Towards understanding R-loop regulation by RNase H1 in *Saccharomyces cerevisiae*

Dissertation

Zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
Der Johannes Gutenberg-Universität Mainz

Fábio Pereira Bento

geb. am 27.02.1992 in Alburitel, Portugal

Mainz, 2023

Dekan:

1. Berichterstatter:
2. Berichterstatter:

Tag der mündlichen Prüfung: 19/04/2023

Table of Contents

Summary

An RNA:DNA hybrid refers to the base pairing of RNA and DNA. In particular, R-loops are three-stranded nucleic acid structures composed of an RNA:DNA hybrid and a displaced DNA strand. In the recent years, several studies have shown that R-loops impact multiple biological processes, such as transcription, telomere maintenance, and DNA repair. Indeed, failure to resolve R-loops in a timely manner results in genomic instability. Hence, cells have evolved several mechanisms to tightly regulate R-loops. One of the most prominent example of an RNA:DNA hybrid removal factor is RNase H1, a monomeric enzyme that degrades the RNA moiety of RNA:DNA hybrids. In human cells, Replication Protein A (RPA) appears to be a regulator of RNase H1 by promoting its recruitment and activity on R-loops. However, it remains unclear if RPA or other proteins influence RNase H1 activity in yeast. RNase H1 is recruited to the chromatin in conditions with high R-loop levels, suggesting that RNase H1 binds to accumulated R-loops. However, how RNase H1 is regulated and responds to high R-loop levels is still elusive.

In this study, we have employed *in vitro* and *in vivo* experiments to understand where, when and how RNase H1 binds to R-loops. We confirm that RNase H1 is recruited to loci with accumulated R-loops, especially in RNAPIII-transcribed genes. Furthermore, we show that RNase H1 can remove R-loops in all cell cycle phases, allowing RNase H1 to maintain R-loop homeostasis in all conditions, including outside of a replication-transcription conflict context.

Using co-immunoprecipitation experiments, we confirm that RPA interacts with the first Hybrid-binding (HB) domain of RNase H1 in a DNA-dependent manner. However, overexpression of RNase H1 without the first HB domain is able to remove R-loops *in vivo*, suggesting that RNase H1 overexpression does not require RPA interaction to remove R-loops in yeast. Moreover, using *in vivo* and *in vitro* assays, we show that the second HB domain of RNase H1 is sufficient to promote binding and removal of R-loops, unlike what was described. We confirm that both HB domains of RNase H1 contribute to a stable interaction with R-loops and likely both HB domains have different functions in RNase H1 binding to substrates. Taken together, these results shed light on how RNase H1 recognizes and responds to stable RNA-DNA hybrids.

Zusammenfassung

Die Basenpaarung von RNA und DNA bildet ein RNA:DNA-Hybrid. R-Loops sind dreisträngige Nukleinsäurestrukturen, zusammengesetzt aus einem RNA:DNA-Hybrid und dem verdrängten DNA-Strang. In den letzten Jahren haben mehrere Studien gezeigt, dass sich R-Loops auf mehrere biologische Prozesse auswirken, wie z. B. auf die Transkription, die Erhaltung der Telomere und die DNA-Reparatur. Wenn es nicht gelingt, R-Loops rechtzeitig aufzulösen, führt dies zu genomischer Instabilität. Daher haben die Zellen mehrere Mechanismen entwickelt, um R-Loops genau zu regulieren. Eines der bekanntesten Beispiele zur Beseitigung von RNA:DNA-Hybriden ist RNase H1, ein monomeres Enzym, das den RNA-Anteil von RNA:DNA-Hybriden abbaut. In menschlichen Zellen scheint dies das Replikationsprotein A (RPA), ein Regulator der RNase H1, zu übernehmen, indem es deren Rekrutierung und Aktivität an R-loops begünstigt. Es ist jedoch noch unklar, ob RPA oder andere Proteine die Aktivität der RNase H1 in Hefe beeinflussen. RNase H1 wird durch erhöhte R-Loop Aufkommen an das Chromatin rekrutiert, was darauf hindeutet, dass RNase H1 lokale Ansammlungen von R-Loops bindet. In Hefe ist die Regulation von RNase H1, sowie die Rekrutierung von RNase H1 an vermehrte R-Loop Ansammlungen jedoch noch nicht geklärt.

In dieser Studie haben wir *in vitro* und *in vivo* Experimente durchgeführt, um zu verstehen, wo, wann und wie RNase H1 an R-Loops bindet. Wir konnten zeigen, dass RNase H1 an Loci mit akkumulierten R-Loops rekrutiert wird, insbesondere in RNAPIII-transkribierten Genen. Darüber hinaus zeigen wir, dass RNase H1 R-Loops in allen Phasen des Zellzyklus entfernt werden kann, so dass RNase H1 das R-Loop Gleichgewicht unter allen Bedingungen aufrechterhalten kann, auch außerhalb eines Replikations-Transkriptionskonflikts.

Mit Hilfe von Co-Immunopräzipitationsexperimenten bestätigen wir, dass RPA mit der ersten Hybrid-bindenden (HB) Domäne von RNase H1 in einer DNA-abhängigen Weise interagiert. Eine Überexpression von RNase H1 ohne die erste HB-Domäne ist jedoch in der Lage, R-Loops *in vivo* zu entfernen. Dies deutet darauf hin, dass eine Überexpression von RNase H1 keine RPA-Interaktion erfordert, um R-Loops in Hefe zu entfernen. Darüber hinaus zeigen wir mit Hilfe von In-vivo- und In-vitro-Versuchen, dass die zweite HB-Domäne von RNase H1 ausreicht, um die Bindung und Entfernung von R-Loops zu ermöglichen, im Gegensatz zu dem, was bisher publiziert wurde. Wir bestätigen, dass beide HB-Domänen von RNase H1 zu einer stabilen Interaktion mit R-Loops beitragen. Höchstwahrscheinlich, nehmen bei der

Bindung von RNase H1 zu den Substraten, beide HB-Domänen unterschiedliche Funktionen ein. Zusammengenommen geben diese Ergebnisse Aufschluss darüber, wie RNase H1 stabile RNA-DNA-Hybride erkennt und auf sie reagiert.

1   Introduction

Early on, it was understood that the interaction between DNA and RNA is at the core of molecular biology. In the early 1960s, the first hybridization of DNA and RNA molecules was achieved and this was the first step toward understanding the importance of the interaction between DNA and RNA. The discovery that RNA and DNA could form a hybrid helix was relevant to understanding the activity of RNA polymerase, how reverse transcriptase works, as well as telomerase, retrotransposons, and other interactions in which DNA and RNA combine. In this thesis, we studied RNA:DNA hybrids and how they are regulated by RNase H1, using *Saccharomyces cerevisiae* as a model organism. In the introduction section, I will highlight the most relevant aspects of RNA:DNA hybrids to facilitate a discussion of the experimental approach.

1.1   RNA:DNA hybrids

The term RNA:DNA hybrids refers to the base pairing of ribo- and deoxyribo-nucleotides. RNA:DNA hybrids occur naturally during replication and transcription. For example, in lagging strand synthesis, 11 nucleotide long RNA primers are generated by the DNA primase. These hybrids also form inside the active site of the RNA polymerase (RNAP), where an 8bp RNA:DNA duplex forms at the transcription bubble[1].

When an RNA:DNA hybrid leads to the displacement of single-stranded DNA (ssDNA) and hence the formation of a three-stranded RNA and DNA structure, it is referred to as an R-loop (Figure 1). R-loops were first identified in the 1970s[2,3], but recent studies indicate how important they are in multiple cellular processes[4,5]. For a long time, R-loops were considered transcription by-products that could interfere with transcription and cause genome instability[6,7]. However, recent studies suggest that R-loop-induced genomic instability is related with failure to resolve R-loops in a timely manner[8,9].

RNA:DNA hybrids are more stable than double-strand DNAs (dsDNAs)[10] and the relative stability of these hybrids depends on the oligomeric length, the content of deoxypyrimidines/deoxypurines, and the A-T/U proportion[11]. Once formed, removal of an R-loop may be a costly energy-consuming process. Therefore, cells evolved multiple mechanisms to regulate hybrids and preserve genome integrity.

1.1.1 Non-R-loop hybrids

A class of RNA:DNA hybrids, not considered R-loops as they do not result in displacement of a DNA strand, has been known to have key roles in transcription and DNA replication. During DNA synthesis, an RNA primer synthesized by DNA polymerase α constitutes an RNA:DNA hybrid essential for replication of the lagging strand. This RNA primer is used by DNA polymerase δ to polymerize mature Okazaki fragments[12].

RNA:DNA hybrids may also arise from misincorporated ribonucleotide triphosphates into the newly synthesized DNA strand. The resulting molecule is composed of a DNA backbone with intercalated ribonucleoside monophosphates (rNMPs), which increase the risk of DNA hydrolysis, especially in alkaline conditions[13]. Although misincorporated rNMPs can compromise genome integrity when not properly excised, they have also been proposed to promote mismatch repair and non-homologous end-joining pathways (NHEJ)[14–16]. In the case of mismatch repair, the rNMPs can act as a signal to distinguish the strands and direct the repair machinery on the newly synthesized strand.

1.1.2 R-loops

R-loops are three-stranded structures consisting of an RNA:DNA hybrid and the displaced strand of DNA. R-loops are typically associated with transcription, and occur *in cis*, at the site of transcription. One possibility for R-loop formation during transcription is termed the 'extended hybrid' model, based on which the nascent RNA remains annealed to its DNA template due to strong interaction between RNA and DNA[10]. However, it is unlikely that R-loops are simply an extension of the usual 8bp RNA:DNA hybrid formed within the transcription bubble, since it does not reconcile with the crystal structure of RNAP that shows RNA and DNA exit from two distinct channels[17]. R-loops are likely formed through a second model, known as the thread back model, which proposes that the 5' end of RNA re-anneals with the DNA. This is supported by the observation that DNA behind the transcribing polymerase is negatively supercoiled[18]. This negatively supercoiled DNA has increase tendency to unwind, which may allow the nascent RNA to re-anneal with the template strand[19]. Moreover, previous studies demonstrated that co-transcriptional R-loops require a free RNA end and a GC skew[20,21]. This model is also supported by the observation that mutants defective in transcription elongation, splicing and relaxation of supercoiled DNA show elevated R-loop levels[22–25]. Defects in termination and elongation factors might stall the RNAP, prolonging the negatively supercoiled state of the DNA and the proximity of nascent RNA. Absence of splicing

factors may expose the RNA, making it more accessible to hybridize with DNA. Lack of topoisomerases increase the unwound DNA by preventing relaxation of the negative supercoils[26]. Based on these observations, the thread back mechanism is the predominant model for hybrid formation.

However, this co-transcriptional model for R-loop formation does not explain why R-loops are not exclusively found at sites of ongoing transcription[27]. Besides this, hybrids accumulate in mutants that impact RNA post-transcriptionally, like RNA export and degradation. This suggested that some R-loops may persist or be formed post-transcriptionally. Indeed, a recent study showed that R-loop can form *in trans*, when the RNA is transcribed at a spatially distinct locus[25]. The formation of R-loops *in trans* is promoted by Rad51 and Rad52, proteins involved in homologous recombination (HR)-mediated repair after double-strand break (DSB)[28]. *In vitro*, the bacterial homologue of RAD51, RecA mediates the formation of hybrids *in trans* post-transcriptionally[29,30]. Interestingly, in some cases ATPase activity may not be required for hybrid formation. The CRISPR-Cas system can cause deformation of the duplex DNA, leading to DNA unwinding and hybridization of the guide RNA[31]. Therefore, unwinding of the double-stranded DNA may be sufficient for R-loop formation *in trans*, notably in negatively supercoiled DNA[32].

Besides DNA supercoiling, other features of RNA sequence or its DNA template contribute to R-loop formation. DNA nicks and high GC content have been shown to affect R-loop formation *in vitro*[19,27,33]. Single strand nicks may increase R-loop formation likely due to the nicked non-template DNA being transiently displaced, thus facilitating the RNA binding to the template strand[19]. Additionally, in yeast, GC rich sequences tend to harbor more R-loops, but these GC rich genes are typically more highly expressed[27,33]. It is important to note that yeast genome does not have much GC skew, except its telomeres. In general, sequence per se is not a crucial determinant of R-loop accumulation in yeast, but rather the transcription rate at the locus. Indeed, a locus without R-loops can be converted into an R-loop-rich locus by changing the promoters to boost rates of transcription[20]. In yeast, genome-wide sequencing of hybrids show a strong presence in telomeres, retrotransposons and highly expressed genes, such as the ribosomal RNA (rRNA) and tRNA genes and other ncRNAs[27,33]. In human cells and plants, R-loops accumulate in repetitive sequences, such as transposable elements, ribosomal DNA (rDNA), centromeres and telomeres, similar to yeast[34–37]. Moreover, in humans and plants, R-loops are predominantly at promoter regions with CpG islands, specifically between the transcription start site (TSS) and the first intron-exon junction[38]. Indeed, measurements of R-loop half-life shows that hybrid lifetimes correlates with GC content, with more stable hybrids having higher GC content[39]. Interestingly, in humans, despite the lack of correlation between R-loop levels and transcriptional outputs, there is a general link between

R-loop formation and the RNAPII pausing status at TSS regions in a GC content-dependent manner[20]. Overall, a strong GC skew correlates with high expression and with R-loop enrichment. In addition to promoter regions, genome-wide studies found R-loop accumulation at transcription termination sites[40,41]. Also, hybrids are formed at sites of DNA damage, including at dysfunctional telomeres and DSB[42]. Although R-loops are generally thought to form behind the polymerase, a recent study suggest that they can form anterior to RNA polymerase II during polymerase backtracking[43].

Accumulation of R-loops in GC-rich sequences may be caused by strong thermodynamic binding features between G-rich RNA and the complementary DNA sequence[10,44]. Moreover, the formation of DNA secondary structures on the displaced strand, such as G-quadruplexes (G4), can increase R-loop stability[45–47], possibly by preventing access of R-loop resolving proteins or by decreasing the reannealing capacity of the DNA. Indeed, a recent study identified a subset of RNA:DNA hybrids with GC skew that are partially resistant to RNase H, an R-loop resolving protein[39]. These G-rich sequences are prone to form G4s, confirming a strong correlation between G4s and R-loop resistance RNase H. Additionally, an increase in G4-forming sequences correlated with longer hybrid lifetimes[39]. Besides G4s, an *in vitro* study using atomic force microscopy (AFM) showed that other secondary structures can form at the displaced DNA strand of the R-loop[48]. These 'R-loop objects' could cause physical changes on the DNA, such as bending the DNA[9], which can lead to *in trans* R-loop formation. Such DNA secondary structures can constitute another layer of R-loop regulation as well as potential regulatory functions by the R-loop[9].

One conserved characteristic of R-loops is their transient nature. In yeast, a strong nuclear RNA:DNA hybrid signal can be detected only upon loss of R-loop resolving enzymes, such as RNase H enzymes[49]. This suggests that hybrids frequently arise, but are rapidly removed. Moreover, R-loop quantification in human cells estimates their abundance at 300 R-loops/cell, with an average half-life of 11 to 15 minutes. This suggests that cells resolve on the order of 27000 R-loops per day[39,50].

Deregulation of R-loops levels has been associated with multiple human disease, including neurological disorders, cancer, and autoimmune diseases. In these situations, R-loops may drive genome instability, alter gene expression, or alter cellular signaling pathways[42,51,52]. Furthermore, progeria in human and other model organisms is associated with defects in genes involved in R-loop biology, such as WRN, ERCC1, EPF and XPG[53–56].

Multiple authors have been trying to categorize R-loops depending on its impact on genome stability and/or its regulatory functions. Some authors differentiate R-loops as physiological and pathological or unscheduled or regulatory R-loops[4,9].The general consensus

is that hyper stabilized R-loops can lead to genome instability, but overall transient and tightly regulated R-loops have regulatory functions essential for genome stability, from transcription termination to DNA repair.

1.1.3 Methods of detection

The methodology for detecting RNA:DNA hybrids has been critical for understanding R-loop formation and regulation. Methods that involved measuring genome instability in combination with RNase H1 overexpression were widely used to observe the indirect effects of stabilized R-loops, as well to identify genes involved in R-loop resolution. Other direct methods have been developed to detect RNA:DNA hybrids. These include electron microscopy[47], isolation and analysis of nucleic acids sensitive to RNase H, while resistant to RNase A[22] and chromatin/DNA immunoprecipitation (ChIP/DIP) and immunofluorescence using monoclonal antibody produced by S9.6 hybridoma[24,26,27,33,57] or an inactivated RNase H[34]. Also, R-loops can be inferred *in vivo* by the mutation profile caused by sodium sulfide[58]. This acts specifically on ssDNA to convert C to U. The mutation profile of the R-loop displaced ssDNA allows the inference of the average length of R-loops, as long as such feature is abolished by RNase H1 treatment. Most direct techniques to detect R-loops, outside of bisulfide based sequencing techniques, probe for the RNA:DNA hybrid itself and thus cannot distinguish between an R-loop or non-R-loop hybrid[59]. Hence, in some scenarios, it is not clear what structure is actually formed.

The most widely adopted method for R-loop mapping is DRIP-seq[34], which uses next generation sequencing to map R-loops isolated by S9.6 imunoprecipitation[26]. DRIP-seq has proved to be a consistent and reproducible method for sequencing R-loops. However, there are some limitations. Fragmenting the genome with restriction enzymes introduces bias and limits resolution[60], which can be addressed by sonicating DNA rather than using restriction enzymes. Moreover, DRIP-seq is limited by its strand insensitivity. This has been be addressed by strand specific DNA library preparation as in ssDRIP-seq[35]. Alternatively, DRIP-RNA-seq[61], RDIP-seq[37], and DRIPc-seq[50] address strand specificity by sequencing the RNA moiety of the hybrid instead of the DNA. Finally, bis-DRIP-seq[38] combines in situ ssDNA bisulfite footprinting with S9.6 hybrid pulldown, which theoretically improves specificity by targeting both the hybrids and ssDNA. Although none of these methods are currently widely adopted, they provide solutions for some of DRIP-seq limitations.

Any variant of DRIP-seq presumes that S9.6 has an unbiased and specific affinity for hybrids. However, S9.6 has some affinity for dsRNA and show biases in hybrid sequence

recognition[62,63]. Although sequence bias of S9.6 has no effect on the composition of sequences detected by DRIP-seq, it could affect the recovery of short hybrids[59,64]. Binding to dsRNA is mostly relevant to methods that sequence hybrid RNA, like DRIPc-seq, and can be mitigated using RNase III treatment before pulldown[65]. Besides this, different DRIP-seq variants show variable results since it has been shown that R-loops are associated with dA:dT tracks[66] as well regions of high GC skew[34].

An alternative approach to DRIP-seq uses a catalytic inactive form of RNase H1, which is able to recognize but not process RNA:DNA hybrids[67]. DRIVE-seq[34] is conceptually similar to DRIP-seq, using tagged, catalytically inactive RNase H to pulldown hybrids *in vitro*. However, its adoption has been limited, as it is less sensitive than DRIP-seq. RNase H1 is more widely used to capture RNA:DNA hybrids by overexpression of catalytically inactive RNase H1 *in vivo*, allowing to pulldown hybrids in their native context. Theoretically, this limits the opportunity for hybrids to dissociate or shift in position before immunoprecipitation. This is the case for R-ChIP, in which catalytically inactive RNase H1 is stably expressed in cells and immunoprecipitated from cross linked chromatin[20]. Expressing the RNase H construct endogenously improves sensitivity and allows for *in situ* capture of hybrids, although its stable expression could in principle alter the turnover of R-loops. Catalytically inactive RNase H1 has also been used in a CUT&RUN-based approach named MapR, for targeting micrococcal nuclease (MNase) to R-loops, allowing to isolate them out of chromatin[68]. Similarly, RNase H1 hybrid binding domain has been also used in a CUT&Tag approach to map RNA:DNA hybrids[69].

Interestingly, the methods mentioned above show differences in hybrids distribution patterns. On one hand, DRIP-based studies show that R-loops mainly form along transcribed gene bodies, with hotspots along GC skewed CpG island promoters and termination regions[50]. On the other hand, bis-DRIP- seq and R-ChIP-seq both involve an *in situ* step and show R-loops to be concentrated at G-rich promoters associated with pausing of RNAPII[20] and almost entirely absent from the 3' end of genes. As much of the DRIPc-seq and DRIP-seq signal is sensitive to RNase H[34,61,70], the difference in signal outside of promoters are probably not due to off target binding. Likely, these differences reflect differences between capturing hybrids *in situ* and after cell lysis. RNase H recognizes a promoter-proximal subset of hybrids in a chromatin context, which could reflect a biological role of these sites or increased accessibility at a promoter by open chromatin[71].

It has been discussed that RNase H1-based approaches are more specific than S9.6-based approaches due to S9.6 specificity to RNA-RNA duplexes besides RNA:DNA hybrids. Indeed, S9.6 shows high affinity for RNA:DNA hybrids (0.6nM) and RNA-RNA duplexes (2.7nM)[62,63], while the isolated hybrid-binding domain of human RNase H1 has 0.2μM K_D for

RNA:DNA hybrids and 4.9μM K_D for dsRNA[72]. Both S9.6 and RNase H1 ability to distinguish RNA:DNA hybrid and dsRNA substrates is questioned, with RNase H1 showing 25 fold lower affinity for dsRNA, while S9.6 showing 4.5 times lower affinity for dsRNA compared with RNA:DNA hybrids. Based on this biochemical studies, RNase H1 shows higher specificity towards RNA:DNA hybrids compared to other substrates, but S9.6 is more sensitive to RNA:DNA hybrids than RNase H1 hybrid binding (HB) domain. Another advantage of RNase H1-based methods is that they capture R-loops *in vivo*, whereas DRIP-based approaches capture hybrids *ex vivo*, after nucleic acid extraction. However, overexpression of catalytically inactive RNase H1 may affect the physiological turnover and stability of R-loops, due to a dominant negative effect on the catalytic active endogenous RNase H1[73,74]. This could potentially change the patterns of R-loop distribution and contribute to differences with S9.6-based approaches.

