at 42°C. Strains were grown in LB and normalized to OD_{600} of 0.6. Cells were irradiated with UV for 30 minutes and OD_{600} was measured immediately after UV treatment and every hour from there on. Rectangular box represents impeded growth phase. Error bars represent the standard deviation (n=3). (B) Genomic DNA (300 ng) isolated from UV–treated *E.coli* GJ 12055, GJ complemented with MS_RHII-RSD and pSK 760 plasmids at three different time points (0, 1 and 4 hours post UV exposure) were dot blotted using anti DNA-RNA hybrid primary antibody S9.6.

Figure. 4. qRT-PCR analysis of *M. smegmatis* RNase H genes and *in vivo* detection of R-loop after UV-stress (A) RNA isolated from *M. smegmatis* at the 4^{th} hour post UV exposure was reverse transcribed to cDNA. qRT-PCR was performed to analyse the expression levels of four RNase H genes MSMEG_5562, MSMEG_2442, MSMEG_4305 and MSMEG_5849. The threshold cycle (Ct) values were normalized with respect to a house keeping gene, *rpoC*, encoding β′ subunit of RNA polymerase (Mukherjee & Chatterji, 2005). Fold change values are means ±SD of 4 different experiments performed in triplicates. (B) Genomic DNA (300 ng)

isolated from UV–treated and untreated *M. smegmatis* wild type and Δ*ms_rhII-rsd* at three different time points (0, 1 and 4 hours post UV exposure) were dot blotted using anti DNA-RNA hybrid primary antibody S9.6.

Figure. 5. Assays with RHII-mutant for RNAse H and (p)ppGpp synthetase activity. (A) *In vitro* fluorescence assay performed with Fluorescein-labeled-poly rA: Dabsyl-labeled poly dT annealed hybrid (AH) as the substrate. In this substrate, dabsyl moiety quenches the fluorescence. Hydrolysis of RNA:DNA hybrid leads to fluorescence. *E. coli* RNase HI was used as positive control. The maximum and minimum fluorescence are shown by Poly A and AH, respectively. (B) *In vivo* complementation assay using *E. coli* Δ*rnhA* Δ*rnhB* (ts) strain GJ12055 (GJ) electrotransformed with plasmids pSK760 (plasmid containing *E. coli rnhA* gene, positive control), pET21b (empty vector, negative control), MS_RHII-RSD and RHII-mutant. Serially diluted samples (10^{-1} to 10^{-6}) were plated in two sets; one set incubated at 30°C and another at 42°C overnight. (C) *In vitro (*p)ppGpp synthesis assay. *(*p)ppGpp synthesis assay was performed with 1 μM of purified MS_RHII-RSD and RHII-mutant protein using GTP and [γ-^{32}P]ATP as the substrate. Reaction mixture without any protein serves as a negative control. The reactions were spotted onto PEI-cellulose sheet, chromatography done using 1.5 M KH_2PO_4 as solvent.

Figure. 6. Assays with RSD-mutants for (p)ppGpp synthesis and RNase H activity. (A) *In vitro* (p)ppGpp synthesis assay was performed as described in Fig. 5C with 1μM of purified MS_RHII-RSD and RSD-mutants. (B) *In vitro* fluorescence assay was performed as described in Fig. 5A. (C) *In vivo* complementation assay was performed as described in Fig. 5B using *E. coli* GJ 12055 (GJ) electrotransformed with plasmids of MS_RHII-RSD and RSD-mutants.

with 0.2% SDS and 100 mM beta-mercaptoethanol (β-ME) in two different sample preparations.

Figure. 7. Effect of oligomerization on MS_RHII-RSD activity. (A) RNase H assay with native and 0.2% SDS-treated MS_RHII-RSD. *In vitro* fluorescence assay was performed as described in figure 5A. (B) (p)ppGpp synthesis assay was performed as described in Fig. 5C with native and 0.2 % SDS treated MS_RHII-RSD. N-terminal variant of Rel_{Msm} (Rel NTD) was used as a positive control.

Figure. 8. Model proposed for the physiological function of MS_RHII-RSD.

UV stress leads to increased R-loop formation and replication-transcription conflicts. Under UV-stress, MS_RHII-RSD expression is upregulated than the conventional RNase HI and HII. Its RHII activity removes the R-loops and the stalled RNA polymerase is destabilized indirectly by (p)ppGpp. Thus, MS_RHII-RSD plays an important role during R-loop induced replication stress response in *M. smegmatis.*

Supplementary Table - S1. ESI-MS spectra mass of Iodoacetamide labelled MS_RHII-RSD

Purified MS_RHII-RSD protein was treated with Iodoacetamide (57 Da) and the reaction mixture was then analyzed by LC-ESI-MS for any change in mass by 57 Da (per free Cysteine). 64259 Da and 64203 Da denote the mass of the protein with 2 and 1 free Cysteine residues respectively. 64148 Da is the mass of the unlabeled protein. MS_RHII-RSD has only 2 cysteines and the labeling of both cysteines by iodoacetamide indicated the absence of any disulphide linkage.