Recently, a comparison over multiple loci showed consistency between bisulfite-based and S9.6-based techniques, but differences with RNase H1-based approaches[75]. The results of bisulfite footprinting have been corroborated by AFM-based visualization of R-loops[48] as well by mathematical modeling[32]. Besides this, bisulfite approaches confirmed that R-loops are a characteristic of 3' end of genes as described in S9.6-based techniques[75]. This could suggest that S9.6-based and bisulfite-based techniques provide a wider distribution pattern for R-loops than RNase H1-based approaches. However, in DRIP-based approaches, we could be observing loss of short, unstable R-loops upon DNA fragmentation prior to DRIP[32,64]. As a result, such small promoter-associated R-loops, which are stabilized *in vivo* by DNA topology, may be best captured by in situ approaches relying on crosslinking, such as R-ChIP[59].

Other S9.6-based approaches, like immunofluorescence or dot-imuno-blots, suffer with the interference of other RNA species. In this case, enzymatic treatment to remove contaminant RNA, such as RNase T1 and RNase III treatment, is necessary to have more reliable results[76]. It is important to note that RNase A should not be used, as this enzyme has RNase H activity that is hard to control for.

1.1.4 R-loops regulatory functions

RNA:DNA hybrids and R-loops are intermediates of transcription and DNA replication. One of the first examples of regulatory R-loops was at the immunoglobulin heavy chain locus. R-loops evolved as a natural source of Ig class-switch recombination (CSR) and form within the G-rich switch regions. The action of cytidine deaminase AID at the displaced G-rich ssDNA is the first step to generate DSBs for CSR[58,77].

Loss of R-loop processing pathways can interfere with transcription and chromatin accessibility. Indeed, cells use R-loops in several biological processes to regulate gene expression, DNA replication and DNA repair.

1.1.4.1   *Epigenetic regulation*

The first observations that R-loops were enriched in active mammalian promoters suggested that R-loops could be involved in gene regulation[34]. These regulatory R-loops are involved in gene regulation and chromatin accessibility by both activating or silencing gene expression[9,65].

Considering the nature of R-loops, it would be expected that they are enriched at loci with increased chromatin accessibility[37]. Indeed, the histone post-translational modification resemble those associated with transcription at promoters, such as H3K4me1, H3K4me3 and H3K27ac, and transcription elongation, like H3K36me3[34,50]. Moreover, R-loops may facilitate transcription by protecting the DNA from methylation and induce chromatin decondensation[78]. R-loops can inhibit DNA methylation by preventing binding of DNA (cytosine-5)-methyltransferases (DNMTs)[34,79]. Another example is GADD45A that binds to RNA:DNA hybrids and promotes DNA demethylation and gene expression through the formation of a regulatory R-loop at the TCF21 promoter[9,80]. In mouse embryonic stem cells, R-loops inhibit repressive chromatin-modifying enzymes and recruit activating chromatin-modifying complexes to promote differentiation genes[61].

However, R-loops can also be associated with histone modifications that promote heterochromatin assembly and chromatin compaction[51,81–83]. In yeast and human cells, R-loops are associated with a heterochromatin marker, H3S10 phosphorylation[82]. Besides reducing R-loop levels, overexpression of RNase H reduces H3S10 phosphorylation, promoting chromatin compaction.

Regulatory R-loops are also linked to antisense transcription through production of an antisense lncRNA, which in turn could recruit transcription regulators. For example, the hybrid formed by the antisense lncRNA of RASSF1 leads to the recruitment of Polycomb repressive complex 2 (PCR2), which inhibits the promoter of RASSF1A[84]. Taken together, these studies strongly suggest that R-loops are associated with specific histone modifications and can affect gene expression by modifying the chromatin.

1.1.4.2 Roles of R-loops in transcription regulation

Besides affecting the chromatin signature, R-loops facilitate transcription of some genes. In budding yeast, lncRNAs encoded in the GAL locus form R-loops that promote expression of GAL genes[85]. Moreover, in an antisense scenario, VIM antisense RNA 1 (VIM-AS1) activates the VIM gene, when VIM-AS1 transcription forms an R-loop that promotes recruitment of transcription activators[86]. On the other hand, R-loops may impair transcription. In *Arabidopsis thaliana*, an R-loop represses the promoter of the lncRNA COOLAIR. ATNDX binds to the R-loop and prevents the recruitment of R-loop resolving enzymes, thus affecting transcription of the COOLAIR lncRNA[87]. *In vitro*, R-loops block transcription[88], and genome-wide studies of transcription in R-loop-forming loci show that R-loops interfere with productive transcription[89–91]. It is unclear what distinguishes R-loops promoting transcription from those blocking transcription. Likely, regulatory R-loops have a quick turnover in cells and do not persist long enough to promote toxicity by stalling RNAP.

Besides promoter regions, R-loops are also enriched in G-rich RNAPII termination sites, where they are proposed to mediate efficient transcription termination[40,41,81,92]. One possibility for how R-loops affect transcription termination is through RNAPII pausing at 3' end. It has been proposed that R-loops may be essential for RNAPII to pause downstream of the poly(A) site[40]. For instance, a common feature of mammalian genes is the presence of G-rich sequences downstream of the poly(A) signal as well as potential G4-forming sequences at the 3'-UTR regions[93]. R-loops formed over G-rich pause sites are removed by Senataxin to allow degradation of 3' transcript by XRN2 and consequently RNAPII termination[40]. Consistently, *in vitro* T7 RNAP transcription has a tendency to pause at G-rich sequences, in which stable R-loops are formed[94]. Moreover, RNAPII is known to oscillate between elongation and backtracking and it has been shown that overextended R-loops can cause transcription termination of backtracked RNAPII[95]. Additionally, R-loops at some gene terminators may trigger antisense transcription, generating dsRNA that recruits the RNA interference machinery and establishes repressive heterochromatin through the H3K9me2 mark to reinforce RNAPII pausing[81]. However, a genome-wide study of R-loop mapping with histone modifications did not find correlations between other repressive chromatin marks and R-loop formation, instead it found a correlation with a mark of active transcription, H3K4 methylation[50]. R-loops might be responsible for the recruitment termination factors, such as PAF1C, although this mechanism remains unclear.

Interestingly, convergent genes were found to be associated with high levels of R-loops at terminators, suggesting that R-loops may be particularly important to prevent transcriptional read-through into adjacent genes[50]. Indeed, R-loops at 3' ends of genes are enriched specially

in gene-rich regions, where read-through RNA could be most problematic. Though on a genome-wide level, it is still unknown what the effects of accumulated R-loops are on transcription read-through.

In contrast to the studies discussed above, deficiency in R-loop resolving enzymes, including SETX, DHX9, DDX5 and XRN2, induces global R-loop accumulation and impair transcription termination at genes transcribed by RNAPII or RNAPIII[40,96–101]. However, it would be important to confirm whether these enzymes absence cause directly or indirectly R-loop accumulation in these sites.

Altogether, these studies emphasize how important it is to have the perfect R-loop balance, because either accumulation or no formation of R-loops impact biological processes in a similar manner.

1.1.5 R-loops and genome instability

Loss of genome integrity can lead to a host of complications, from cancers to neurological disorders. First evidences for R-loops as a source of genomic instability was initially provided by the THO mutants[22]. Yeast THO R-loop-forming mutants show transcription associated hyperrecombination phenotype, elevated chromosome and plasmid loss. Mutants affected in different steps of transcription and mRNA processing show similar transcription and genome instability patterns as those of THO mutants[102].

In principle, R-loops can cause DNA damage by multiple mechanisms. R-loops result in stretches of exposed ssDNA, which are chemically more unstable than dsDNA and accessible to the action of DNA damaging agents or certain enzymatic activities, thus likely more prone to mutagenesis, recombination and DSBs[103,104]. Indeed, mutations during transcription are mainly found in the non-transcribed DNA strand (the ssDNA in the R-loop)[105]. Analysis of nucleotide substitution rates along human genes suggest that promoter-associated R-loops lead to higher rates of cytosine deamination in the non-template strand near the 5' end of genes[106]. Human activation induced cytidine deaminase (AID),which is involved in CSR and somatic hypermutation in activated B-cells, target the ssDNA in R-loops[107]. In yeast, AID increases mutations in transcribed genes, when it is expressed in THO mutants that accumulate R-loops[108], thus confirming that mutagenic agents or enzymes induce mutagenesis in the exposed ssDNA of R-loops. Furthermore, DNA nicks, such as those generated by reactive oxygen species (ROS), can stabilize and promote R-loop formation[19]. Targeting ROS to actively transcribed genes or telomeres promotes G4 and R-loop formation, resulting in replication-independent damage at these sites[109–111]. Finally, R-loops could

promote mutagenic DNA replication by serving as primers to initiate non-canonical replication. This origin-independent replication can be observed in *E. coli*[112] and in the rDNA loci of yeast cells lacking Top1 and RNase H enzymes[113]. Nevertheless, more evidence need to be provided to understand the accuracy of replication primed by R-loops.

R-loops were considered drivers of DNA damage and genome instability based on observations made in conditions where hybrids prevention or removal is compromised. Recent work has demonstrated that the recombination promoted by R-loops is essential for correct DNA repair at breaks and at dysfunctional telomeres, as well for to ensure proper centromere function and chromosome segregation. Overall, R-loops have both beneficial and harmful impacts on the cells.

1.1.5.1 R-loops in DNA repair

DSBs are one of the most toxic forms of DNA damage and are a primary driver of genome instability. Cells have a variety of strategies to repair DSBs, including HR and NHEJ[114,115]. Early observations showed that unscheduled R-loops can cause DSB formation, but recently DSBs have also been shown to cause high R-loop levels. Indeed, evidence shows that, while hybrids deregulation can impede DSB repair, RNA:DNA hybrids are important intermediates to promote DNA repair. However, it is important to emphasize that hybrids are not a feature of all DSBs, since it has been shown site-specific breaks that are not associated with RNA:DNA hybrids[116].

1.1.5.1.1 R-loops block DNA repair

On one hand, hybrids have been proposed to hinder the repair of breaks once formed. Recently, in budding yeast, it was shown that hybrids accumulate at transcriptionally active breaks[117]. When R-loop resolution enzymes were lost, transcription impaired recombination and increase genomic instability at induced DSBs, which could be rescued by RNase H1 overexpression. This suggested that hybrids can interfere with DSB repair processes.

One possibility by which R-loops inhibit DNA repair by affecting end resection processes. EXO1 and DNA 2 promote long-range resection, in a process requiring the BLM helicases[118,119]. The loss of R-loop resolution factors, such as DDX5, DHX9, and HNRNPD*/SAF-A, leads to hybrids accumulation and impaired end resection at induced DSBs[101,120–122]. RNA:DNA hybrids accumulation decreases RPA and RAD51 loading and increases genomic instability. Therefore, accumulated hybrids at break sites may block

nucleases from binding to the DNA ends, preventing resection. They may also mask ssDNA recognition by RPA, thereby inhibiting RAD51 nucleofilament formation and homology search. Indeed, RNase H2 and SETX have also been implicated in removing hybrids at DSB sites to allow RAD51 loading[123,124]. Additionally, *in vitro* experiments showed that long hybrids at DNA ends cannot be unwound by BLM and inhibit EXO1 activity[125].

In budding yeast, accumulated R-loops caused by depletion of RNase H enzymes and Sen1 leads to a DSB downstream of the R-loop, with an hybrid in one end of the break[126,127]. Resection in the end-containing a hybrid was inhibited, while it proceeded normally in the hybrid-free end, indicating that the hybrid needs to be removed for resection to occur and allow efficient repair. Consistent with this, RNase H1/H2 overexpression in *S. pombe* enhanced resection[128]. Taken together, these studies suggest that hybrid removal is a critical event of DSB repair and if hybrids become deregulated, they can interfere with resection, leading to defects in repair and genome instability.

1.1.5.1.2 R-loops facilitate DNA repair

In contrast to the idea that R-loops hinder DNA repair is their essential role in promoting DNA damage response[129]. RNA:DNA hybrids have been proposed to be required for both HR and NHEJ, and their presence or absence may directly influence the DSB repair process.

In *S. cerevisiae*, Sen1 loss increases hybrid levels in DSB sites, which leads to increase KU occupancy at break ends, activating an alternative resection pathway that promotes classical and microhomology-mediated NHEJ[130]. In human cells, KU and RBM14 facilitates RNAPII and 53BP1 recruitment to DSBs, as well as hybrid formation[131–133], suggesting that KU promotes formation and stabilization of hybrids at DSBs, creating a positive feedback loop that promotes NHEJ. In agreement, loss of other R-loop resolution factors, such as DDX1 or EXOSC10, increased NHEJ efficiency while decreasing HR efficiency[121,134,135]. These observations suggest that stabilized hybrids in DSBs favor NHEJ by helping to both restrict resection and recruit NHEJ factors.

On the other hand, R-loops and hybrids have also been reported to direct cells toward HR. In *S. pombe*, it was shown that hybrids are necessary transiently at DSBs to prevent excessive resection and modulate RPA loading for efficient HR[128]. This study suggest that hybrid formation and turnover is essential to control resection and have HR-mediated DSB repair. Furthermore, a novel HR reporter in mammalian cells shows that HR was stimulated by local transcription[28]. HR was dependent on RAD51 associated protein 1(RAD51AP1) to promote invasion of ssRNA into template DNA. Concurrent invasion of ssDNA at the same

region would also occur, forming a DR-loop. As a result, DR-loops enhanced RAD51-mediated invasion of the template DNA by ssDNA to generate a D-loop, an important intermediate in DSB repair. A recent study also shows that RNAPIII produces hybrids important for the regulation of DNA repair[136]. RNAPIII is recruited to DSBs, where it transcribes *de novo* hybrids in an MRE11 and CtIP-dependent manner. Loss of RNAPIII activity and hybrid formation impairs RPA and RAD51 loading, suggesting that hybrid formation is necessary for HR. Additionally, the hybrid mediated by RNAPII transcription may protect the 3' DNA strand from an exonucleolytic attack, preventing loss of genetic material.

R-loops also promote an alternative HR pathway know as transcription-associated homologous recombination repair (TA-HRR)[137]. In TA-HRR, R-loops form near break sites in actively transcribed loci and are recognized by RAD52, which promotes hybrid removal by XPG to generate ssDNA, as well as recruitment of HR factor, such as BRCA1. Similarly, RAD52 is recruited to ROS-stabilized hybrids by the TC-NER factor CSB[110,111]. RAD52 recruits RAD51 to promote strand invasion and HR. Together, these suggest that RNAPII or RNAPIII-dependent *de novo* hybrid formation promotes recruitment of downstream repair factors to break sites, to ensure faithful DNA repair. Overall, RNA:DNA hybrids at DSBs exemplify the importance of their transient nature to regulate a biological process without errors.

1.1.5.2 R-loops and replication stress

During S phase, the transcription and replication machineries need to access the same template, which might lead to collisions in certain situations and/or specific genomic regions[138,139]. For example, active replication origins are usually present in promoter regions of highly expressed human genes[140,141] and long human genes require more than one cell cycle to be fully transcribed[142]. Therefore, transcription-replication collisions (TRCs) during S phase are unavoidable. Additionally, TRCs and replication fork stalling can be caused by paused, stalled and backtracked RNA polymerases, which may result in chromosomal breakages and rearrangements[138,143]. Notably, numerous studies demonstrated TRCs are one of the main sources of R-loop-induced replication stress and DNA damage[22,142,144–146]. Indeed, the absence of R-loop-resolution factors cause replication stress and DNA damage in S phase cells, suggesting that R-loops interfere with fork progression[22,127]. Early studies also demonstrated that replication forks slow at sites with an induced R-loop and RNase H1 overexpression restores normal fork progression[23,145,147].

Replication forks can encounter transcription machinery when moving in the same direction, known as co-directional collision (CD), or when moving in opposite directions

towards each other, defined as head-on collision (HO). In mammalian cells, an episome containing an R-loop-forming region exacerbated the damage associated with both HO and CD conflicts upon transcription induction compared with an episome without an R-loop[90]. This suggests that the presence of R-loops is more of a problem than transcription itself. Moreover, although both types of encounters can interfere with progression of the replication fork, they have distinct outcomes, with HO conflicts being more detrimental. HO encounters activate ATR, while CD conflicts induced ATM damage signaling pathway.

Interestingly, R-loops levels are affected by TRC orientation. In particular, HO collisions enhance hybrid formation , while CD collisions do not[90]. One possibility for hybrid formation in a HO collision could be due to topological stress created in this context. Converging transcription and replication machineries leads to accumulation of positive supercoiling ahead of the respective machineries and negative supercoiling accumulate behind them. The cumulative positive supercoiling between the machineries may cause the fork to stall, when not compensated by topoisomerase activity, while the negative supercoiling favors R-loop formation behind the machineries. Similarly, in *Bacillus subtilis*, resolution of positive supercoiling formed in a HO encounter by topoisomerase promotes R-loop formation by allowing diffusion of negative supercoiling behind RNAPII. By contrast, upon CD conflicts topological stress is relieved[91]. Moreover, in CD collisions, the presence of a 5' flap in the RNA:DNA hybrid allows CMG to unwind the hybrid, which could explain the reduction of R-loop levels in CD conflicts[148].

A recent study in *S. cerevisiae*, demonstrated that naked R-loops containing a G4 on the non-template strand, but lacking proteins bound including RNAP, could impact replication fork progression[148]. At both HO and CD R-loops, leading strand replication forks pause but can bypass these structures and restart DNA replication on the leading strand downstream of the R-loop. In particular, leading strand hybrids impede fork progression at CD conflicts, but restart of leading strand replication is not a direct consequence of RNA:DNA duplexes, rather it is dependent on transcription to resume replication. However, at HO conflicts, RNA:DNA hybrids on the lagging strand are not the direct cause of fork stalling and uncoupling. Instead, the G4 formed in the leading strand as a result of R-loop formation on the lagging strand causes fork stalling or uncoupling. Thus, G4 formation in the displaced ssDNA of an R-loop may be important determinant of an R-loop impact on the replication fork[148,149]. Similarly, studies in a bacterial replication system show that naked R-loops can be bypassed in both orientations. In addition, they demonstrate that a RNAP bound R-loop poses a greater threat to fork progression than a naked R-loop, particularly in the HO orientation[150]. In yeast, a dCas9-associated R-loop also blocks the fork in both orientations[149]. Finally, although both RNA:DNA hybrids and G4s can inhibit lagging strand synthesis at HO collisions causing gaps in nascent

lagging strands, R-loop transcripts can prime leading strand restart by Pol α after CD collisions between R-loop and replisome[148].

1.1.5.3 Telomeric R-loops

Telomeres are prone to R-loop formation due to their high GC content and possible G4 structures formed in the displaced strand. Telomere repeat-containing RNA (TERRA) is a lncRNA transcribed at telomeres from the C-rich strand and it has a strong tendency to form R-loops both *in cis* and *in trans* at chromosome ends[151–156]. Telomeric R-loops are transient and cell cycle-dependent, being formed in early S-phase and removed by the RNase H enzymes in late S-phase, when telomere replication happens. However, telomeric R-loop levels inversely correlate with telomere length[157,158]. At long telomeres, R-loop levels are kept low, while at short telomeres, R-loops are stabilized and RNase H enzymes do not localize efficiently to the telomeres to remove them. In yeast, stabilized R-loops at short telomeres promote recombination with the long telomeres to prevent premature senescence[157]. On the other hand, R-loops removal from short telomeres by overexpressing RNase H1 impairs recombination, resulting in the accelerated senescence[76]. Telomeric R-loops are essential to maintain telomeres length in human cancer cells that depend on a recombination mechanism know as alternative lengthening of telomeres (ALT)[36,155]. Therefore, similar to DSBs, R-loops at short telomeres drive recombination. However, at telomeres the 5' end hybridizes with TERRA, while at DSBs it is the 3' overhang that pairs with RNA. Still, in both telomeres and DSBs the presence of an R-loop promotes RAD51 loading. Additionally, similar to DSBs and other biological processes, a balanced and controlled R-loop levels are essential for proper telomere maintenance.

1.1.5.4 R-loops in mitochondria

The impact of R-loops on genome stability extends to the mitochondrial DNA (mtDNA), where R-loops are important intermediates in mtDNA replication[57,159]. Before replication, R-loops form on the origin of replication of the heavy strand. These R-loops are processed by RNase H1 to generate the RNA primer to start DNA synthesis[160,161]. Thus, regulation of mitochondrial R-loops is essential to ensure that the mtDNA is properly replicated and that mitochondrial levels are maintained. Indeed, loss of RNase H1 causes R-loop accumulation and elevated levels of mitochondrial recombination. Moreover, R-loop accumulation decreases mitochondrial copy number and causes embryonic lethality [160,162]. Accumulation of ROS can

generate 8-oxoguanine, which may inhibit RNase H1 recruitment to mitochondrial origin of replication, mimicking RNase H1 deficiency[163]. Although the molecular mechanism by which RNase H1 loss destabilizes the mitochondrial genome is not clear, a recent study suggests that mtDNA replication initiation is promiscuous when R-loops accumulate[161].

1.1.6 R-loop regulation

Given the potential genomic instability caused by persistent R-loops and the necessity for regulatory R-loops to be transient, cells have evolved several mechanisms that regulate R-loops genome-wide. One way to regulate R-loop functions and prevent deleterious consequences is by restricting R-loop formation spatially and temporally. For instance, R-loops that control gene expression could be spatially separated from the TSSs of the genes they regulate. R-loop levels also could change during the cells cycle and/or R-loop formation being restricted to a specific cell cycle phase. In the case of TCF21 gene, expression of an antisense lncRNA, R-loop accumulation and TCF21 expression happen in a sequential manner during cell cycle[80], preventing detrimental effects of R-loops while allowing processes to be initiated through R-loops. Additionally, some enzymes that process R-loops like RNase H2 are tightly cell cycle regulated, which would confer a temporal R-loop regulation[76].

Another possible R-loop regulation method is through preventing their formation. In this case, multiple factors and pathways can be involved, including RNA processing and export machinery, which suppress hybrid formation by sequestering nascent RNA from the template. One hypothesis for this separation to happen is by targeting the RNA to the nuclear pores. Indeed, increased physical distance to nuclear pores increases R-loop levels in yeast, likely because transcripts accumulate in the nucleoplasm and re-anneal with their DNA templates[164]. In addition, proper packaging of nascent RNAs and coordination of transcription also prevent unscheduled R-loops. For this reason, mutations in RNA binding proteins (RBP) and transcription factors increase R-loop levels in yeast[22,82,165]. RBPs and other factors may act as RNA chaperones that prevent R-loop formation by promoting an optimal ribonucleoprotein incapable of hybridizing with its DNA template, as is the case for THO[4] (Figure 1).

During transcription, negative DNA supercoiling is generated behind the RNAP, which potentially opens up the DNA template and facilitates nascent RNA association to DNA strands. Topoisomerases regulate this negative supercoiling, thus preventing R-loop formation[26]. Additionally, modifications of the RNA strand of the hybrid by RNA methyltransferases METTL3 or TRDMT1 helps regulate R-loop formation (Figure 1)[166–168]. Finally, proper assembly of chromatin would also limit RNA:DNA hybrid formation. For

instance, histone deacetylation coordinates chromatin compaction through its interaction with human THO complex and prevent unscheduled R-loop formation during transcription[169].

Once RNA:DNA hybrids are formed, a host of additional proteins unwind or degrade the RNA within the R-loop to regulate R-loop levels. Cells contain numerous non-redundant R-loop resolving factors (Figure 1). One possible explanation is that each factor acts on a specific subset of RNA:DNA hybrids. Moreover, it is interesting to think that the limited potency of these factors may suggest that each factor is regulated spatially or temporally to constrain its ability to remove R-loops[170].
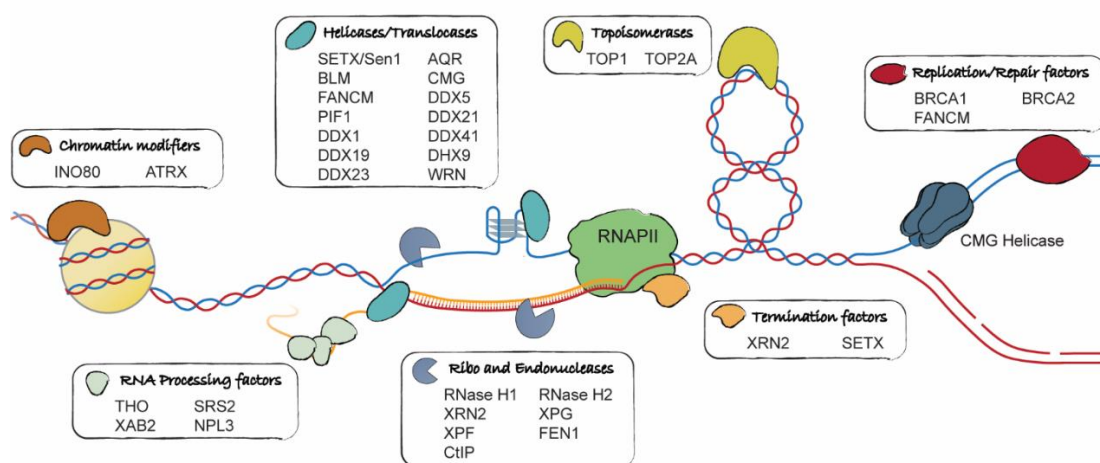
Figure 1 – Factors that regulate R-loop levels. Factors that prevent the formation of R-loops or play a role in their resolution as reference throughout the text. The red and blue lines are DNA, and RNA is depicted by yellow line. The structure on the ssDNA strand of the R-loop represents a G-quadruplex. Adapted from Brickner et al[42].

1.1.6.1 R-loop resolution

R-loops levels can be controlled by resolution of their three-stranded structure. One possibility for this to happen is through RNA:DNA helicases, which unwind RNA:DNA hybrids or limit their formation. Several studies have been identifying an increasing number of RNA:DNA helicases that lead to an accumulation of R-loops, when depleted from cells. These include Senataxin, FANCM (Mph1 in yeast), AQR, DDX19, DDX21, DDX23, DDX1, DDX5 (Dbp2 in yeast), BLM (Sgs1 in yeast), DDX41, yeast Pif1 and others[40,96,135,171–182] (Figure 1). Although loss of these helicases causes R-loop accumulation, it has not been confirmed whether all these helicases unwind hybrids *in vitro* and *in vivo*. First, *in vitro* assays are required to access the properties of these helicases and their preference for hybrids over other molecules. Additionally, *in vivo* experiments that would include overexpressing the helicases

to show a reduction of hybrids and any R-loop associated phenotype are essential to confirm a direct effect on R-loops *in vivo*. However, this may be challenging as these helicases typically are essential components in other biological processes or require other factors to act upon R-loops. For instance, RNAPII pausing at R-loop-rich regions triggers a signaling cascade that leads to the phosphorylation of DDX23 RNA helicase and its recruitment to pause sites to resolve R-loops[178]. Dbp2 regulates R-loop levels likely by cooperating with Sen1[173]. DDX19 is a nucleopore-associated mRNA export factor that is able to unwind RNA:DNA hybrids *in vitro* and acts on R-loops formed upon replication stress or DNA damage. Some helicases, such as Pif1, BLM and Sgs1, DDX9, or FANCM, it was shown to have *in vitro* RNA:DNA unwinding activity. However, other helicases, such as RECQL5, may act differently. R-loops accumulate in RECQL5-depleted cells, but no helicase activity is required for R-loop suppression[183]. Other study suggested that RECQL5 promotes TOP1 activity[184], which may explain its role in R-loop suppression.

One of the best characterized R-loop-resolving helicases is Senataxin, encoded by SETX in human or Sen1 in yeast. Sen1 was initially identified as a DNA and RNA helicase with 5'-3' RNA:DNA unwinding activity *in vitro*[185]. Later it was shown *in vivo* that Sen1 helicase inactive *sen1-1* mutants accumulate R-loops at highly expressed loci[24] and inactivation of human SETX leads to increased R-loops levels at transcription termination pause sites[40]. Together with Senataxin essential role in transcription termination of RNAPII and RNPIII genes, studies suggest that Sen1/SETX unwinds R-loops during transcription termination[99,186–188]. Additionally, studies also demonstrated Senataxin is recruited to DSBs and replication forks to resolve hybrids and ensure efficient DSB repair and promote fork progression, respectively[124,130,189].

1.1.6.2 R-loop degradation

R-loops can be removed by degradation of the RNA strand. This important function is carried out by RNase H enzymes[190,191]. RNases H are metal-dependent endonucleases that hydrolyze RNA residues in RNA:DNA heteroduplexes with some sequence preference but not strict sequence specificity[192,193]. RNase H enzymes are divided into two main classes: RNase H1 and H2. These classes have common structural features of the catalytic domain but different ranges of substrates that they cleave.

The presence of two distinct, but evolutionary conserved RNase H proteins may suggest they have some non-overlapping function. Indeed, RNase H1 and RNase H2 display different cleavage pattern on hybrid substrates[194]. However, in yeast, R-loop levels increase only when

both RNases H are inactivated[195], implying they can replace each other to degrade the same R-loops. Interestingly, cells lacking RNase H enzymes have increased R-loop formation, although the presence of several RNA:DNA helicases. Once again, this suggests that the different R-loop removal pathways may target different subset of RNA:DNA hybrids. A genome-wide study suggest that RNase H enzymes target R-loops formed in highly transcribed and longer genes, such as retrotransposons and other genes transcribed by RNAPI and RNAPIII, specifically tRNA genes[27]. In addition, RNase H1 deletion also caused R-loop accumulation in mitochondria genes[33].

Although in human the absence of one of the RNase H enzymes lead to severe diseases, RNase H enzymes are not essential in yeast, which could be explained by the existence of compensatory pathways and the different impact that R-loop accumulation has in yeast versus other higher eukaryotes.

1.1.6.2.1 RNase H2

RNase H2 is a trimeric complex composed by one subunit with catalytic activity, RNase H201 (or RNase H2A in human), and two auxiliary subunits, RNase H202 and RNase H203 (or RNase H2B and RNase H2C, respectively)[191,196–198]. RNase H2 has the ability to degrade the RNA in R-loops and it is addionally able to perform ribonucleotide excision repair (RER). This is possible due to RNase H2 only requiring to recognize 2'-OH group of one ribose molecule for its activity[196,199,200]. In the absence of RNase H2 function, RER is defective and topoisomerase enzymes can aberrantly process rNMPs, which leads to deletions and formation of DSBs[201–203]. Furthermore, RNase H2 hydrolyzes Okazaki fragments, leaving a single ribonucleotide of the RNA primer attached to the newly synthesized DNA[204]. While RNase H1 deletion has mutation rates comparable with wild type cells, the loss of RNaseH2 results in increased genome instability and mutation rates[205]. This is consistent with RNase H2 accounting for the majority of the RNase H activity in cells and having dual role in R-loop and rNMP removal[205–207].

In yeast, a recent study demonstrated that restriction of RNase H2 activity to the G2/M-phase of the cell cycle is sufficient to rescue R-loop associated damage, suggesting a major role for RNase H2 activity post-replicatively, while RNase H1 seems to respond in a cell cycle-independent manner to scenarios that accumulate R-loops[76]. Moreover, RNase H2 exhibit important differences in biochemical properties compared with RNase H1. The preferred site of cleavage is on the 5' site of RNA:DNA hybrid. In the case of RNase H2, substrate specificity is relaxed in the presence of Mn^{2+} ions[208]. The subunit RNase H2B/RNase H202 interacts with proliferating cell nuclear antigen (PCNA) via a PCNA-interacting peptide (PIP) at its C-

terminus[209]. PCNA is a protein essential for DNA replication and repair, thus interaction with PCNA supports the role of RNase H2 in replication and repair processes. Overall, the auxiliary subunits, RNase H2B/RNase H202 and RNase H2C/RNase H203, may provide a platform for the assembly of the complete enzyme, processivity and interaction with PCNA and other cellular proteins[208,209].

In humans, mutations in any of the components of human RNase H2 cause Aicardi-Goutières syndrome (AGS), which leads to neurological dysfunction and psychomotor retardation[198]. Recently it was shown that acute immune response in AGS patients is likely due to DNA hypomethylation and increased R-loop levels[210].

1.1.6.2.2 RNase H1

RNase H1 is a monomeric enzyme that degrades the RNA moiety of RNA:DNA hybrids by interacting with the 2'-OH group of four consecutive ribose molecules. In higher eukaryotes, RNase H1 is important for mtDNA replication[161], and indeed RNase H1 null mouse embryos arrest development because they are incapable of amplifying mtDNA[211]. In humans, no genetic disease has been identified to be associated with RNH1 mutations, likely due to this enzyme being essential for embryonic development. On the other hand, yeast RNase H1 is not required for mtDNA replication, likely due to replication in yeast mitochondria occurring by a different mechanism than found in other organisms[206,212].

Excess RNase H1 is often toxic for vertebrate cells[23,169,213,214]; however, the molecular reasons for this effect remains unclear. One hypothesis for this toxicity may be due to a dose-dependent indirect perturbation on the nascent transcriptome and proteome[65]. For instance, in human cells, the overexpression of RNase H1 was shown to affect the transcription of antisense RNA[215] and the stability of some DNA damage response factors[216], while in mouse embryonic stem cells it was shown to derepress Polycomb-repressed genes[217]. On the contrary, in yeast RNase H1 overexpression does not seem to affect transcriptome and proteome changes, although more studies are required.

All eukaryotic RNase H1 have highly conserved regions at their N- and C-termini separated by a variable sequence[191]. The conserved N-terminal region contains the hybrid binding (HB) domain, while the catalytic domain (H-domain) is located in the C-terminus (Figure 2). Most eukaryotic cells have a single copy of HB domain, but *S. cerevisiae* contains two similar HB domains[218]. Human RNase H1 also contains a mitochondrial targeting sequence (MTS) upstream of the HB domain, and the protein is expressed in two isoforms that are localized to the nucleus and mitochondria[211]. On the other hand, yeast does not have MTS.
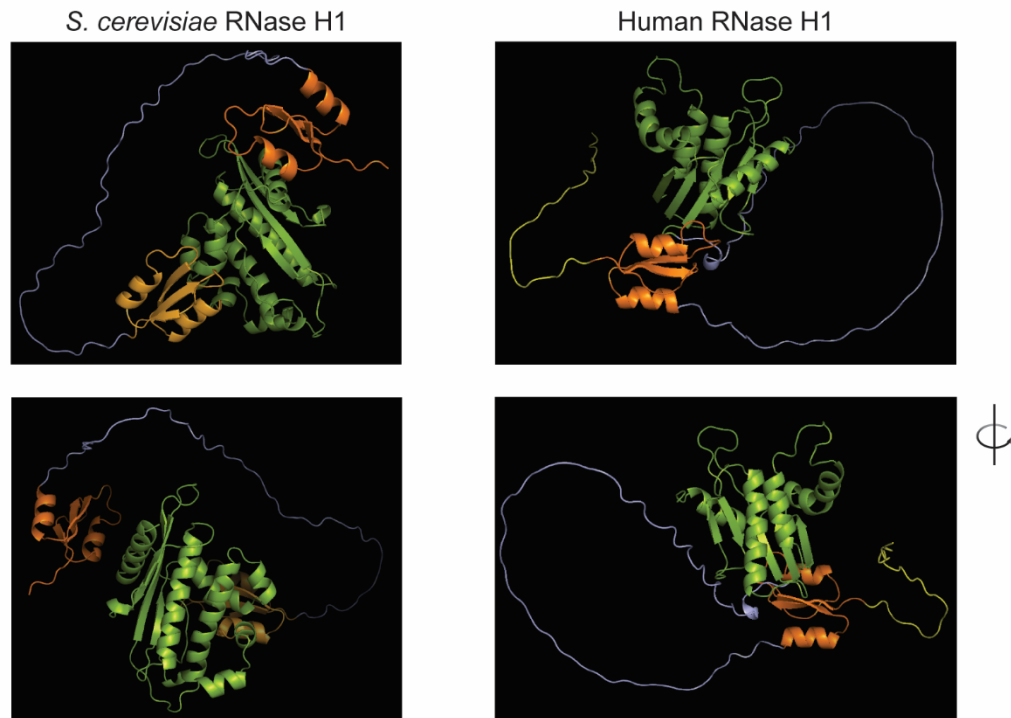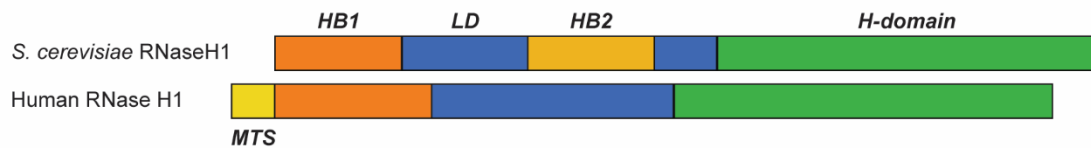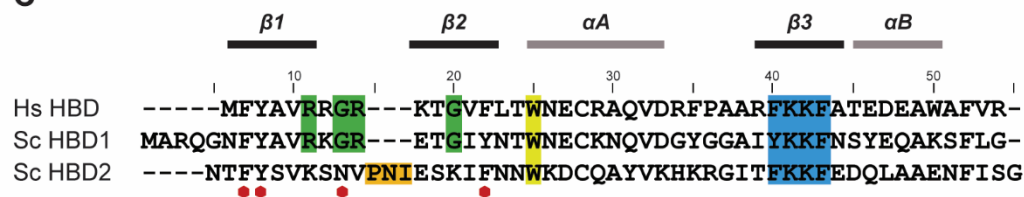
Figure 2 – RNase H1 structure in eukaryotes. (A) Tertiary structure of S. cerevisiae and human RNase H1 predicted by AlphaFold. Mitochondrial targeting sequence (MTS) is depicted in yellow, hybrid binding domains (HB) are represented in orange (both dark and light in case of S. cerevisiae RNase H1), in blue is the connection or linker domain (LD), and the RNase H or catalytic domain (H-domain) is green. **(B)** RNase H1 organization found in S. cerevisiae and human and consists of HB domain, LD and H-domain. Human RNase H1 contains additionally a MTS, while yeast contains two HB domains. **(C)** Alignment of the HB domains of S. cerevisiae (Sc) and human (Hs) RNase H1. The black bars represent the β strands, and the grey bars are α helices of human HB domain. In blue is the FKKF motif, which contains K59 and K60 amino acids in human. Yellow highlights amino acid W43 in human. Green marks conserved amino acids (R11, G13, R14 and G17) in HB1 domain of S. cerevisiae discussed in the text. Red dots are amino acids (F7, Y8, G17 and Y19) conserved in HB1 domain of S. cerevisiae discussed in the text. Orange highlight corresponds to PNI amino acids mentioned in the text.

The HB domain was first described in *S. cerevisiae* RNase H1 as a motif that binds dsRNA[219]. However, later it was shown that HB domain of human RNase H1 has 25-fold preference for RNA:DNA (K_D ~0.2µM) compared with the same sequence of dsRNA(K_D ~4.9µM)[72,180]. Sequence conservation and NMR structure of the HB domain allowed to understand which amino acids were essential for interaction with substrates[218–220]. In addition, mutations of highly conserved amino acids confirmed their function[221].

HB domain comprises a three-stranded antiparallel β-sheet and two α-helices, which are located onto opposite faces of the β-sheet (Figure 2A, C)[220]. From the structures of *S. cerevisiae* and human RNase H1 HB domains, it is apparent that several of the conserved hydrophobic amino acids form the core of the domain and define its overall dimensions[72,220]. The HB domain interacts with two consecutive 2'-OH groups of the RNA strand through the loop between αA and β3. A shallow positively charged groove containing the conserved FKKF motif and Y29 residue interact with the DNA strand (Figure 2C). In this case, two of the conserved residues, W43 and F58 in human, contact two deoxyribose rings of the DNA strand. If RNA is present instead, the 2'-OH groups of the riboses clash with the two aromatic amino acids, which explains HB domain preference for RNA:DNA hybrids over RNA:RNA duplexes. The most positively charged surface of the HB domain is not directly involved in substrate binding, but rather acts as a first point of contact directing the hybrid to the formation of specific interactions with the HB domain[222]. Once properly engaged, the HB domain enhances the RNase H activity of the protein[72,221]. Indeed, mutagenesis or low salt concentrations that reduce specific interactions leads to lower catalytic activity of the enzyme[219,221]. For instance, at low salt concentrations, nonspecific charge-charge interaction occur between HB domain and other substrates besides RNA:DNA hybrids, such as dsRNA, ssDNA and dsDNA[219]. On the other hand, nonspecific interactions are overcome at high $MgCl_2$ or physiological salt concentrations, allowing optimal RNase H activity. Moreover, weaker binding to hybrids caused by deletion of the HB domain or mutations in the HB domain amino acids involved in interacting with the DNA strand of RNA:DNA (W43, K59 and K60 in human RNase H1) impairs RNase H activity and processivity, indicating that binding through the HB domain and catalysis by the RNase H domain are related processes[72,221]. This data allowed to propose a model where the HB domain may first approach the RNA:DNA substrate facilitating interactions and hydrolysis by the RNase H domain. After catalysis and dissociation of the RNase H domain, the HB domain remains connected to the substrate to promote new interactions of the catalytic domain with nearby sites of the hybrid, resulting in several rounds of hydrolysis per HB domain binding event and short RNA fragments that can easily dissociate from DNA[72,191].

Both HB domains of *S. cerevisiae* are very similar with an identical 3D structure. However, two important features in their sequence distinguish them: highly conserved R11, G13, R14 and G17 amino acids in HB1 domain are replaced by N, V and K, respectively, in HB2 domain, and three amino acids, P, N and I are present between β1 and β2 strands only on HB2 domain (Figure 2C). Early *in vitro* studies with HB domains of *S. cerevisiae* suggest that RNase H1 binds to RNA:DNA hybrids or dsRNA through its first HB domain. In particular, HB1 domain binds to substrates through its most N-terminal region and mutations on F7, Y8, G17 and Y19 are enough to abolish interaction with dsRNA and RNA:DNA duplexes(Figure 2C)[218,219]. The second copy of the HB domain does not form stable complexes with dsRNA or hybrids when alone, due to the presence of the extra P, N and I amino acids (Figure 2C). However, HB2 domain enhances the overall binding of RNase H1 when in combination with the first HB domain[218]. The HB domain from human and the first HB domain of *S. cerevisiae* have two highly conserved arginine residues important to stabilize RNA:DNA hybrids in the loop between αA and β3. Since yeast HB2 domain does not have arginine residues in these two positions, it is possible that when the first HB domain interacts with the hybrid, the second HB domain may be able to contribute to binding by assuming the correct structure when in close proximity to the RNA:DNA, like an induced fit[191]. Considering that W43, K59 and K60 amino acids from human RNase H1 are conserved in both HB domains of yeast (W21/126, K38/142 and K39/143) and the structure of RNase H shows identical position of these amino acids in both organisms, it would be expected that mutations in these residues would confer a defective HB domain in yeast[221].

The loop between β1 and β2 has a different position in the HB domain complex with RNA:DNA compared with the HB domain alone, suggesting that substrate binding changes conformation of the domain[222]. Interestingly, other studies showed that modifications that reduce the flexibility of the DNA strand of RNA:DNA duplexes make the hybrids resistant to cleavage by RNase H1[223,224]. In addition to RNase H1 structure, this result confirms RNase H1 recognizes both the RNA and DNA strands. Furthermore, it suggests that both protein and substrate need a certain level of mobility in order for RNA:DNA hybrids degradation to happen. Thus it is possible that less flexible DNA forms are more resistant to R-loop degradation by RNase H1.