Supplementary Table – S2

Circular dichroism (CD) analysis of MS_RHII-RSD and its variants. Purified proteins (2 μM) in 10 mM Tris-Cl were taken and the CD spectra recorded. The helical content of the proteins were estimated using the K2D2 software. The full length MS_RHII-RSD protein and its mutant variants showed a high alpha helicity whereas the individual domains had a low helicity.

Supplementary Table – S3. List of strains and plasmids used in this study

Supplementary Table – S4. List of primers used in site directed mutagenesis

Supplementary Table – S5. List of primers used in generation and confirmation of Δ*ms_rhII-rsd* strain.

Supplementary Figure. S1. Growth profile of UV-treated *E. coli* strains. (A) Growth curve of UV-treated and untreated *E. coli* GJ 12055, GJ complemented with MS_RHII-RSD and pSK 760 plasmids at 30°C. Strains were grown in LB and normalized to OD_{600} of 0.6. Cells were irradiated with UV for 30 minutes and OD_{600} was measured immediately after UV treatment and every hour from there on. Error bars represent the standard deviation (n=3).

Supplementary Figure. S2. Growth profile of *M. smegmatis* wild type (WT) and Δ*ms_rhII-rsd*. Strains were grown in MB7H9 and normalized to OD_{600} of 0.6. Cells are irradiated with UV for 30 minutes. Then OD_{600} was measured immediately after UV exposure, after 1 hour, and then at every 4 hour interval from there on. Error bars represent the standard deviation (n=3).

Supplementary Figure. S3. SDS - PAGE analysis of the pure fractions of MS_RHII-RSD and its mutant variants. The recombinant protein MS_RHII-RSD wild type, RHII-mutant and RSD mutants (D356A, D357A, R358A, D359A) were purified to homogeneity using Ni-NTA and the normalized volume of the elutions were loaded in 12% gel.

Supplementary Figure. S4. Oligomerization studies with MS_RHII-RSD and its mutant variants and analysis of interaction mode in hexameric MS_RHII-RSD. (A) Gel filtration chromatography profile of MS_RHII-RSD, RHII-mutant and RSD-mutant (D356A). All the wild type and mutant proteins eluted between 12.6 to 12.9 ml and the corresponding elution volumes were extrapolated to estimate the molecular weight from the standard curve. (B) Native PAGE analysis of MS_RHII-RSD, RHII-mutant and RSD-mutants. 20 µg of each protein was loaded

onto 8% Native PAGE. Native PAGE marker, BSA and ferritin were used as molecular markers. Isoelectric point (pI) of protein sample is mentioned below each lane. (C) Native PAGE analysis of 20 μg of MS_RHII-RSD untreated and treated with increasing percentage of SDS concentration (0.05, 0.1, 0.5, 1, 1.25, 1.5 and 2 in lanes 1-7 respectively) were loaded onto 8% Native PAGE. BSA and ferritin were used as molecular markers. (D) Native PAGE analysis of MS_RHII-RSD treated with 0.2 % SDS and 100 mM beta-mercaptoethanol (β-ME) in two different sample preparations.

Supplementary Figure. S5. Circular Dichroism (CD) analysis of MS_RHII-RSD treated with 0.2 % SDS.

Supplementary Figure. S6. (p)ppGpp synthesis assay using GDP as the substrate was performed as described in Fig. 5C with native and 0.05 % SDS treated MS_RHII-RSD. N-terminal variant of Rel_{Msm} (Rel NTD) was used as positive control.

Supplementary Figure. S7. Circular Dichroism (CD) analysis of MS_RHII-RSD, RHII-mutant, RSD-mutants (D356A), RHII domain and RSD domain proteins.

Supplementary Figure. S8. Western blotting for analysis of MS_RHII-RSD expression levels after UV stress. *M. smegmatis* was grown till OD_{600} 0.6, exposed to 30 minutes UV stress and grown for 4 hours similar to the condition followed in qRT-PCR analysis. Crude lysates of *M. smegmatis* culture (UV treated and untreated) were blotted against anti-MS_RHII-RSD polyclonal antibodies as described before (Murdeshwar and Chatterji, 2012) . *M. smegmatis* sigmaA factor which is a house keeping gene was used as the internal loading control. Densitometric analysis was performed using MultiGauge software. Prof. B. Gopal, IISc., is acknowledged for the anti-sigA antibodies.

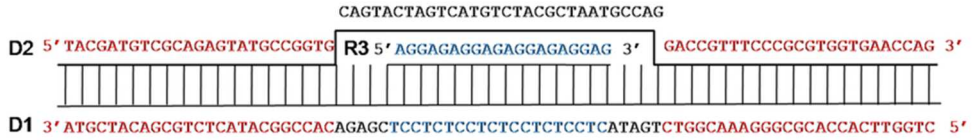


Figure. 1. Pictorial representation of R-loop formation. Two single-stranded DNA of 80 bases (D1 and D2) with 30 non-complementary bases in the middle were annealed to form double-stranded DNA. 20 bases containing RNA (R3) complementary to template DNA (D1) was then added to form R-loop.
36x5mm (600 x 600 DPI)

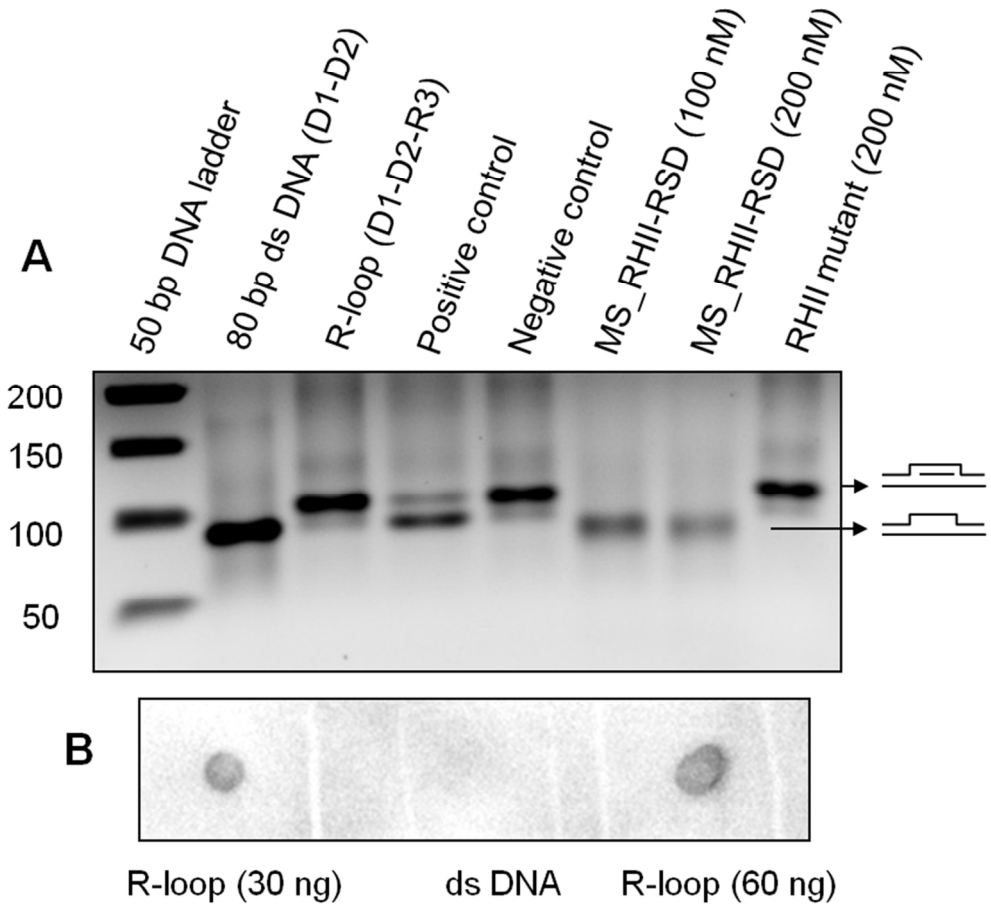


Figure. 2. R-loop hydrolysis assay and dot blot analysis of synthesized R-loops. (A) R-loop was synthesized by annealing two single-stranded DNA (D1 and D2) to form a double stranded DNA (D1-D2) followed by mixing with complementary RNA (R3) in R-loop synthesis buffer. This was confirmed by band mobility shift in 2 % Agarose gel electrophoresis. R-loop (D1-D2-R3) was incubated with MS_RHII-RSD protein at 37°C for 1 hour in Tris buffer, pH 7.4 containing KCl and MnCl2 and analysed for R-loop hydrolysis. E. coli RNase HI (100 nM) was used as a positive control and the reaction mix lacking any protein and RNase HII activity dead mutant (RHII mutant) were used as a negative control. Indicates R-loop; indicates ds DNA product after R-loop hydrolysis. (B) Dot blotting to confirm R-loop (D1-D2-R3) synthesis. The synthesized R-loop was spotted onto nylon membrane and UV cross-linked. Anti DNA-RNA hybrid antibody S9.6 was used as primary antibody and HRP signal was developed using chemiluminescence. ds DNA (D1-D2) was used as a negative control.
136x129mm (300 x 300 DPI)