1.1.6.2.2.2 CATALYTIC DOMAIN

The C-terminal catalytic domain of RNase H1 comprises around 150 amino acids and is responsible for catalyzing hydrolysis of the phosphodiester bond of the RNA strand in a RNA:DNA hybrid. Cleavage requires at least four consecutive ribonucleotides[191,225]. This was

confirmed by the structure of RNase H domain in complex with RNA:DNA duplex[226], which shows that the hybrid fits in two shallow grooves of the enzyme, the catalytic site and the basic protrusion, and the 2'-OH of four consecutive riboses interact with the catalytic site. The DNA sits in the basic protrusion, and one of the phosphates fits into a pocket, which requires a distortion of the DNA backbone. This conformational change in the DNA cannot be accomplished by RNA, thus confirming that both strands contribute to the specificity of RNase H1. Indeed, RNase H specifically recognizes the A form RNA strand and the B form DNA strand[226]. Since the sequence of the DNA and/or proximity to an end may influence its malleability, it may result in a variety of sequences preferences related with DNA flexibility[223]. Moreover, the basic character of the basic protrusion is important to attract the substrate to the enzyme allowing binding in the proper manner. Hence, engagement of the hybrid in both HB and catalytic domains requires the presence of basic residues to attract the substrate. Overall, although RNase H1 binds to dsRNA, it has a stronger preference towards RNA:DNA hybrids and it is unable to cleave dsRNA substrates, strengthening the role of RNase H1 only on RNA:DNA hybrids.

RNase H enzymes require divalent metal ions for catalysis. RNase H1 has a preference for Mg^{2+}, with optimum condition of 1mM Mg^{2+} and pH 7-8[67,227]. However, hydrolysis can be also supported by Mn^{2+}[228,229] and it is inhibited by the presence of Ca^{2+}[222,227,230].

1.1.6.2.2.3 LINKER DOMAIN

All eukaryotic RNase H1 enzymes contain a variable sequence connecting the N- and C-terminal regions, known as connection or linker domain. Due to this domain containing differences in length and amino acid sequence between organisms, it has been proposed that it is important mainly to provide flexibility and allow the HB and RNase H domains to move freely in and around the substrate. For instance, the connection domain of human RNase H1 is 64 amino acids and if fully extended and completely flexible, it would allow the HB domain and RNase H domain to be separated by >20nm, this would correspond to 60 nucleotides separating HB domain and catalytic domain. *S. cerevisiae* contains two linker domains, a longer one (53 amino acids) connecting HB1 and HB2 domains and a shorter one (26 amino acids) between HB2 domain and the catalytic domain. Although little is known about the linker domain functions, variability in the length and composition of the linker domain may suggest that this region interacts with proteins or complexes[191].

Although RNase H1 interactome has not been explored extensively, a few proteins are known to interact with RNase H1. Three of the known RNase H1 interactors are the subunits of replication protein A (RPA)[231,232]. In human cells, RNase H1 was identified as an interactor of RPA-coated ssDNA. However, DNA and RNA do not seem to be essential for this interaction. RPA interacts specifically with a conserved basic ridge in the HB domain of RNase H1. RPA promotes association and activity of RNase H1 on R-loops *in vitro*. *In vivo*, overexpression of a catalytically active RNase H1 mutant that cannot bind RPA could not suppress genome instability in several contexts, unlike the wild-type RNase H1[233]. Thus, RPA appears to be a regulator of RNase H1 in human cells. In yeast, RNase H1 regulation and interaction with RPA remains unclear. Although little is known of RNase H1 regulation, overexpression of RNase H1 can suppress R-loop accumulation in cells without others R-loop resolving factors, suggesting the presence of dosage-dependent mechanism to regulate RNase H1.

1.2 Scope of the thesis

Several studies show the relevance of RNA:DNA hybrids in multiple biological processes. Hybrids were shown not only to cause genome instability, but also to participate in physiological processes in the cells, such as DSB repair. Although R-loops may generate DSBs, it is also involved in promoting their efficient repair[128]. A timely removal of R-loops is essential for DSB repair completion. One of the proteins essential for tight regulation of RNA:DNA hybrids is RNase H1.

Genome wide studies using defective RNase H1 to map R-loops may suggest a bias to a subset of R-loops compared with other methods to measure RNA:DNA hybrids. Little is known why RNase H deletion causes increase in a specific subset of R-loops and what causes such preference. Structure of RNase H1 produced a lot of information of how the enzyme binds and degrades RNA:DNA hybrids[192]. In human cells, RPA was identified as a regulator of RNase H1[233]. However little is known if RPA is enough to explain RNase H1 preference for a subset of R-loops.

This work aims to gain a deeper understanding of RNase H1 regulation in *S. cerevisiae*. We confirm where and when RNase H1 binds to R-loops. To study how RNase H1 may be regulated, we created multiple mutants and evaluated their affinity for R-loops *in vivo* and *in vitro*. Furthermore, we were interested in identifying factors that could impact R-loop removal

by RNase H1 and we expanded the understanding of RNase H1 and RPA interaction. Our data may provide new insights into how RNase H1 binds and remove R-loops.

2 Results

2.1 RNase H1 binds to accumulated R-loops

Previous genome-wide studies showed that *rnh1 rnh201* double mutants have high R-loop loads specifically in RNPIII transcribed genes, such as tRNAs and 5S rDNA, retrotransposons and in other highly transcribed and longer genes, suggesting that RNase H enzymes target R-loops in this subset of genes[27,33]. To further understand where RNase H1 binds in the genome, we used a galactose-inducible system and overexpressed 6xHA-tagged RNase H1 when cells were in exponential phase (Figure 3A, B). Subsequently, we performed R-ChIP in wild-type (WT) and *rnh1 rnh201* mutant cells that express an exogenous catalytically inactive RNase H1 (D193N), which binds to RNA-DNA hybrids but does not resolve them (Figure 4). Additionally, we performed DRIP experiments using the S9.6 monoclonal antibody that detects RNA-DNA hybrids, in wild-type (WT) and *rnh1 rnh201* mutant cells that either overexpress RNase H1 (RNH1) or express the catalytic-dead version of it (Figure 5). Our approach to study RNase H1 binding and R-loop removal involved a variety of different loci, including two tRNAs, SUF2 and SUF11, 5S rDNA, 18S rDNA, RPL15a and Actin. Both tRNAs and 5S rDNA are transcribed by RNAPIII, 18S rDNA is transcribed by RNAPI and RPL15a and Actin are transcribed by RNAPII.
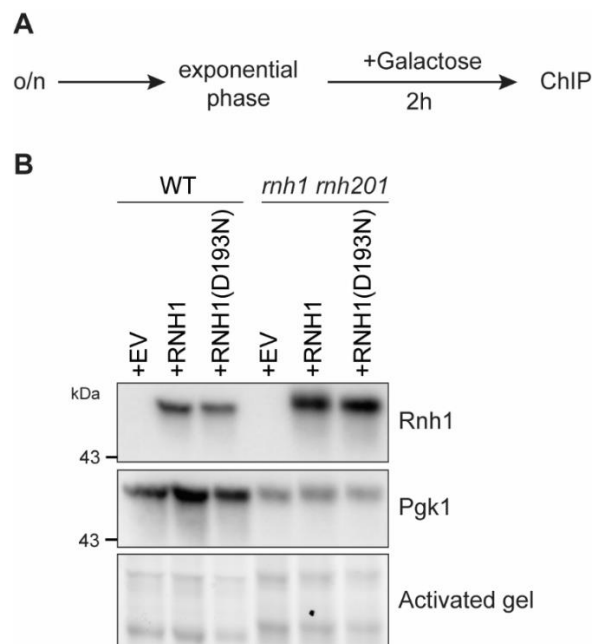
Figure 3 – Overexpression of RNase H1 in wild-type and rnh1 rnh201 mutants. (A) Experimental setup as described in text. **(B)** Western blot of RNase H1 overexpression. Rnh1 and Pgk1 were detected with anti-HA and anti-Pgk1 antibodies, respectively.

Our R-ChIP data show that RNase H1 is significantly enriched in the tRNAs, SUF2 and SUF11, and rRNA genes, 5S and 18S, in both wild-type and *rnh1 rnh201* double mutants (Figure 4). Interestingly, RNase H1 is more enriched in *rnh1 rnh201* mutants than wild-type in SUF2, SUF11 and 5S rDNA (Figure 4). R-loop levels measured by DRIP show that SUF2 tRNA has more R-loops in the absence of RNase H enzymes which are removed upon RNase H1 overexpression (Figure 5). Strikingly, RNase H1 overexpression decreases R-loop levels only in SUF2 (Figure 5), suggesting that RNase H1 removes R-loop once they accumulate in the absence of RNase H enzymes. Moreover, RNPIII-transcribed genes, SUF2, SUF11 and 5S rRNA, show increased R-loop levels when catalytic-dead version of RNase H1 is overexpressed in *rnh1 rnh201* double mutants compared to wild-type cells (Figure 5). This R-loop stabilization relates to an increased recruitment of RNase H1 (D193N) in these same genes in *rnh1 rnh201* mutants (Figure 4). Together these results may suggest that RNase H1 has a strong affinity to RNAPIII-transcribed genes, but it removes R-loops only when they accumulate above wild-type levels.

RNase H1 is significantly recruited to the 18S rDNA locus in both wild-type and rnh1 rnh201 mutants, but RNase H1 enrichment decreases in the absence of RNase H enzymes (Figure 4). Similarly, R-loop levels decrease in *rnh1 rnh201* double mutants compared to wild-type cells (Figure 5). Indeed, overexpression of both active and defective RNase H1 versions did not affect R-loop levels in the 18S rRNA gene, suggesting that RNase H1 is not active in the 18S rDNA locus, while being recruited due to elevated R-loop levels. In RNAPII-transcribed genes, RPL15a and Actin, RNase H1 is not significantly enriched (Figure 4), which is associated with similar R-loops between wild-type cells and *rnh1 rnh201* mutants (Figure 5), indicating that RNase H1 does not target R-loops in these loci.

Overall, these results suggest that RNase H1 is recruited to accumulated R-loops, particularly in RNAPIII-transcribed genes. However, RNase H1 weak binding to loci that are characterized by elevated R-loop levels, such as 18S rDNA, implies that RNase H1 does not bind to all loci with stable R-loops and additional regulators or conditions may limit RNase H1 binding and/or activity.
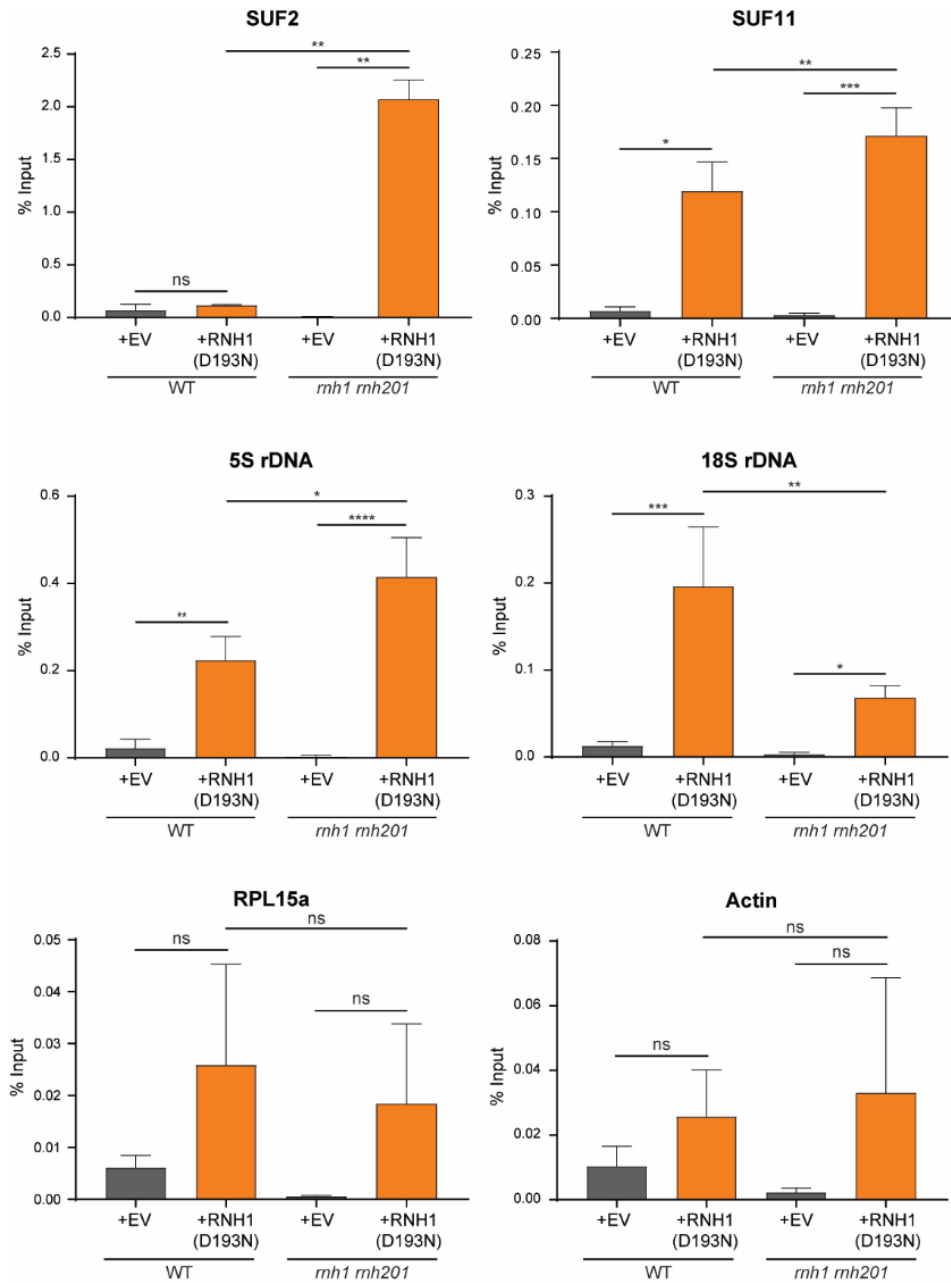
Figure 4 – RNase H1 is recruited to tRNA and rRNA genes. R-ChIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with HA antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns- non-significant). WT, wild-type.
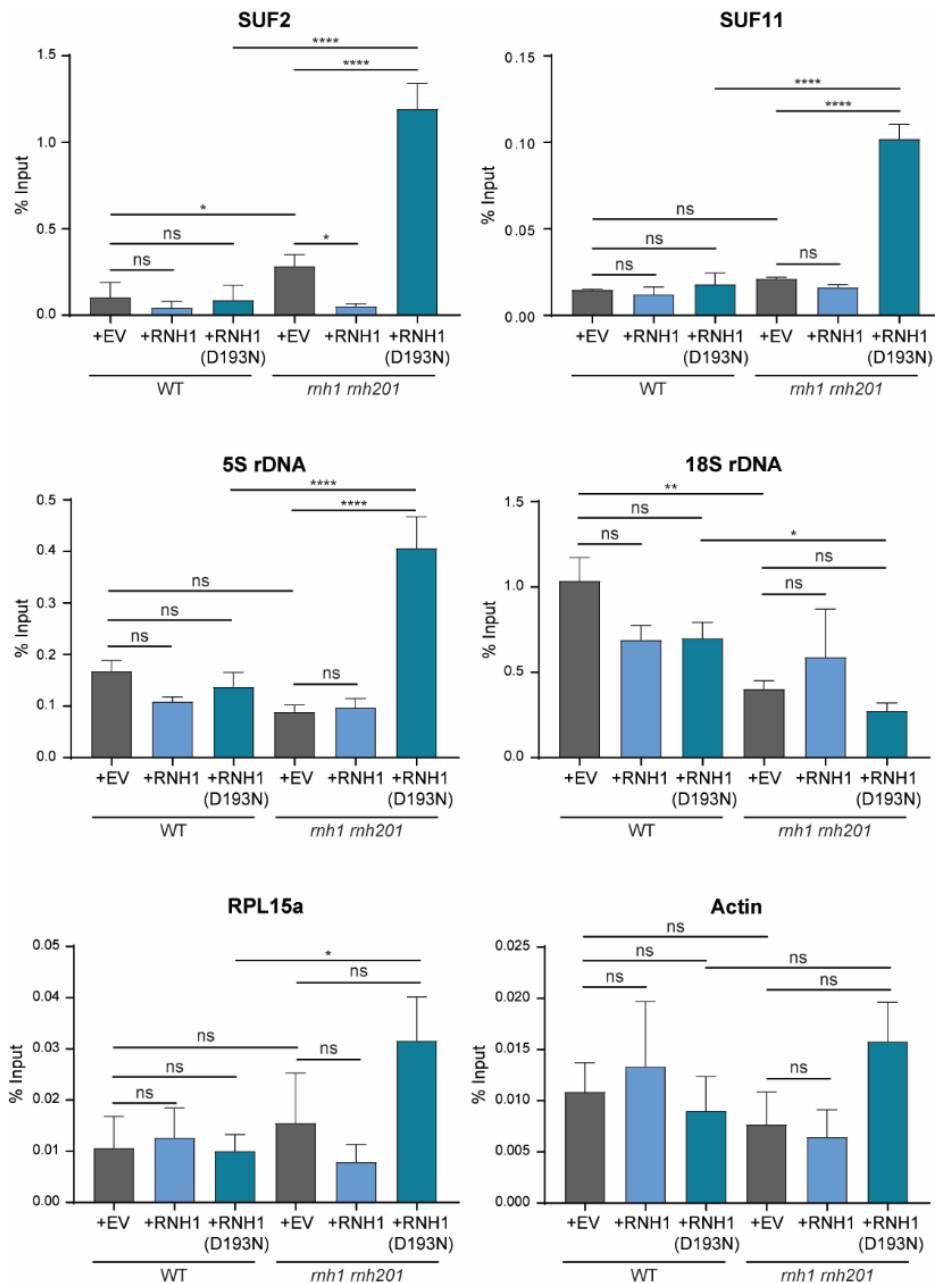
Figure 5 – RNase H1 removes accumulated R-loops. DRIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with S9.6 antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns- non-significant). WT, wild-type.

2.2 R-loop removal through the cell cycle by RNase H1

Although it has been demonstrated that RNase H1 expression can remove R-loops in S- and G2-phase[76], it remains unclear whether RNase H1 removes R-loops in G1-phase. To study RNase H1 activity through all cell cycle phases, we created multiple cell cycle alleles of

RNase H1 that limit its expression to either G1-, S- or G2/M-phase. As a readout, we took advantage of the sensitivity of the double mutant *rnh1rnh201* to conditions that accumulate R-loops, such as addition of the genotoxic agent methyl methanesulfonate (MMS) to the medium[76], or the absence of Sen1, which is accomplished using a degron system to degrade Sen1 upon addition of auxin (Figure 6). We observed that cells are viable when expressing RNase H1 in G1, S or G2/M phase in the absence of RNase H2 and Sen1 (Figure 6A). Similar results were also observed in presence of MMS (Figure 6A). Collectively, these data suggest that RNase H1 is capable of removing R-loops in all cell cycle phases.
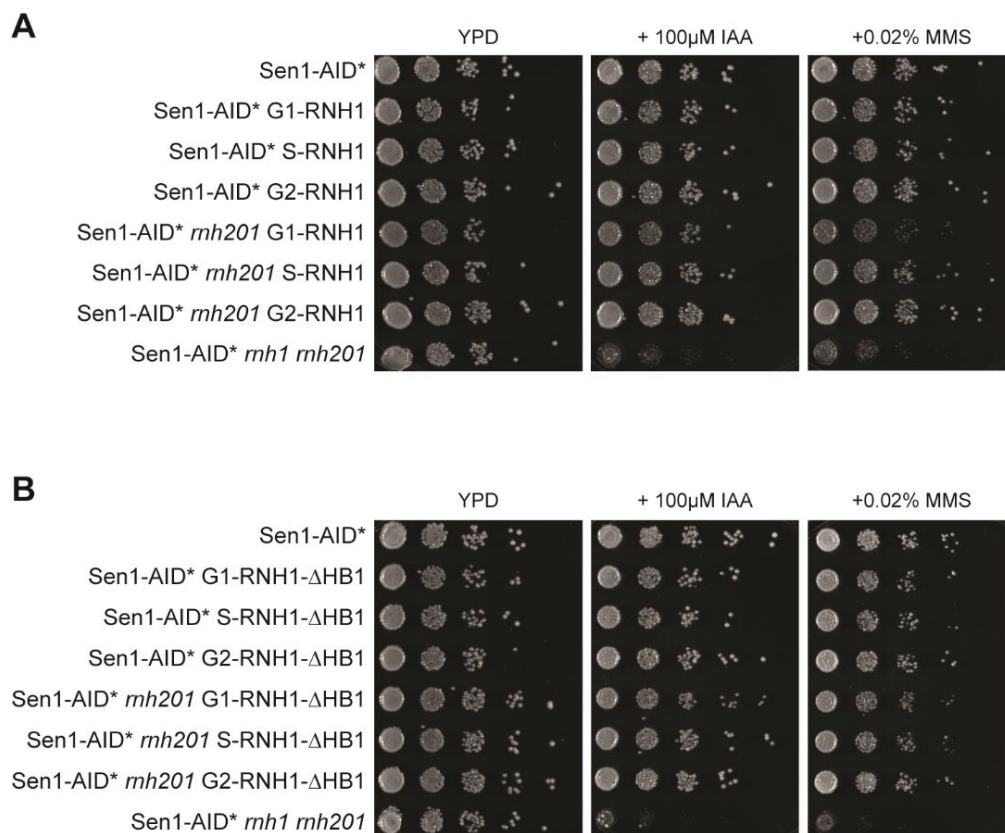
Figure 6 – RNase H1 removes R-loops in all cell phases. Cells of the indicated genotypes were spotted in serial dilutions onto YPD, auxin (IAA) and MMS-containing YPD plates. The plates were imaged after 48 hours of incubation at 30°C. **(A)** Spotting of RNase H1 cell cycle alleles. G1-RNH1, S-RNH1 and G2-RNH1 are only expressed in G1, S or G2 phase, respectively. **(B)** Spotting using cell cycle alleles of RNase H1 truncation mutant without the first hybrid-binding domain.