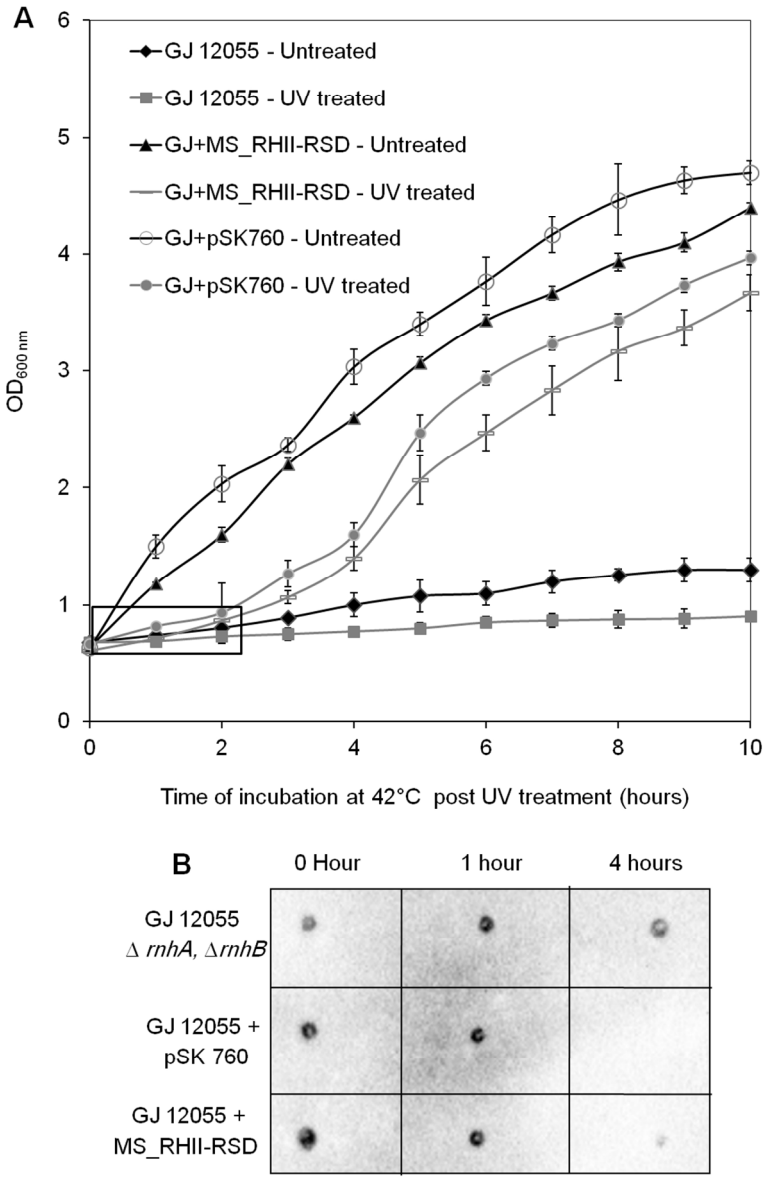


Figure. 3. Growth profile and in vivo R-loop analysis of UV-treated E. coli strains. (A) Growth curve of UV-treated and untreated E. coli GJ 12055, GJ complemented with MS_RHII-RSD and pSK 760 plasmids at 42°C. Strains were grown in LB and normalized to OD600 of 0.6. Cells were irradiated with UV for 30 minutes and OD600 was measured immediately after UV treatment and every hour from there on. Rectangular box represents impeded growth phase. Error bars represent the standard deviation (n=3). (B) Genomic DNA (300 ng) isolated from UV–treated E.coli GJ 12055, GJ complemented with MS_RHII-RSD and pSK 760 plasmids at three different time points (0, 1 and 4 hours post UV exposure) were dot blotted using anti DNA-RNA hybrid primary antibody S9.6.
266x393mm (300 x 300 DPI)

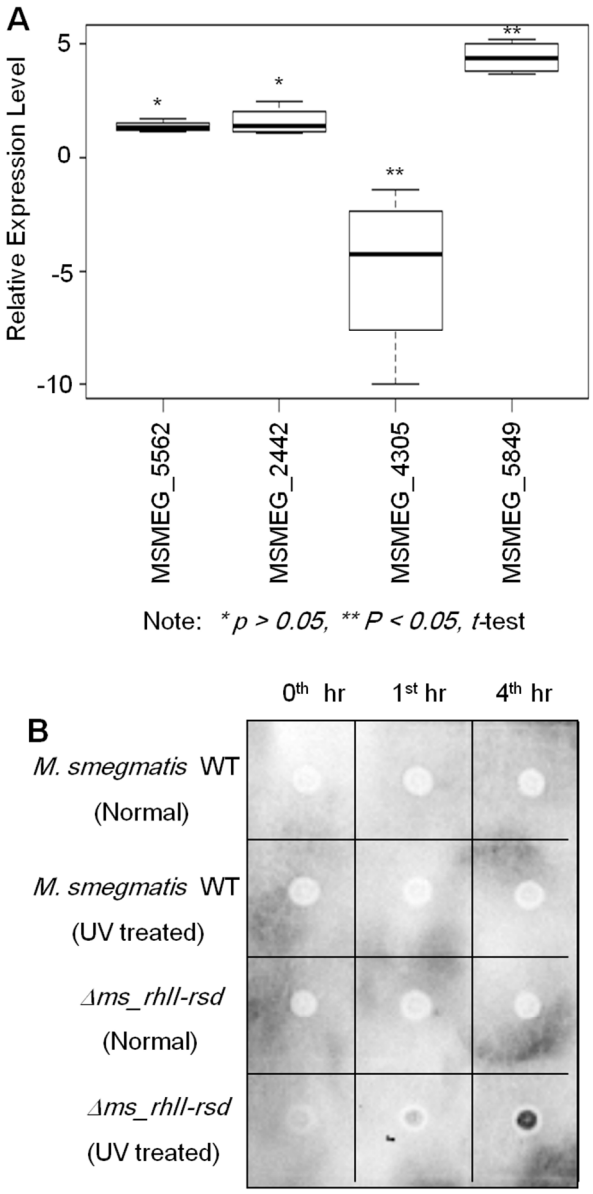


Figure. 4. qRT-PCR analysis of M. smegmatis RNase H genes and in vivo detection of R-loop after UV-stress (A) RNA isolated from M. smegmatis at the 4th hour post UV exposure was reverse transcribed to cDNA. qRT-PCR was performed to analyse the expression levels of four RNase H genes MSMEG_5562, MSMEG_2442, MSMEG_4305 and MSMEG_5849. The threshold cycle (Ct) values were normalized with respect to a house keeping gene, rpoC, encoding β' subunit of RNA polymerase (Mukherjee & Chatterji, 2005). Fold change values are means ±SD of 4 different experiments performed in triplicates. (B) Genomic DNA (300 ng) isolated from UV–treated and untreated M. smegmatis wild type and Δms_rhII-rsd at three different time points (0, 1 and 4 hours post UV exposure) were dot blotted using anti DNA-RNA hybrid primary antibody S9.6.
237x468mm (300 x 300 DPI)

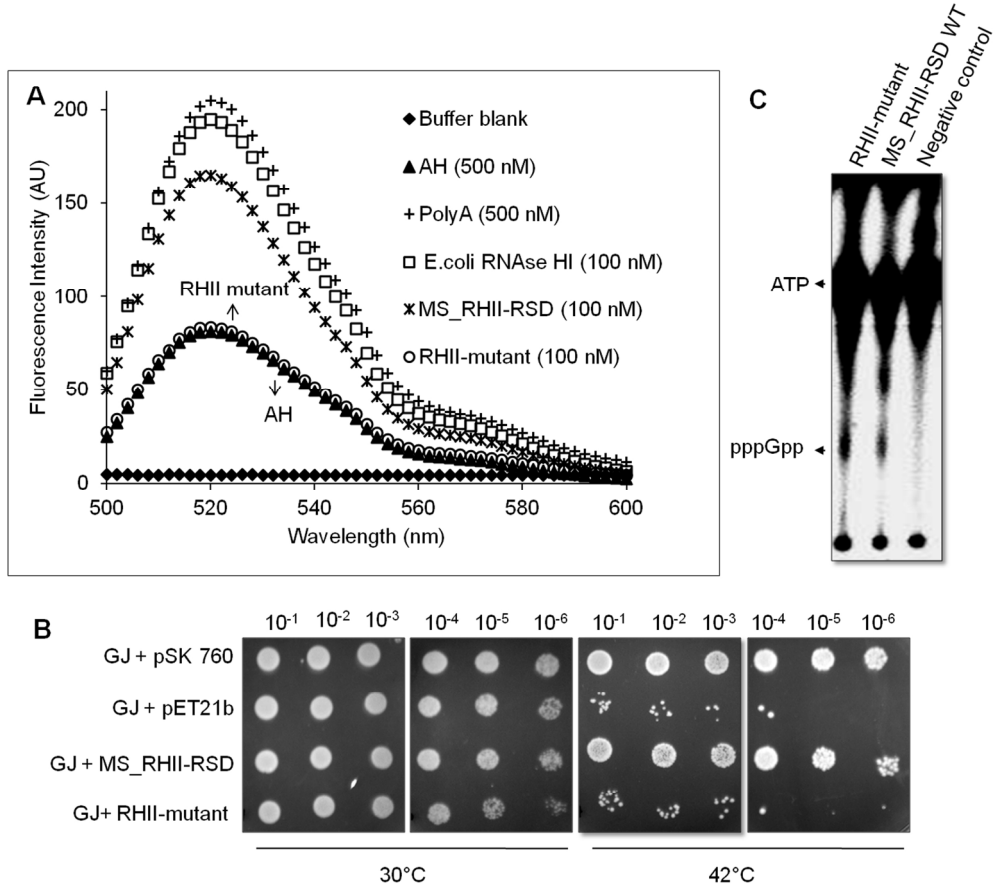


Figure. 5. Assays with RHII-mutant for RNAse H and (p)ppGpp synthetase activity. (A) In vitro fluorescence assay performed with Fluorescein-labeled-poly rA: Dabsyl-labeled poly dT annealed hybrid (AH) as the substrate. In this substrate, dabsyl moiety quenches the fluorescence. Hydrolysis of RNA:DNA hybrid leads to fluorescence. E. coli RNase HI was used as positive control. The maximum and minimum fluorescence are shown by Poly A and AH, respectively. (B) In vivo complementation assay using E. coli ΔrnhA ΔrnhB (ts) strain GJ12055 (GJ) electrotransformed with plasmids pSK760 (plasmid containing E. coli rnhA gene, positive control), pET21b (empty vector, negative control), MS_RHII-RSD and RHII-mutant. Serially diluted samples (10-1 to 10-6) were plated in two sets; one set incubated at 30°C and another at 42°C overnight. (C) In vitro (p)ppGpp synthesis assay. (p)ppGpp synthesis assay was performed with 1 μM of purified MS_RHII-RSD and RHII-mutant protein using GTP and [γ-32P]ATP as the substrate. Reaction mixture without any protein serves as a negative control. The reactions were spotted onto PEI-cellulose sheet, chromatography done using 1.5 M KH2PO4 as solvent.
214x193mm (300 x 300 DPI)