2.2.1 R-loop removal by RNase H1 in G1-phase

During S-phase, it has been proposed that replication and transcription collisions can be the cause of R-loop formation and RNase H1 might resolve these R-loops[22,90,127]. To better understand the role of RNase H1 beyond the replication-transcription conflicts we mainly

focused on RNase H1 activity during G1-phase. To achieve this, we arrested WT and *rnh1 rnh201* cells in G1 with α-factor, overexpressed RNase H1 or its catalytic dead version and monitored RNA-DNA hybrids and RNase H1 recruitment by DRIP and R-ChIP, respectively (Figure 7A). In both WT and *rnh1 rnh201* mutants RNase H1 is less expressed in G1-phase (Figure 7B), which is consistent with the low mRNA expression levels of GAL genes in G1 (Figure 7C), suggesting that the galactose inducible system might affect RNase H1 overexpression in G1-phase.
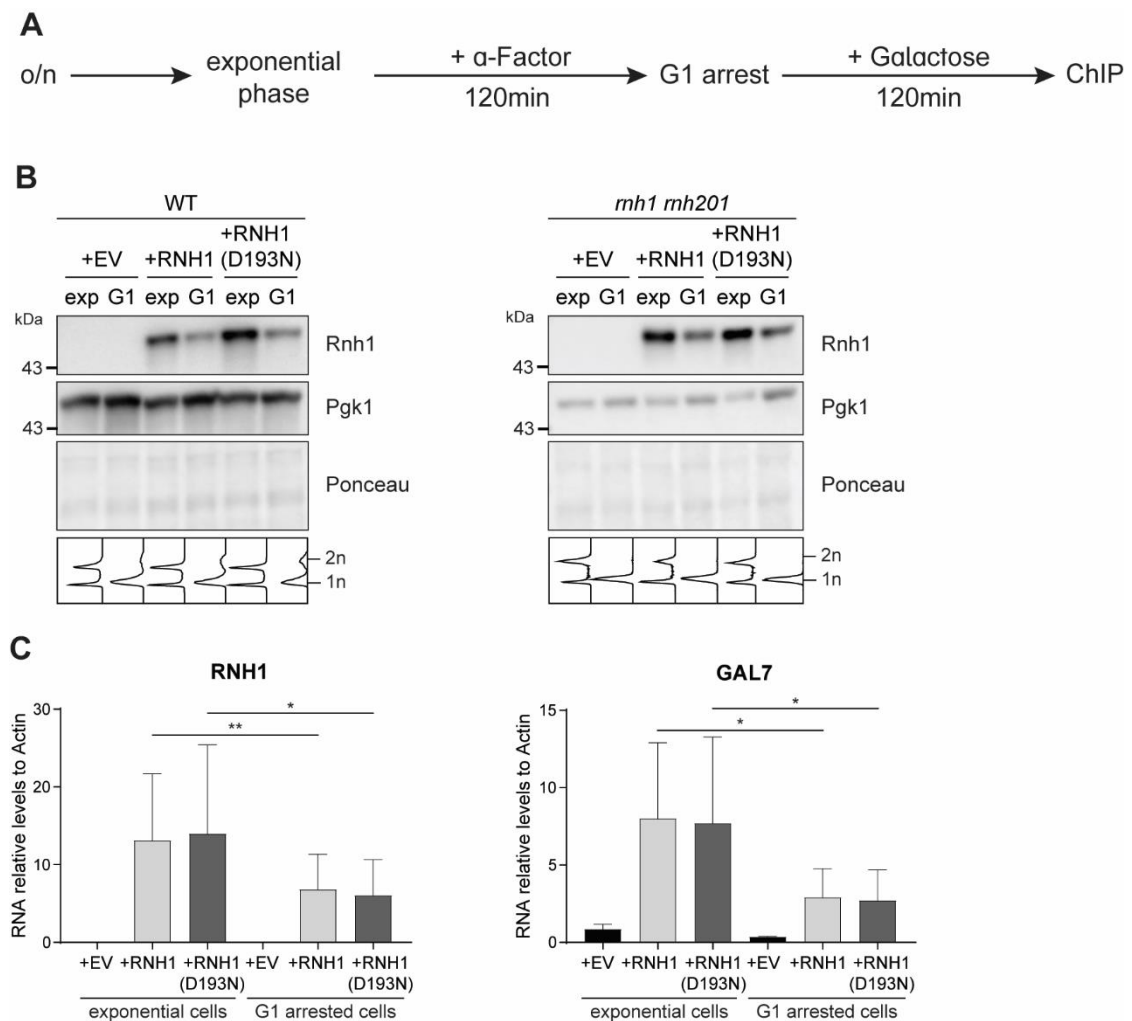
Figure 7 – RNase H1 overexpression in G1-phase. (A) Experimental setup described in text. **(B)** Western blot of RNase H1 overexpression in WT and rnh1 rnh201 double mutants. Rnh1 and Pgk1 were detected with anti-HA and anti-Pgk1 antibodies, respectively. Cell cycle profile of cultures collected for R-ChIP and DRIP. **(C)** RNH1 and GAL7 mRNA quantification by RT-qPCR in rnh1 rnh201 double mutants that overexpress RNH1, RNH1 (193N) or empty vector control. . Data are depicted as mean +SD, n=3. P-values were obtained from Student's t-test (*p<0.05, **p<0.01).
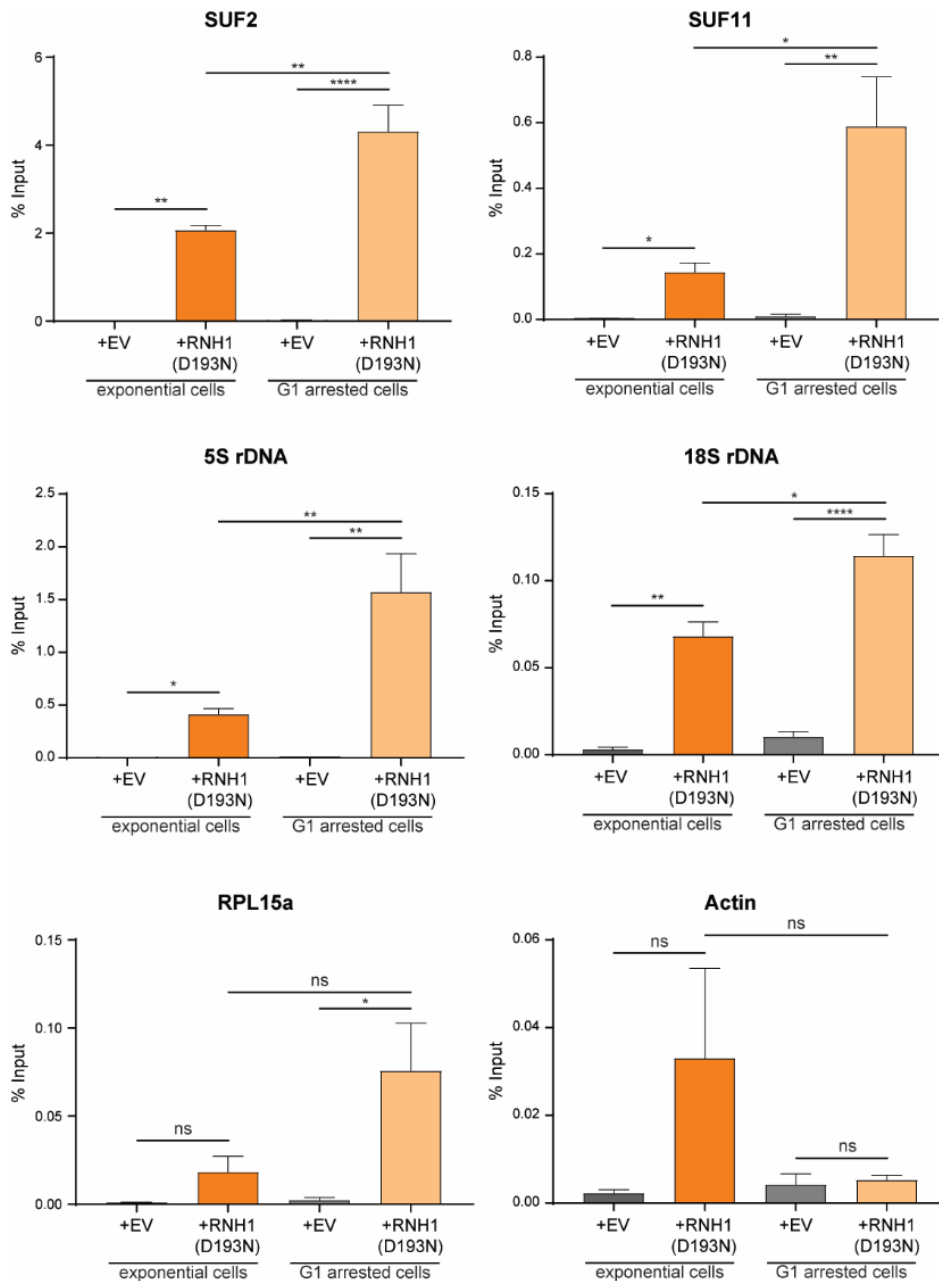
Figure 8 – RNase H1 is recruited to tRNA and rRNA genes in G1-phase. R-ChIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with HA antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns- non-significant).

In *rnh1 rnh201* mutants, RNase H1 is enriched in all selected loci, except Actin, in G1-arrested cells (Figure 8), indicating that RNH1 is recruited to R-loops in G1-phase. Moreover, RNase H1 recruitment to tRNAs, rRNA genes and RPL15a gene is significantly increased in G1-arrested cells compared to cells in exponential phase (Figure 8), which agrees with the stabilization of R-loops by RNase H1 (D193N) overexpression in G1-arrested cells over exponential cells (Figure 9). In addition, overexpression of RNase H1 leads to removal of

accumulated R-loops in G1-phase in SUF2 and RPL15a genes. All together, these data suggest that RNase H1 maintains its activity on accumulated R-loops during G1-phase. Furthermore, in this experiment we confirm that RNase H1 is also capable of targeting RNAPII-transcribed genes, like RPL15a, in conditions that lead to R-loop accumulation.
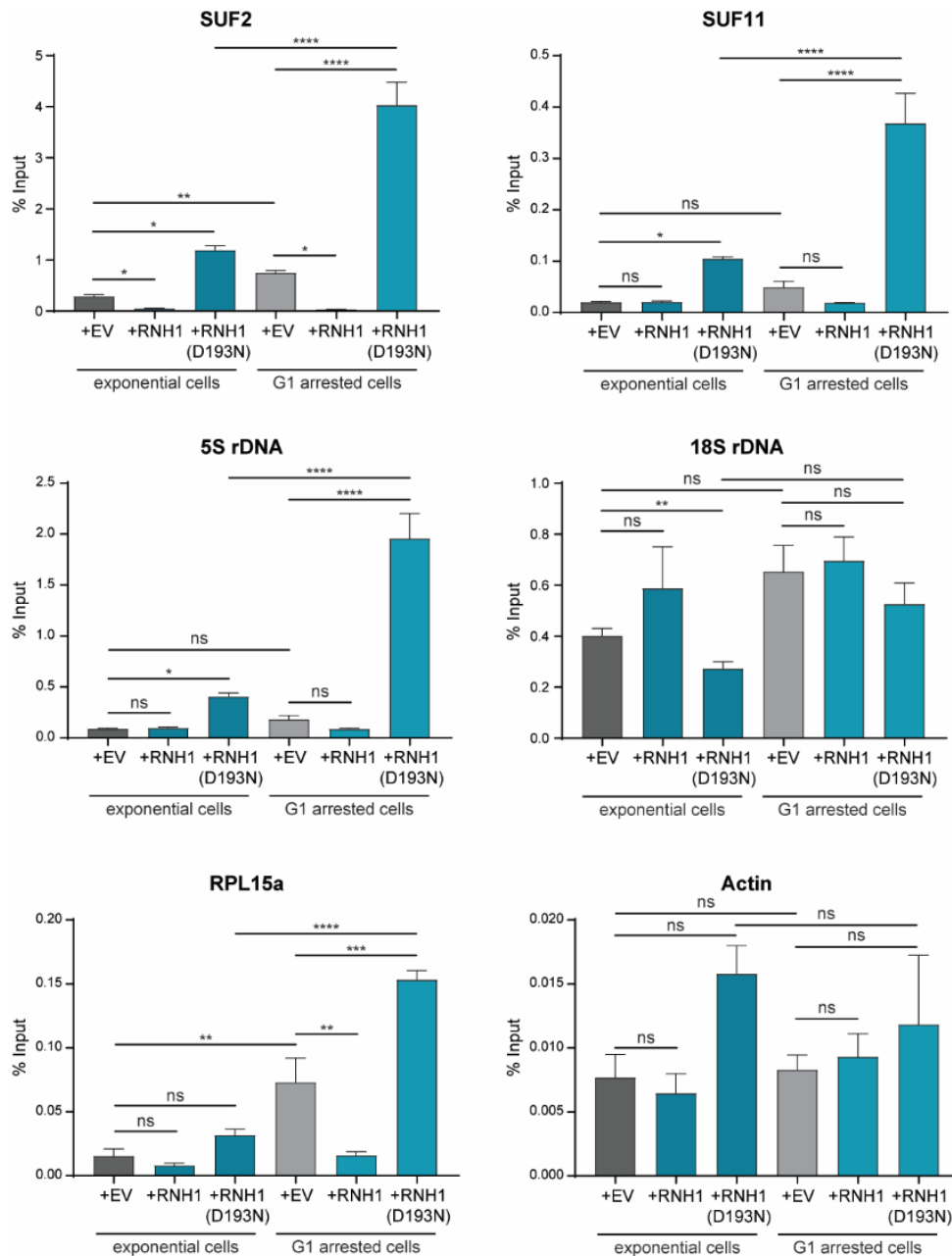

Figure 9 – RNase H1 removes accumulated R-loops in G1-phase. DRIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with S9.6 antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns- non-significant).

To confirm if RNase H1 is responsible for R-loop removal in the absence of stabilized R-loops during G1, we performed R-ChIP and DRIP experiments in exponential and G1-arrested wild-type cells, after overexpressing RNase H1 or RNase H1 (D193N) (Figure 7A). RNase H1 is recruited to the SUF11, 5S rDNA, 18S rDNA and RPL15a loci in wild-type cells arrested in G1-phase, but RNase H1 is more enriched in G1-arrested cells compared to exponential cultures only in the 5S rRNA gene (Figure 10). Interestingly, only 18s rDNA locus contains more R-loops in G1-arrested compared to cycling cells, in which overexpression of RNase H1 significantly reduced R-loop levels (Figure 11), suggesting that RNase H1 targets 18S rDNA locus in conditions that lead to R-loop accumulation. Similar to exponential cells, overexpression of catalytic inactive RNase H1 does not stabilize R-loops in G1-arrested wild-type cells.
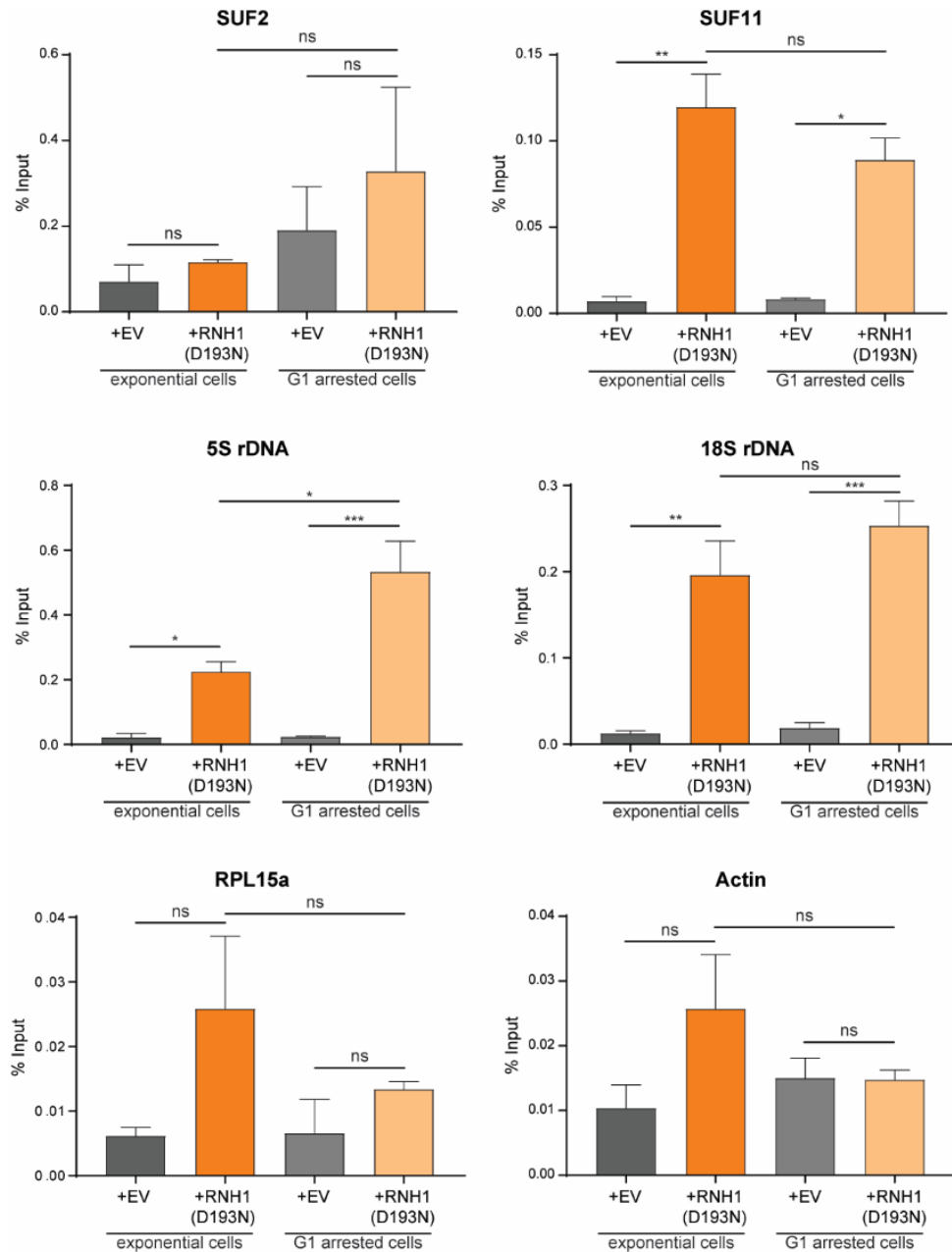
Figure 10 – RNase H1 binds to tRNA and rRNA genes in G1-arrested wild-type cells. R-ChIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with HA antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns- non-significant).
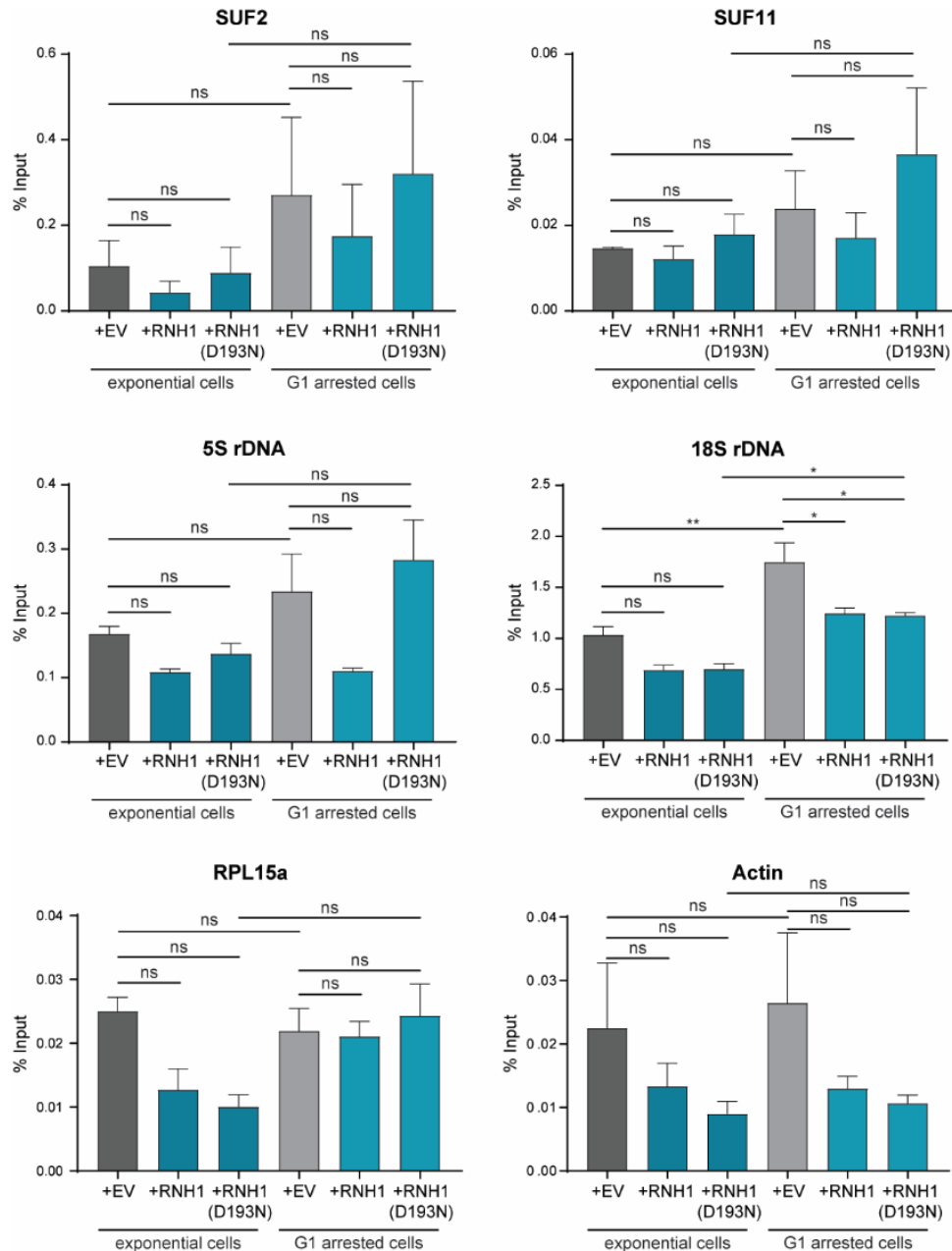
Figure 11 – RNase H1 overexpression does not reduce R-loop levels in G1-arrested wild-type cells. DRIP performed in WT or rnh1 rnh201 mutants that overexpress RNase H1 (RNH1), RNase H1 (D193N) or empty vector control (EV). Chromatin immunoprecipitation with S9.6 antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$, ns- non-significant).