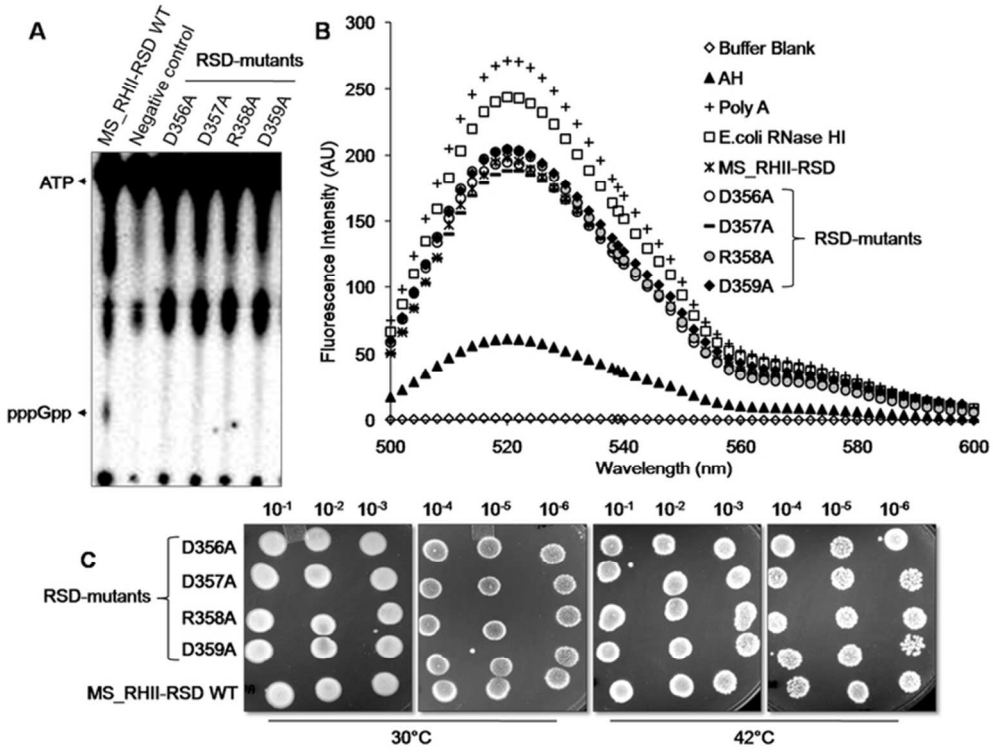


Figure. 6. Assays with RSD-mutants for (p)ppGpp synthesis and RNase H activity. (A) In vitro (p)ppGpp synthesis assay was performed as described in Fig. 5C with 1μM of purified MS_RHII-RSD and RSD-mutants. (B) In vitro fluorescence assay was performed as described in Fig. 5A. (C) In vivo complementation assay was performed as described in Fig. 5B using E. coli GJ 12055 (GJ) electrotransformed with plasmids of MS_RHII-RSD and RSD-mutants.
with 0.2% SDS and 100 mM beta-mercaptoethanol (β-ME) in two different sample preparations.
66x50mm (300 x 300 DPI)