2.3 Endogenous RNase H1 removes accumulated R-loops

RNase H1 is a very low expressed protein, and it has been reported that overexpression of its endogenous levels may force its recruitment into R-loops[205]. In order to understand whether low levels of RNase H1 would allow the enzyme to be still recruited to R-loops in G1,

we performed ChIP and DRIP of endogenously tagged RNase H1 or RNase H1 (D193N) in both G1-arrested and exponential cells. Similar to ectopically expressed RNase H1 (Figure 7B), the protein levels of endogenous RNase H1 are lower in G1 compared to exponential phase (Figure 12A).

We also confirmed by spotting assay that endogenous expressed HA-tagged RNase H1 is functional and does not cause any growth defects (Figure 12B). As it was expected, *rnh1 rnh201* mutant is highly sensitive to MMS, while mutants bearing RNH1-6xHA *rnh201* grew like wild type cells in MMS (Figure 12B). Furthermore, in the absence of RNase H2, the cells that express endogenous catalytic defective RNase H1 are sensitive to MMS (Figure 12B), suggesting that endogenous catalytic inactive RNase H1 does not resolve R-loops. Notably, strains with RNase H1 (D193N) and lacking RNase H2 exhibit similar R-loop levels to *rnh1 rnh201* double mutants (Figure 14), indicating that endogenous catalytic inactive RNase H1 behaves more similar to the absence of RNase H1 and does not lead to strong R-loop stabilization in all loci like the overexpression of RNase (D193N).
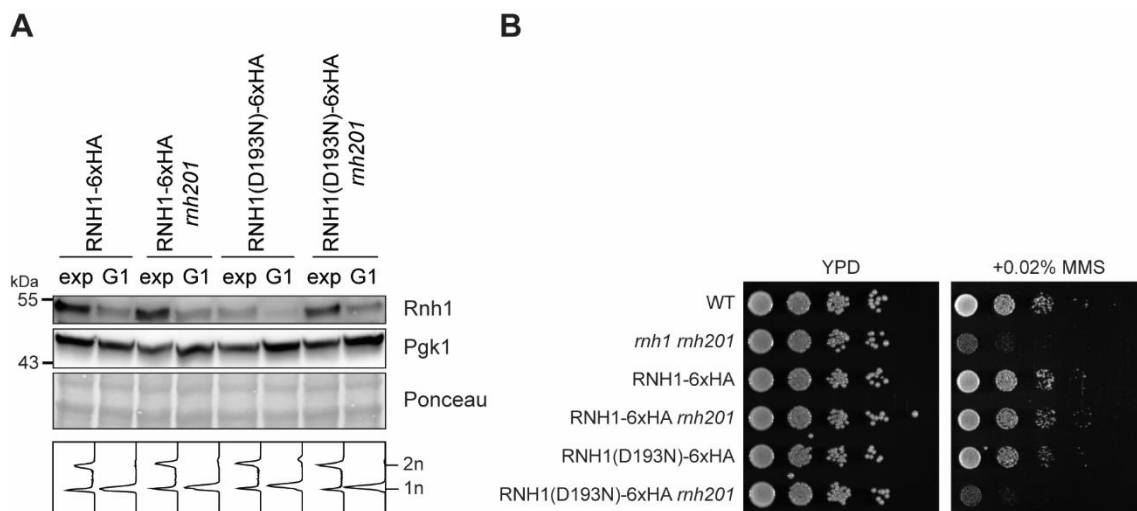
Figure 12 – Endogenous RNase H1 expression in G1-arrested and exponential cells. (A) Western blot of endogenous RNase H1 tagged with 6xHA. Rnh1 and Pgk1 were detected with anti-HA and anti-Pgk1 antibodies, respectively. Cell cycle profile of cultures collected for R-ChIP and DRIP. **(B)** Spotting of endogenous tagged RNase H1. Cells of the indicated genotypes were spotted in serial dilutions onto YPD and MMS-containing YPD plates. The plates were imaged after 48 hours of incubation at 30°C.

Endogenous RNase H1 is only significantly recruited to the SUF2 and 5S rDNA loci in exponential cells, in particular when RNase H2 is absent and RNase H1 is catalytic inactive (Figure 13). In G1-arrested cells, defective RNase H1 is enriched in the SUF2 locus and its recruitment is significantly increased in G1-arrested cells compared to exponential cells. Other RNAPIII-transcribed genes, SUF11 and 5S rRNA, show a general increase in RNase H1

recruitment G1-arrested cells, but only defective RNase H1 is significantly more enriched in SUF11 locus when cells are in G1-phase (Figure 13). RNase H1 recruitment to the SUF2 locus is correlated with the R-loop levels in the locus, in which strains lacking RNase H2 and/or functional RNase H1 accumulate more R-loops and RNase H1 (Figure 14). Similar to previous experiments, R-loop levels increase in G1-phase (Figure 14), leading to higher binding of RNase H1 to SUF2 locus, suggesting that endogenous RNase H1 is recruited to loci with stabilized R-loops. However, the other loci also contain more R-loops in G1-arrested cells or RNase H1 (D193N) *rnh201* strains, while not showing significant enrichment of RNase H1. One possible explanation is that RNase H1 is very low expressed, leading to low detection and recruitment in the majority of the loci. Indeed, RNase H1 is preferentially recruited to SUF2, which accumulates higher amounts of R-loops compared to other loci tested, suggesting that RNase H1 preferentially binds to loci containing more R-loops.

In general, strains with one RNase H enzyme functional maintain low R-loop levels similar to wild-type cells, in both exponential and G1-arrested cultures (Figure 14), suggesting that both RNase H enzymes maintain R-loop homeostasis and are functional in G1-phase. Overall, these results indicate that although endogenous RNase H1 is low expressed, it maintains reduced R-loop levels in cycling and G1-arrested cells.
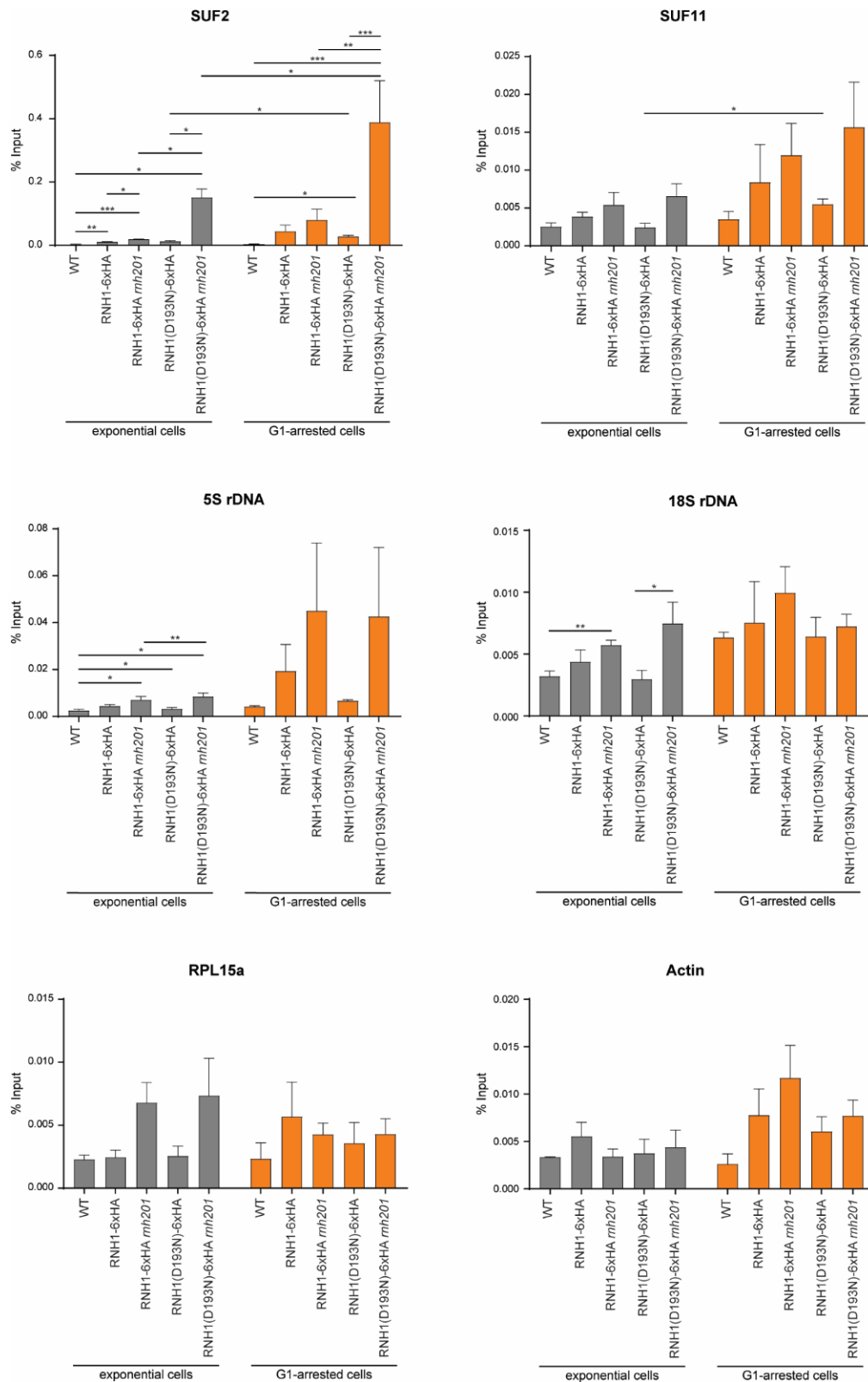
Figure 13 – Endogenous RNase H1 binds to tRNA genes in the absence of functional RNase H1 and RNase H2. R-ChIP performed in WT or mutants with HA-tagged RNase H1 or RNase H1 (D193N). Chromatin immunoprecipitation with HA antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001). WT, wild-type; RNH1, RNase H1.
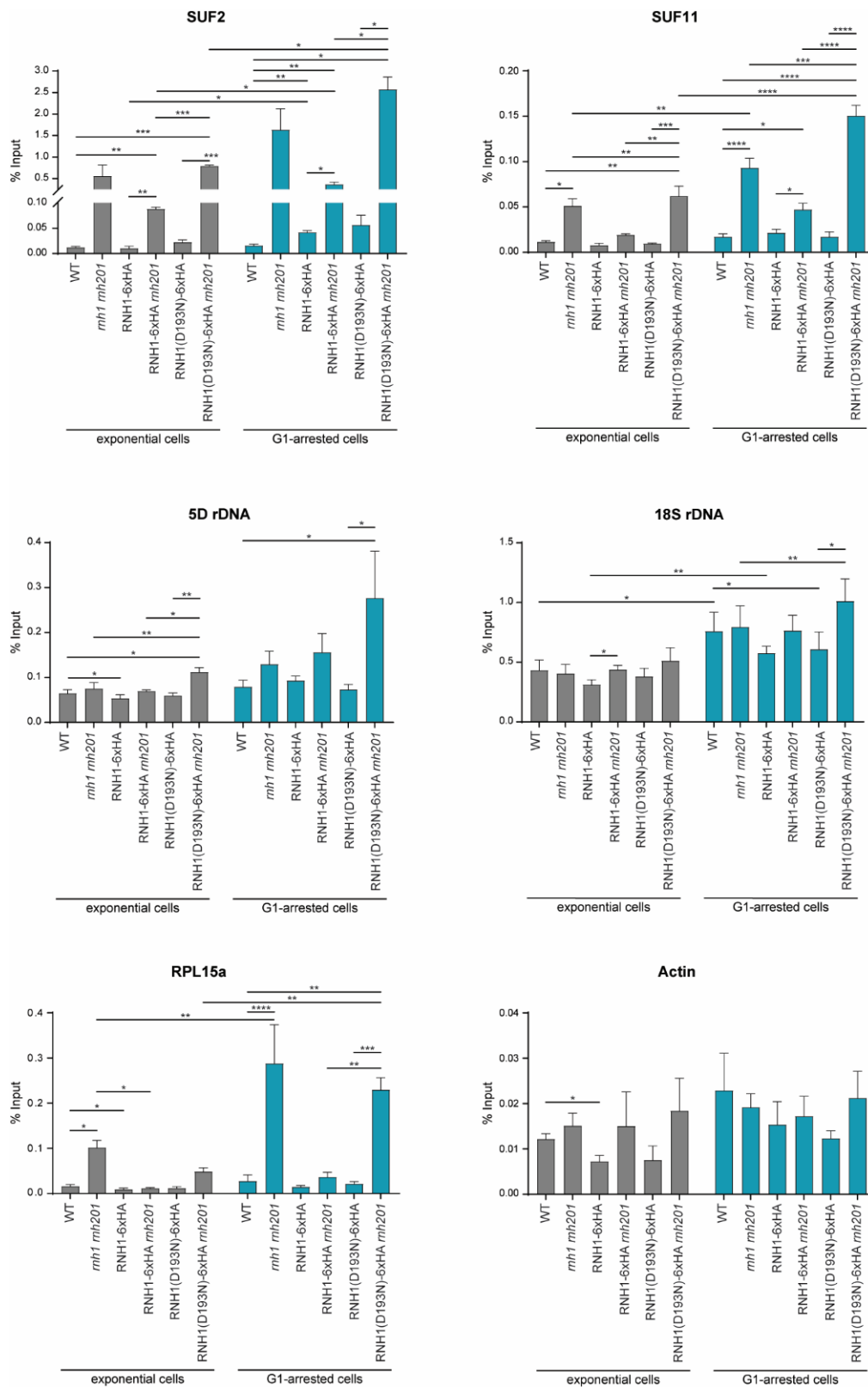
Figure 14 – Endogenous RNase H1 maintains low R-loop levels in G1-phase. DRIP performed in WT or mutants with HA-tagged RNase H1 or RNase H1 (D193N). Chromatin immunoprecipitation with HA antibody and qPCR analysis of the indicated strains at the different loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001). WT, wild-type; RNH1, RNase H1.

2.4 RNase H1 interacts with RPA

In human cells, ssDNA-coated RPA has been identified as an interactor of RNH1. This interaction was described to be important for RNase H1 binding and removal of R-loops[233]. In yeast, a genome-wide proteome screen has identified two subunits of RPA, Rfa1 and Rfa3, as RNase H1 interactors[231]. However, the impact of this interaction on RNase H1 regulation in yeast is still unknown.

First, to understand RNase H1 and RPA interaction, we did a pulldown of HA-tagged RNase H1 and analyzed by western blot if Rfa1 was immunoprecipitated. Immunoprecipitation of endogenous RNase H1 or ectopically expressed RNase H1 captured endogenous Rfa1, confirming that RNase H1 interacts with RPA (Figure 15). Treatment of cell extracts with SM nuclease or DNase I disrupted the interaction between RPA and RNase H1 (Figure 15). On the other hand, removal of RNA or R-loops did not affect RPA-RNase H1 interaction. Thus, RPA and RNase H1 interact independently of RNA and R-loops, whereas the presence of DNA is essential.
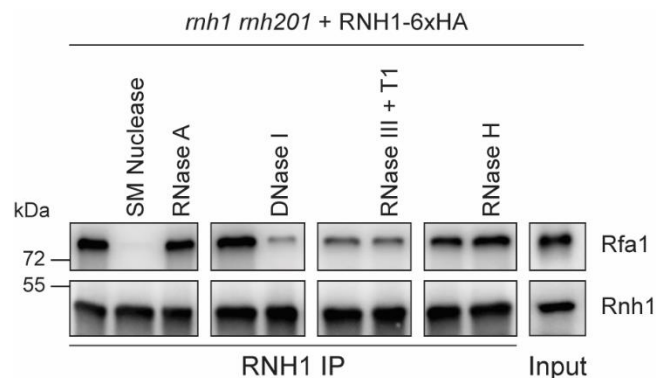
Figure 15 – RNase H1 interaction with RPA is DNA-dependent. HA-tagged RNase H1 ectopically expressed in rnh1 rnh201 double mutants was immunoprecipitated using anti-HA antibody. Each sample was treated with the described nuclease. DNase I, RNase III+T1 and RNase H were in the presence of the respective buffer of each enzyme. HA and Rfa1 in the immunoprecipitates were analyzed by western blot. Cell lysates (12.5%) were loaded as input.

Next, we collected cells progressing synchronously through the cell cycle and immunoprecipitated RNase H1 to study when the interaction of RNase H1 and RPA takes place. Both endogenous and ectopically expressed RNase H1 immunoprecipitation captured Rfa1 in all cycle phases (Figure 16A, B), confirming that the interaction between RNase H1 and RPA occurs in a cell cycle-independent manner. Moreover, we could observe that

endogenous RNase H1 is less expressed is G1-phase (Figure 16A; input), confirming previous observations of RNase H1 expression in G1-arrested cells (Figure 12A).
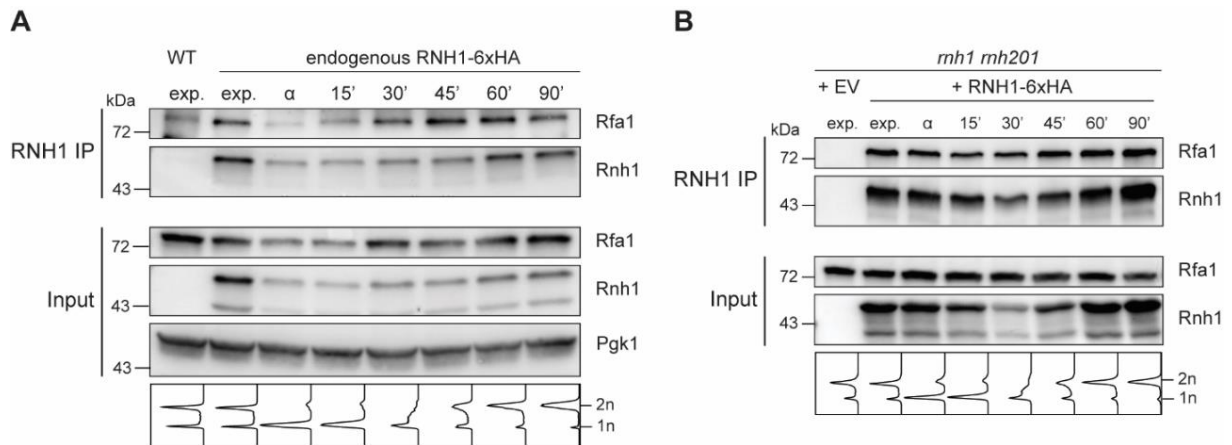
Figure 16 – RNase H1 interacts with RPA in a cell cycle-independent manner. (A) Cells arrested with α-factor and released. Endogenous HA-tagged RNase H1 was immunoprecipitated using anti-HA antibody. HA and Rfa1 in the immunoprecipitates were analyzed by western blot. Cell lysates (12.5%) were loaded as input. Cell cycle profiles were acquired by flow cytometry analysis of DNA content of corresponding sample. **(B)** Cells arrested with α-factor and released. HA-tagged RNase H1 ectopically expressed with the constitutive GPD promoter in rnh1 rnh201 double mutants was immunoprecipitated using anti-HA antibody. HA and Rfa1 in the immunoprecipitates were analyzed by western blot. Cell lysates (12.5%) were loaded as input. Cell cycle profiles were acquired by flow cytometry analysis of DNA content of corresponding sample.

2.5   Characterization of RNase H1 binding to RNA:DNA hybrids

In yeast, RNase H1 is composed by two hybrid-binding domains (HB domains), a catalytic domain and linker domains connecting the other domains. Lack of HB domains prevents binding to and, consequently, removal of RNA-DNA hybrids by RNase H1. It remains unclear how both HB domains act *in vivo* to promote RNase H1 activity. Therefore, we created multiple truncations of HA-tagged RNase H1, which included deletion of the first and/or second HB domain and deletion of both linker domain and first HBD. In addition, we replaced the first HB domain by HB2 or the second HB domain by HB1 (Figure 17A). To understand how RNase H1 interacts with RPA, we overexpressed these constructs with the GAL promoter for 2 hours or with the constitutive GPD promoter in *rnh1 rnh201* double mutants (Figure 17B, C) and performed immunoprecipitation of RNase H1 constructs. While RNase H1 and the constructs that include the first HB domain interact with Rfa1, absence of the first HB domain disrupts RNase H1 interaction with RPA (Figure 18). Also, RNase H1 with two HB2 domains interacts less with Rfa1 than RNase H1 with two HB1 domains or the full length protein (Figure 18A, B). Hence, RPA interacts with RNase H1 primarily through its first HB domain. This is supported

by the pulldown of RPA which shows no interaction with RNase H1 lacking the first HB domain (Figure 18C).
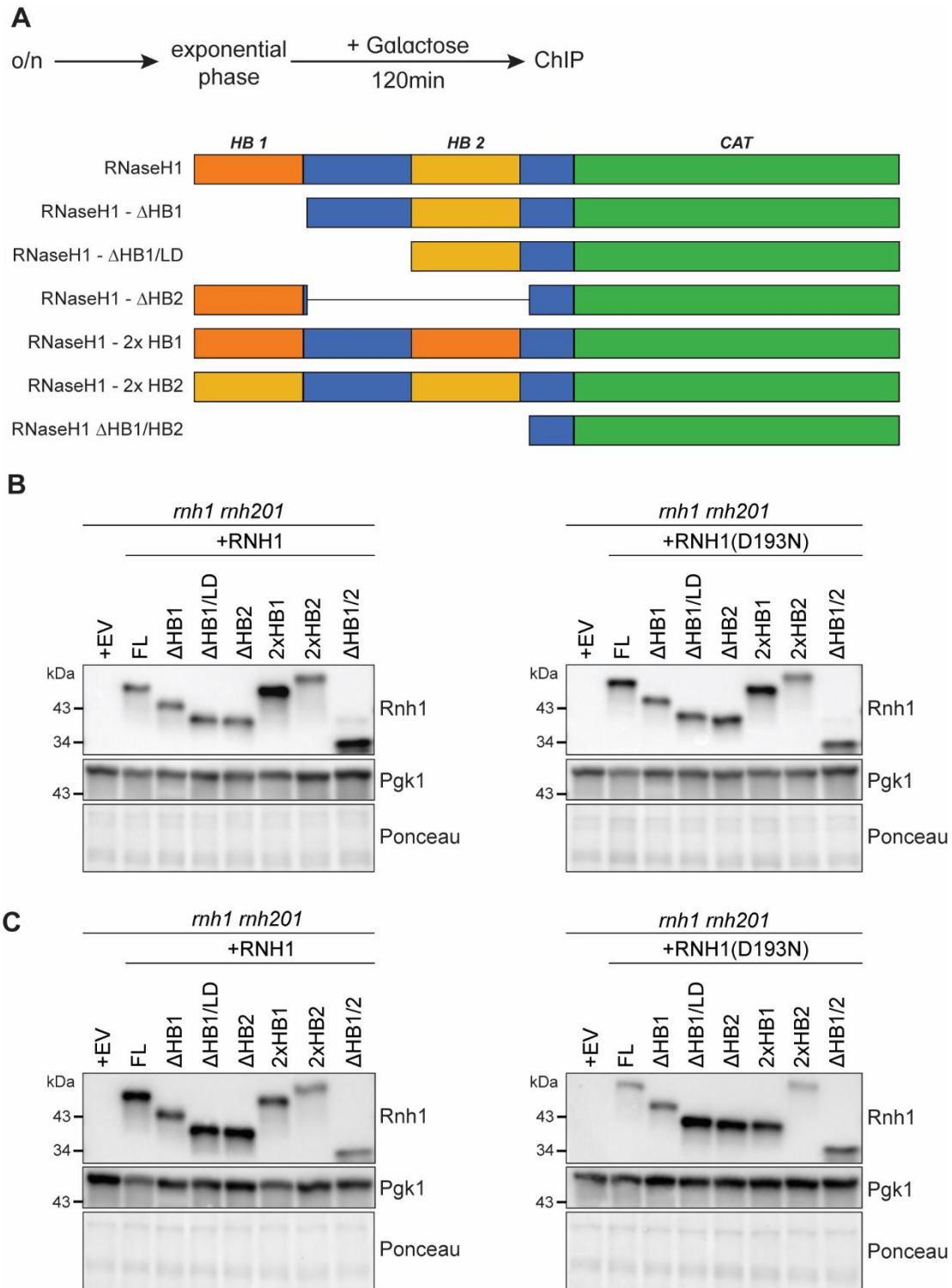
Figure 17 – RNase H1 truncation mutants' expression in rnh1 rnh201 double mutants. (A) Experimental setup described in text. Scheme representing multiple truncations of RNase H1 tagged with C-terminal HA. **(B)** Protein expression of full-length (FL) RNase H1 and truncation versions under

the GAL promoter upon galactose addition for 2 hours. **(C)** Protein expression of RNase H1 FL and truncation versions under the constitutive promoter GPD.
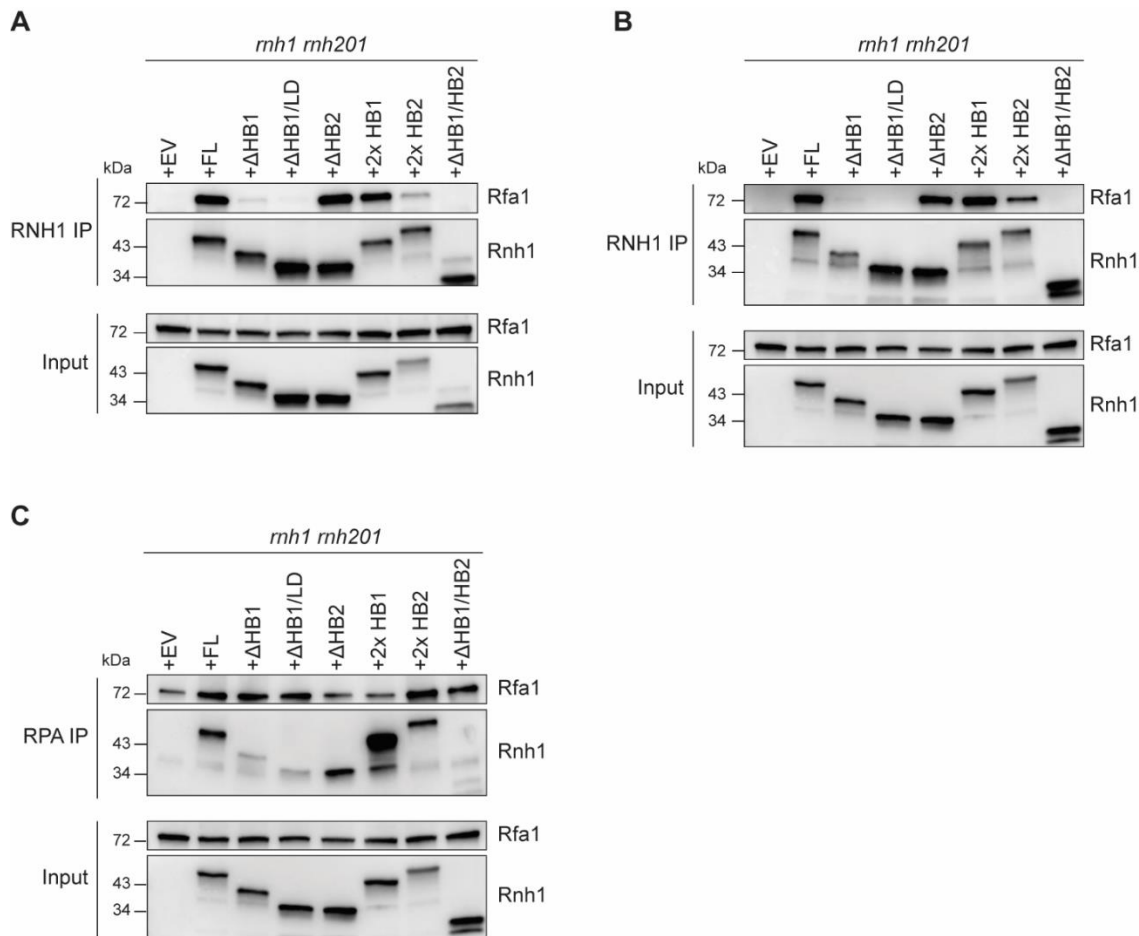
Figure 18 – RNase H1 interacts with RPA through its first hybrid-binding domain. Full-length (FL) RNase H1 and its truncation mutants were ectopically expressed with a GAL promoter **(A)** or GPD promoter **(B)** in rnh1 rnh201 mutants. RNase H1 pulldown with anti-HA antibody. RNase H1 and Rfa1 were detected with anti-HA and anti-RPA antibodies, respectively. **(C)** FL RNase H1 and its truncation mutants were ectopically expressed with a GPD promoter in rnh1 rnh201 mutants. Endogenous RPA was immunoprecipitated with anti-RPA antibody. RNase H1 and Rfa1 were detected with anti-HA and anti-RPA antibodies, respectively.

To study the binding activity of RNase H1 truncation mutants *in vivo,* we overexpressed all the constructs in *rnh1 rnh201* double mutants by addition of galactose for 2 hours in exponential cells and collected samples for R-ChIP and DRIP (Figure 17A). We focused in the SUF2 locus as a readout for RNase H1 binding and activity, due to its high R-loop levels and strong RNase H1 enrichment. As shown previously, full length RNase H1 strongly binds to SUF2 locus, which contains stabilized R-loops in *rnh1 rnh201* mutants (Figure 19B). However, RNase H1 mutants without one or both HB domains (ΔHB1, ΔHB1/LD and ΔHB2) or containing two HB2 domains (2xHB2) are less recruited to R-loops compared to full-length RNase H1

(Figure 19B). Surprisingly, RNase H1 with two HB1 domains strongly binds to SUF2 locus (Figure 19B). This may be due to a strong overexpression of this protein over the other constructs (Figure 17B), which might suggest a dose-dependent effect of RNase H1 overexpression. Moreover, RNase H1- ΔHB2 is significantly enriched over empty vector control, but still less recruited than RNase H1 (Figure 19B), suggesting that HB1 domain has a stronger impact on RNase H1 recruitment to R-loops. Nevertheless, RNase H1 requires both HB domains for efficient binding to hybrids. Except for RNase H1 without both HB domains (ΔHB1/2), overexpression of all others RNase H1 truncations significantly reduces R-loop levels over empty vector control (Figure 19C), implying that ectopically expressed RNase H1 requires at least one HB domain to remove R-loops efficiently. In addition, RNase H1 does not require interaction with RPA to degrade R-loops, since truncation mutants that do not interact with RPA are able to remove R-loops. Furthermore, this confirms that RNase H1 without HB domains cannot bind and process efficiently R-loops, similar to human RNase H1 without HB domain[72,221]. Although, overexpression of RNase H1 (D193N) leads to stabilization of R-loops in SUF2 locus, overexpression of catalytic defective RNase H1 with two HB2 or without one or both HB domains do not accumulate R-loops over the empty vector control (Figure 19D), suggesting that RNase H1 requires both HB domains to have a stable binding capable of stabilizing R-loops.

As expected, ChIP of catalytic active RNase H1 does not show strong enrichment of RNase H1 in SUF2 locus, due to removal of R-loops (Figure 19A). Thus, this ChIP might provide evidence of which catalytic active RNase H1 mutants are removing R-loops efficiently and, consequently, quickly dissociates from its substrate. We show that RNH1-2xHB1 is enriched in SUF2 locus when overexpressed (Figure 19A). Indeed, this mutant seems to remove hybrids less efficiently than wild-type RNase H1, although not statistical significant (Figure 19C), suggesting that RNase H1 with two HB1 domains is less competent removing R-loops and/or dissociating from R-loops.

In addition, we performed RPA ChIP to confirm that RPA recruitment was not affected by different RNase H1 mutants that do not interact with RPA. In general, RPA is highly enriched when elevated levels of R-loops are detected by DRIP (Figure 19E, F). RNase H1 mutants lacking RPA interaction do not show less RPA recruitment to the SUF2 locus compared to full length RNase H1 (Figure 19E), suggesting that RPA binds to R-loops and overexpression of RNase H1 mutants that do not interact with RPA, such as RNase H1-ΔHB1 and RNase H1-ΔHB1/LD, do not impact RPA recruitment to R-loops. However, RNase H1-2xHB1 overexpression leads to higher RPA recruitment to SUF2 locus compared to full-length RNase H1 (Figure 19E), while R-loop levels are not significantly higher (Figure 19C), implying that mutants that interact with RPA, coupled with high expression levels, might affect RPA

recruitment to hybrids. Overall, these results indicate that RPA is recruited to R-loops and RNase H1 has little impact in RPA recruitment to loci with accumulated R-loops. Furthermore, we confirm that RNase H1 activity is independent of its interaction with RPA.
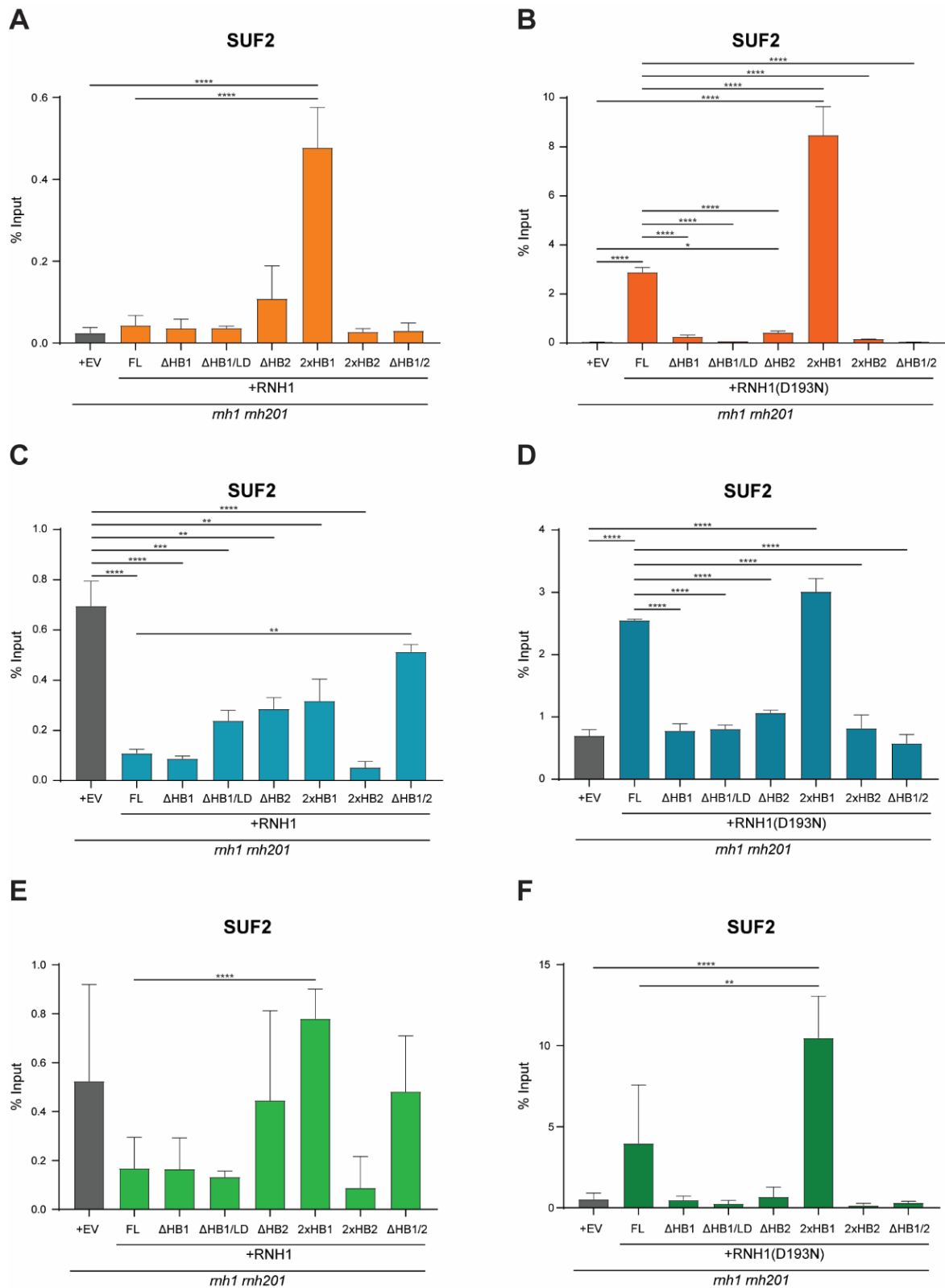
Figure 19 – RNase H1 requires both hybrid-binding domains for a stable binding with R-loops, while one HB domain is enough to remove R-loops upon RNase H1 overexpression. (A) RNase

H1 ChIP of ectopically expressed RNase H1 under a GAL promoter in rnh1 rnh201 mutants. **(B)** RNase H1 (D193N) ChIP (R-ChIP) of ectopically expressed RNase H1(D193N) under a GAL promoter in rnh1 rnh201 mutants. **(C)** DRIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 under a GAL promoter. **(D)** DRIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 (D193N) under a GAL promoter. **(E)** RPA ChIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 under a GAL promoter. **(F)** RPA ChIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 (D193N) under a GAL promoter. Chromatin immunoprecipitation with HA (**A, B**), S9.6 (**C, D**) or RPA **(E, F)** antibody and qPCR analysis of the indicated strains at SUF2 loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001). EV, empty vector; FL, full-length; RNH1, RNase H1.

Alternatively, we overexpressed RNase H1 truncation mutants using the constitutive GPD promoter. As shown before, full-length RNase H1 strongly binds to SUF2 locus and removes R-loops or stabilizes them when catalytic defective version of it is expressed (Figure 20B, C, D). RNase H1 truncations show weaker binding to R-loops compared to wild-type RNase H1 (Figure 20B). Similar to full length RNase H1, truncation versions of RNase H1 remove accumulated R-loops in SUF2 locus, except when lacking both HB domains (Figure 20C), confirming that RNase H1 requires at least one HB domain to remove R-loops. On the other hand, catalytic inactive versions of RNase H1 mutants do not accumulate R-loops like RNase H1 (D193N) (Figure 20D). ChIP of catalytic active RNase H1 mutants shows enrichment of RNase H1-2xHB1 over empty vector control and full-length RNase H1 (Figure 20A). Additionally, although not statistical significant, RNase H1-ΔHB2 seems to stay bound to SUF2 locus, suggesting that lack of HB2 domain decreases efficient R-loop degradation or HB2 domain is important for dissociation from loci after R-loops have been removed.

Moreover, we show that deletion of linker domain, such as in RNase H1-ΔHB1/LD, does not affect R-loop binding and removal comparing to RNase H1-ΔHB1 (Figure 20B, C), suggesting that linker domain between HB1 and HB2 does not affect RNase H1 function *in vivo*. In summary, these data indicate that RNase H1 only needs one HB domain to remove accumulated R-loops, but it requires both HB domains for a stable bind to R-loops. Furthermore, it is tempting to speculate that each HB domain have different functions essential for an efficient removal of R-loops.
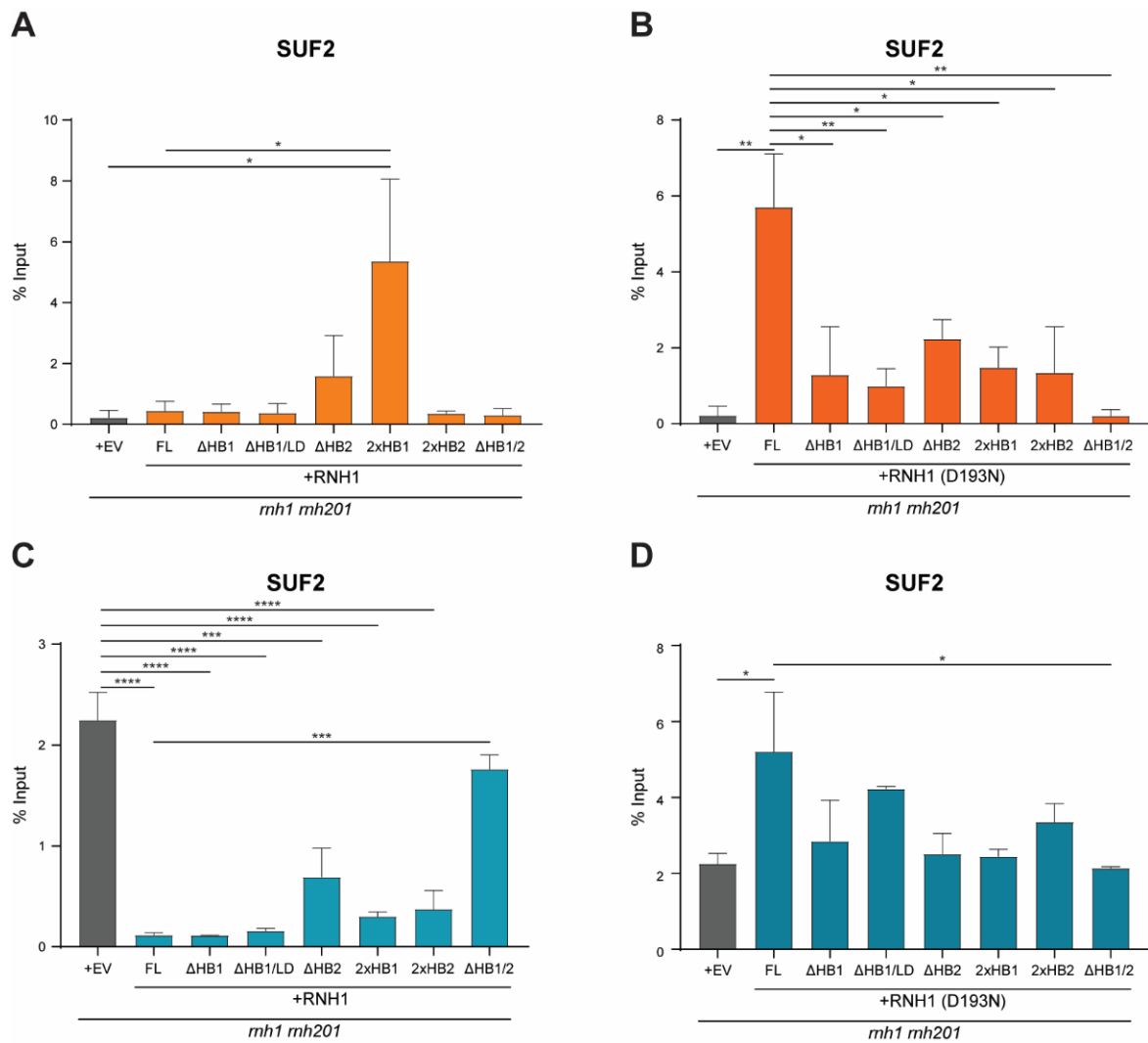
Figure 20 – RNase H1 requires at least one hybrid-binding domain to remove R-loops. (A) RNase H1 ChIP of ectopically expressed RNase H1 under a GPD promoter in rnh1 rnh201 mutants. **(B)** RNase H1 (D193N) ChIP (R-ChIP) of ectopically expressed RNase H1 (D193N) under a GPD promoter in rnh1 rnh201 mutants. **(C)** DRIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 under a GPD promoter. **(D)** DRIP in rnh1 rnh201 mutants with ectopically expressed RNase H1 (D193N) under a GPD promoter. Chromatin immunoprecipitation with HA (**A, B**) or S9.6 (**C, D**) antibody and qPCR analysis of the indicated strains at SUF2 loci. Data are depicted as mean +SD, n=3. P-values were obtained from One-Way ANOVA using the Tukey test to correct for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001). EV, empty vector; RNH1, RNase H1.