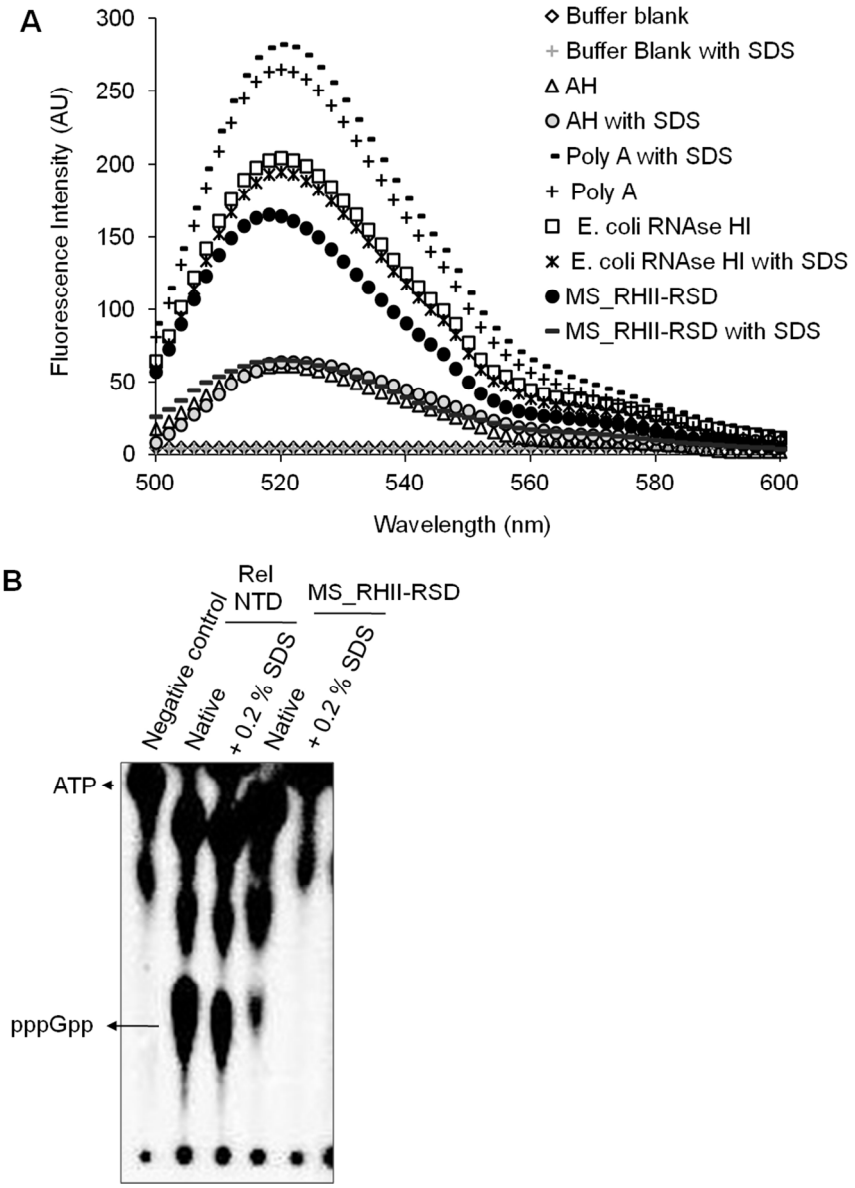


Figure. 7. Effect of oligomerization on MS_RHII-RSD activity. (A) RNase H assay with native and 0.2% SDS-treated MS_RHII-RSD. In vitro fluorescence assay was performed as described in figure 5A. (B) (p)ppGpp synthesis assay was performed as described in Fig. 5C with native and 0.2 % SDS treated MS_RHII-RSD. N-terminal variant of RelMsm (Rel NTD) was used as a positive control.
249x296mm (300 x 300 DPI)

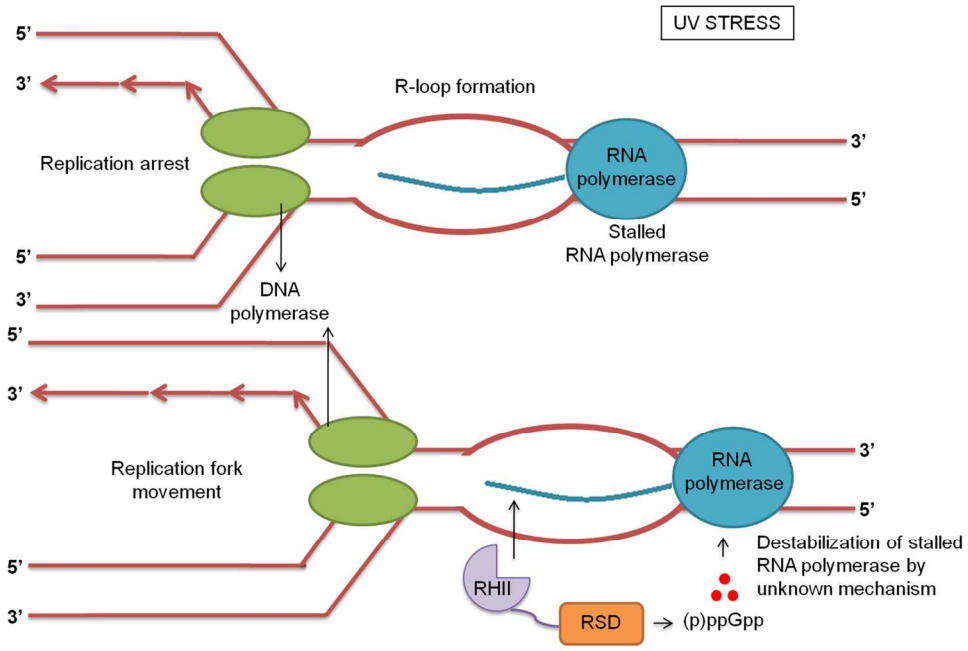
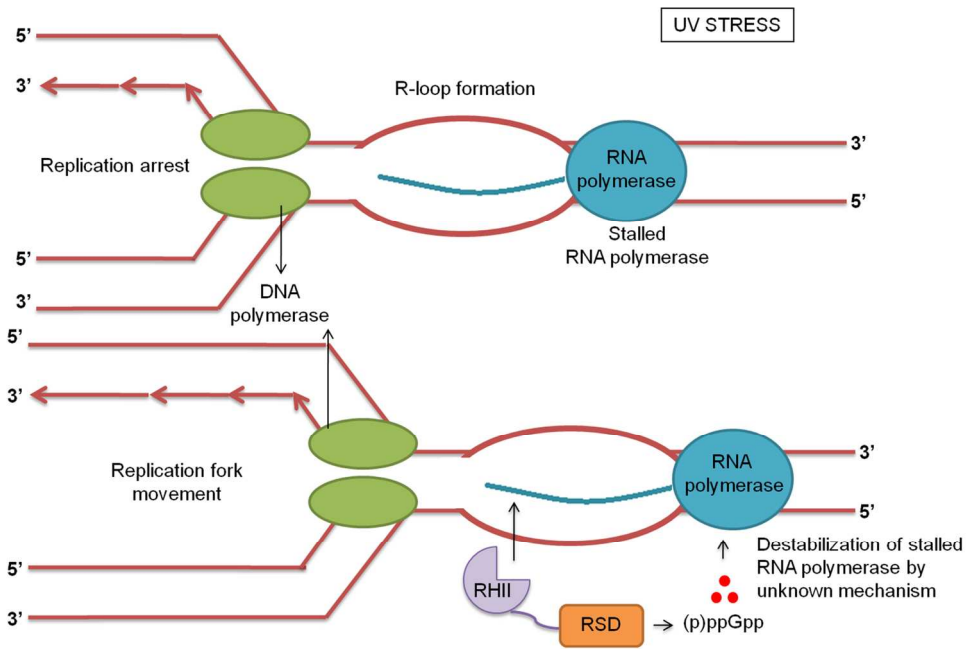
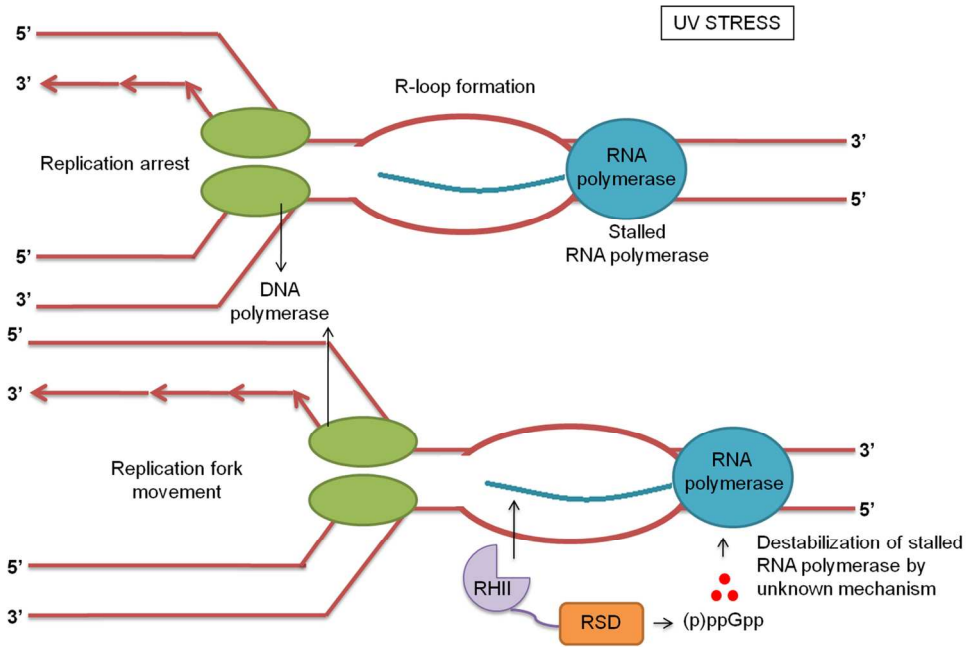


Figure. 8. Model proposed for the physiological function of MS_RHII-RSD.
UV stress leads to increased R-loop formation and replication-transcription conflicts. Under UV-stress, MS_RHII-RSD expression is upregulated than the conventional RNase HI and HII. Its RHII activity removes the R-loops and the stalled RNA polymerase is destabilized indirectly by (p)ppGpp. Thus, MS_RHII-RSD plays an important role during R-loop induced replication stress response in M. smegmatis.

164x105mm (300 x 300 DPI)



Model proposed for the physiological function of MS_RHII-RSD.
UV stress leads to increased R-loop formation and replication-transcription conflicts. Under UV-stress, MS_RHII-RSD expression is upregulated than the conventional RNase HI and HII. Its RHII activity removes the R-loops and the stalled RNA polymerase is destabilized indirectly by (p)ppGpp. Thus, MS_RHII-RSD plays an important role during R-loop induced replication stress response in M. smegmatis.
164x105mm (300 x 300 DPI)



164x105mm (300 x 300 DPI)

Abbreviated Summary

Due to the presence of a RNase HII domain and a (p)ppGpp synthesizing domain together in MS_RHII-RSD, R-loops are degraded and the interrupted RNA polymerase is destabilized by unknown mechanism to rescue the organism from replicative stress. Consistent with this observation, MS_RHII-RSD gene expression was upregulated under UV stress and increased R-loop accumulation was observed in this gene deleted strain. We have further elucidated its domain dependence for the optimal protein function.