2.5.1 Is RNase H1 activity dosage-dependent?

To test whether R-loop removal or stabilization by RNase H1 or RNase H1 (D193N), respectively, would affect cell viability, we did a spotting assay with all RNase H1 mutants expressed in *rnh1 rnh201* double mutants (Figure 21; Figure 22). In addition, we ectopically expressed all RNase H1 truncations using different promoters and copy number plasmids, in order to understand if RNase H1 expression levels would affect survival in the presence of

high R-loop loads. RNase H1 is less expressed using a low copy number plasmid (CEN plasmid) than a high copy number plasmid (2µ plasmid). The inducible GAL promoter also leads to higher expression of RNase H1 than the GPD promoter. Thus, we can observe the effect of RNase H1 at different overexpression levels. As expected, *rnh1 rnh201* double mutants with empty vector control are highly sensitive to MMS (Figure 21). Similarly, RNase H1-ΔHB1/2 also does not rescue growth of *rnh1 rnh201* mutants in MMS, including when it is highly expressed with a 2µ plasmid and/or GAL promoter (Figure 21), confirming that RNase H1 requires HB domains to remove R-loops. Although *rnh1 rnh201* with RNase H1-ΔHB1/2 (2µ plasmid with GPD promoter) grows in MMS after 72h (Figure 21B), we observed no rescue after 48h in MMS (data not shown). Most of RNase H1 mutants rescue growth in the presence of MMS similar to full length RNase H1 using any of the plasmid type or promoter (Figure 21). Interestingly, RNase H1-ΔHB2 does not completely rescue *rnh1 rnh201* double mutants' growth in MMS using a low copy number plasmid with GPD promoter (Figure 21A). Also, this mutant shows slower growth when overexpressed with a low copy number plasmid with GAL promoter (Figure 21C). This is consistent with RNase H1-ΔHB2 mutant removing less efficiently R-loops than other mutants, though not statistical significant (Figure 20C). Strikingly, overexpression of RNase H1 without HB2 using a high copy number rescues *rnh1 rnh201* double mutants growth in the presence of MMS (Figure 21B, D). These data show that RNase H1 mutants with lower R-loop processing activity can be compensated by higher expression levels, suggesting that RNase H1 has a dosage-dependent activity.
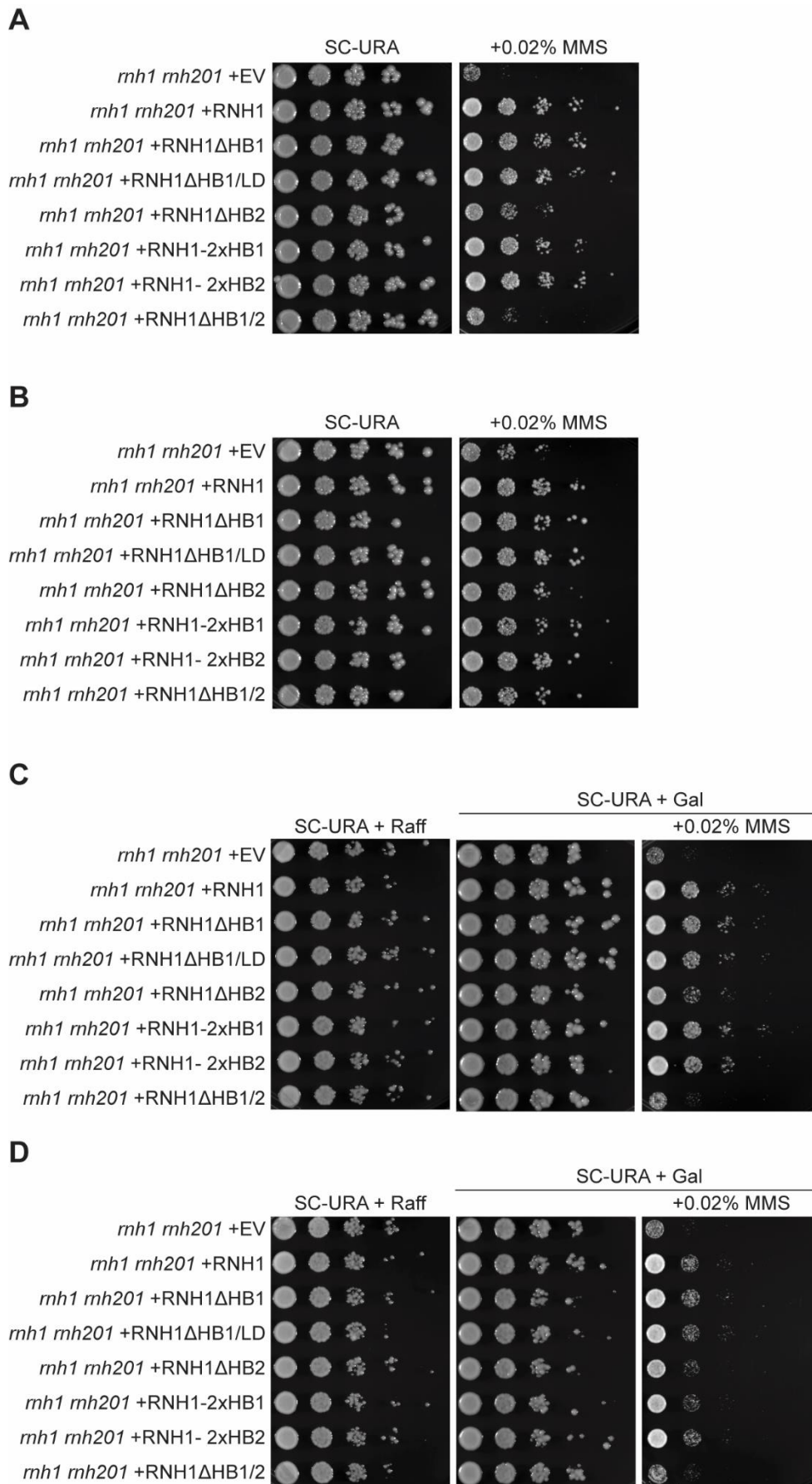
Figure 21 – RNase H1 activity is dosage-dependent. Spotting of rnh1 rnh201 double mutants with **(A)** ectopically expressed RNase H1 truncation mutants from a CEN plasmid with GPD promoter; **(B)**

ectopically expressed RNase H1 truncation mutants from a 2µ plasmid with GPD promoter; **(C)** ectopically expressed RNase H1 truncation mutants from a CEN plasmid with GAL promoter; and **(D)** ectopically expressed RNase H1 truncation mutants from a 2µ plasmid with GAL promoter. Cells of the indicated genotypes were spotted in serial dilutions onto YPD and MMS-containing YPD plates (**A-B**) or SC-URA+Raffinose, SC-URA+Galactose, and MMS-containing SC-URA+Galactose (**C-D**). The plates were imaged after 72h of incubation at 30°C.

Furthermore, overexpression of RNase H1 (D193N) with a low copy number does not decrease cell viability due to stabilization of R-loops (Figure 22C), while a higher overexpression using a 2µ plasmid causes cell death (Figure 22D). Also, constitutive ectopic expression of RNase H1 (D193N) causes increased sensitivity of *rnh1 rnh201* to MMS (Figure 22A), in which high copy number plasmid causes even stronger growth defect (Figure 22 B). Overexpression of RNase H1-2xHB1 also causes cell death using either CEN or 2µ plasmids, since this mutant is expressed at much higher levels compared to the other protein versions (Figure 17B). All together, these results confirm that both removal and stabilization of R-loops are dependent on RNase H1 protein levels, including in mutants with weaker binding to R-loops.
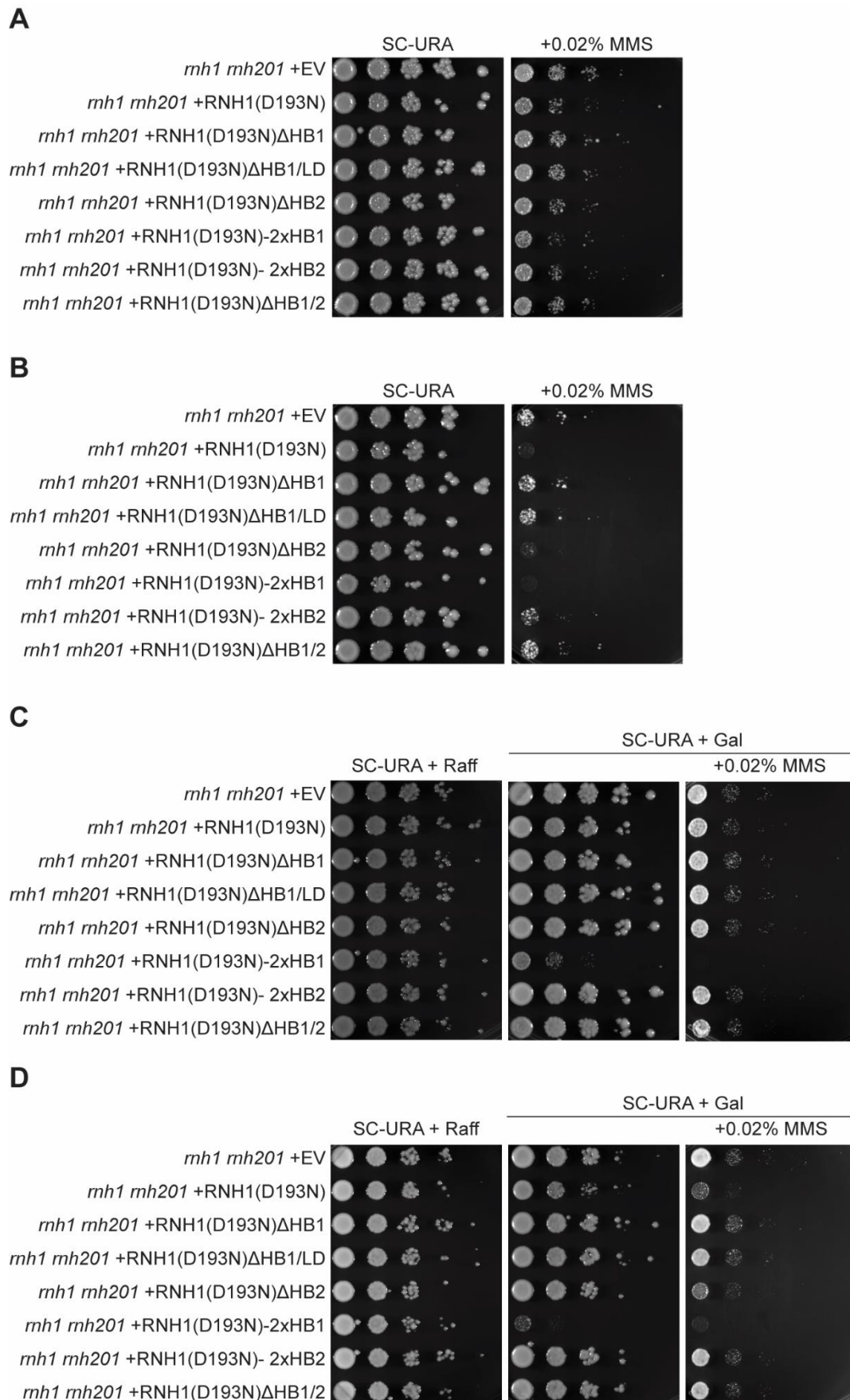
Figure 22 – Defective RNase H1 stabilizes R-loops in dosage-dependent manner. Spotting of rnh1 rnh201 double mutants with **(A)** ectopically expressed RNase H1 (D193N) truncation mutants from a

CEN plasmid with GPD promoter; **(B)** ectopically expressed RNase H1 (D193N) truncation mutants from a 2µ plasmid with GPD promoter; **(C)** ectopically expressed RNase H1 (D193N) truncation mutants from a CEN plasmid with GAL promoter; and **(D)** ectopically expressed RNase H1 (D193N) truncation mutants from a 2µ plasmid with GAL promoter. Cells of the indicated genotypes were spotted in serial dilutions onto YPD and MMS-containing YPD plates (**A-B**) or SC-URA+Raffinose, SC-URA+Galactose, and MMS-containing SC-URA+Galactose (**C-D**). The plates were imaged after 72h of incubation at 30°C.

2.5.2 RNase H1 binding to RNA:DNA hybrids *in vitro*

Previous studies have shown that HB1 domain of RNase H1 is essential for binding to hybrids *in vitro* and HB2 domain does not bind to RNA-DNA hybrids in the absence of HB1 domain[218]. However, we have observed that overexpression of RNase H1 lacking the first HB domain can bind and remove R-loops *in vivo* (Figure 20B, C) and rescue *rnh1 rnh201* sensitivity to MMS (Figure 21). On the other hand, our results from RNase H1 overexpression without HB2 domain suggest that lack of HB2 domain lead to less efficient R-loop degradation than the lack of HB1 domain (Figure 20C). These data confirms a role of HB2 domain *in vivo* in R-loop binding and removal.

To better understand how RNase H1 truncations bind to hybrids, we purified the catalytic-dead versions of full length RNase H1, RNase H1-ΔHB1, RNase H1-ΔHB2, RNase H1-ΔHB1/2 and performed multiple *in vitro* assays. The purified proteins were incubated with different fluorophore-conjugated oligonucleotide substrates and fluorescent polarization (FP) was measured to determine the binding affinity (Figure 25; Table 1). As it was expected, RNase H1 possesses the strongest affinity for R-loops (K_M= 9.9 nM±1.7 nM) and RNA:DNA hybrids (K_M= 20.6 nM±5.9 nM) *in vitro* compared to other nucleic acid substrates (Figure 23A, E; Figure 25; Table 1). Additionally, we constructed multiple hybrid substrates with overhangs that would resemble structures formed during different cellular processes, such as R-loops with a 15 nucleotide 5' ssRNA overhang, a similar structure formed during transcription. Similar to R-loop substrate, RNase H1 binds strongly to R-loop with 5' ssRNA overhang (K_M= 6.6 nM±0.6 nM) (Figure 23B). We found that RNase H1 has a strong affinity with RNA:DNA hybrids with a 5' ssRNA overhang (K_M= 11.8 nM±2.6 nM) or a 3' ssDNA overhang (K_M= 7.6 nM±1.2 nM) (Figure 23F, G). Surprisingly, this binding activity is even stronger to RNA:DNA hybrids without any overhang (Figure 23E), suggesting that RNase H1 might bind to nucleic acids surrounding a hybrid. We also show that loss of one HB domain decreases RNase H1 binding to R-loops and RNA:DNA hybrids substrates, as well as in variants with overhangs (Figure 23). Especially, loss of HB1 domain shows 8 and 11 times lower affinity to R-loops and RNA:DNA hybrids, respectively (Figure 23A, E). Strikingly, RNase H1 without HB2 domain leads to 60 and 53 times weaker affinity of RNase H1 towards R-loops and hybrids, respectively (Figure

23A, E). When RNase H1 does not have both HB domains it cannot bind to all substrates, confirming that RNase H1 interacts with all substrates through its HB domains (Figure 23; Figure 24).

To confirm the FP assays, we incubated the different substrates with 125nM of each protein tested and visualized binding by electrophoresis mobility shift assay (EMSA). Consistent with FP experiments, RNase H1 strongly binds to RNA:DNA hybrid and R-loop substrates, while loss of one HB domain decreases the binding affinity, since the percentage of unbound substrate is highly increased compared to full length RNase H1 (Figure 23C, D, H-J). Thus, the EMSA assays confirm that loss of HB2 domain is more detrimental to RNase H1 binding to hybrids *in vitro* than loss of HB1 domain. As controls, we performed EMSA and FP experiments with ssRNA, dsRNA, ssDNA, dsDNA and dsDNA with ssDNA overhang substrates (Figure 24). RNase H1 shows weaker binding to control substrates compared to hybrid substrates (Figure 25). We found that RNase H1-ΔHB2 does not bind or shows weaker binding than full length RNase H1 or RNase H1-ΔHB1 (Figure 24A, B, E, F, G; Figure 25), suggesting that HB2 domain might bind moderately to other oligonucleotides substrates besides hybrids.

Following that, we purified RPA complex to study the effect of RPA in RNase H1 binding *in vitro*. First, we tested RPA binding to multiple substrates with EMSA and FP assay (Figure 23; Figure 24). As expected, RPA strongly binds to ssDNA *in vitro* (K_M= 2.5 nM±0.3 nM; Figure 24E). Although FP assay shows that RPA strongly binds to dsDNA, EMSA assays does not show any binding to dsDNA by RPA (Figure 24I). This could be due to the fact that FP assay detects binding to non-hybridized ssDNA present in the dsDNA substrate solution. Recently, it was shown that RPA binds to ssRNA and promotes R-loop formation[234]. Consistently, we found that RPA binds to ssRNA (K_M= 552 nM±48.9 nM) stronger than dsRNA or other substrates that do not contain ssDNA (Figure 24A, Table 1). Moreover, we found that RPA strongly binds to R-loop substrates, but weaker to RNA:DNA hybrids *in vitro* (Figure 23A-D; Table 1). This confirms that RPA binds to R-loops through its displaced strand and does not interact efficiently with the RNA:DNA hybrid. Indeed, RPA only strongly binds to substrates that contain exposed ssDNA, such as dsDNA or RNA:DNA hybrid with 3' ssDNA overhang substrates(Figure 25).
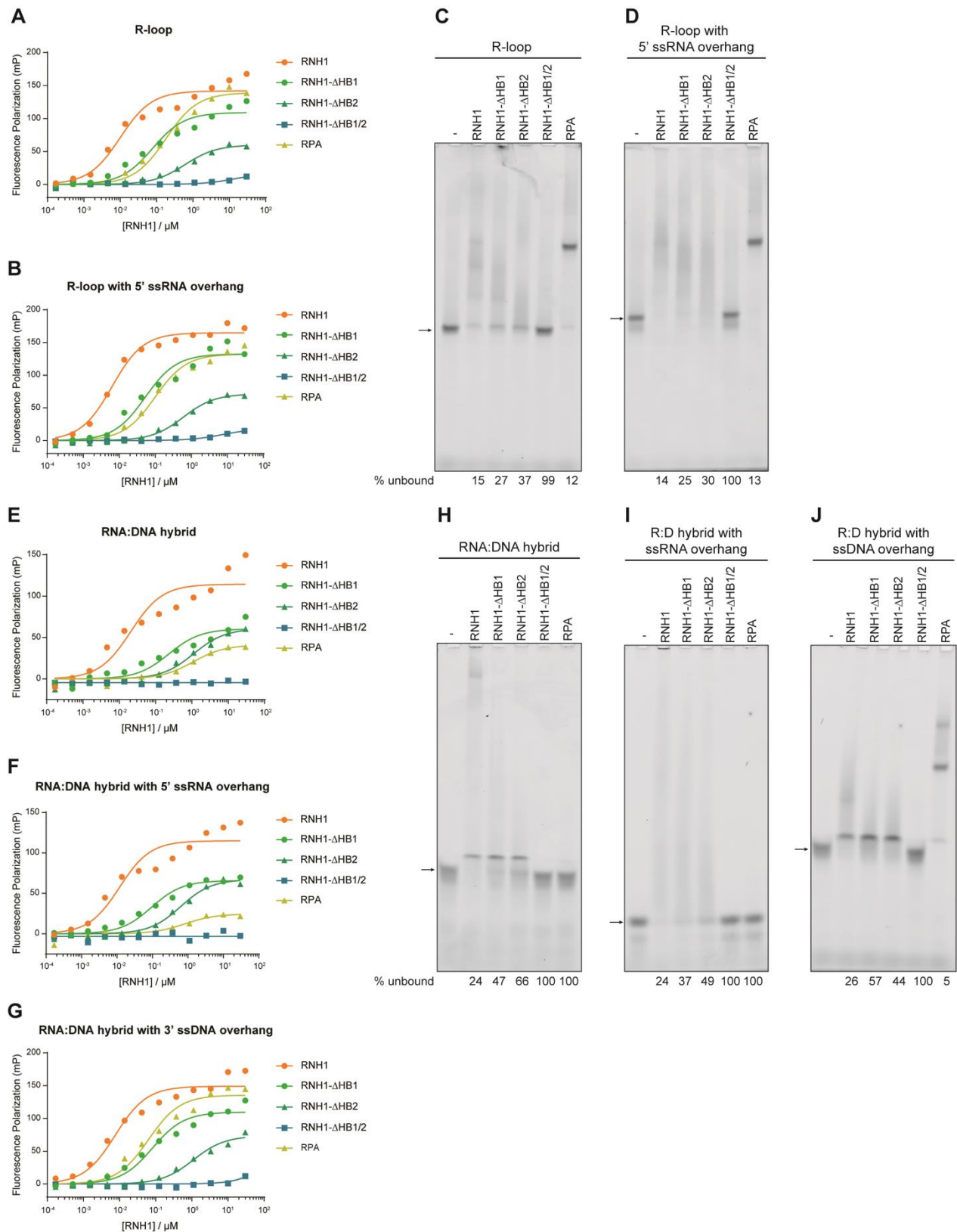
Figure 23 – RNase H1 requires both Hybrid-binding domains for maximum binding with RNA-DNA hybrids *in vitro*. Fluorescent polarization assay of RNase H1 versions and RPA with 5nM of 6-FAM-conjugated R-loop **(A)**, R-loop with 5' ssRNA overhang **(B)**, or RNA:DNA hybrid **(E)**, with 5'ssRNA overhang **(F)** or with 3'ssDNA overhang **(G)** substrate (5nM). Substrate was incubated with serial dilutions of indicate proteins. Data are represented as mean values (n=3). Colored lines represent Michaelis-Menten fits. **(C, D, H-J)** Indicated substrates (25nM) were incubated with RPA or RNase H1

versions (125 nM). Binding to substrates was detected on a native polyacrylamide gel. Arrow indicates unbound substrate.
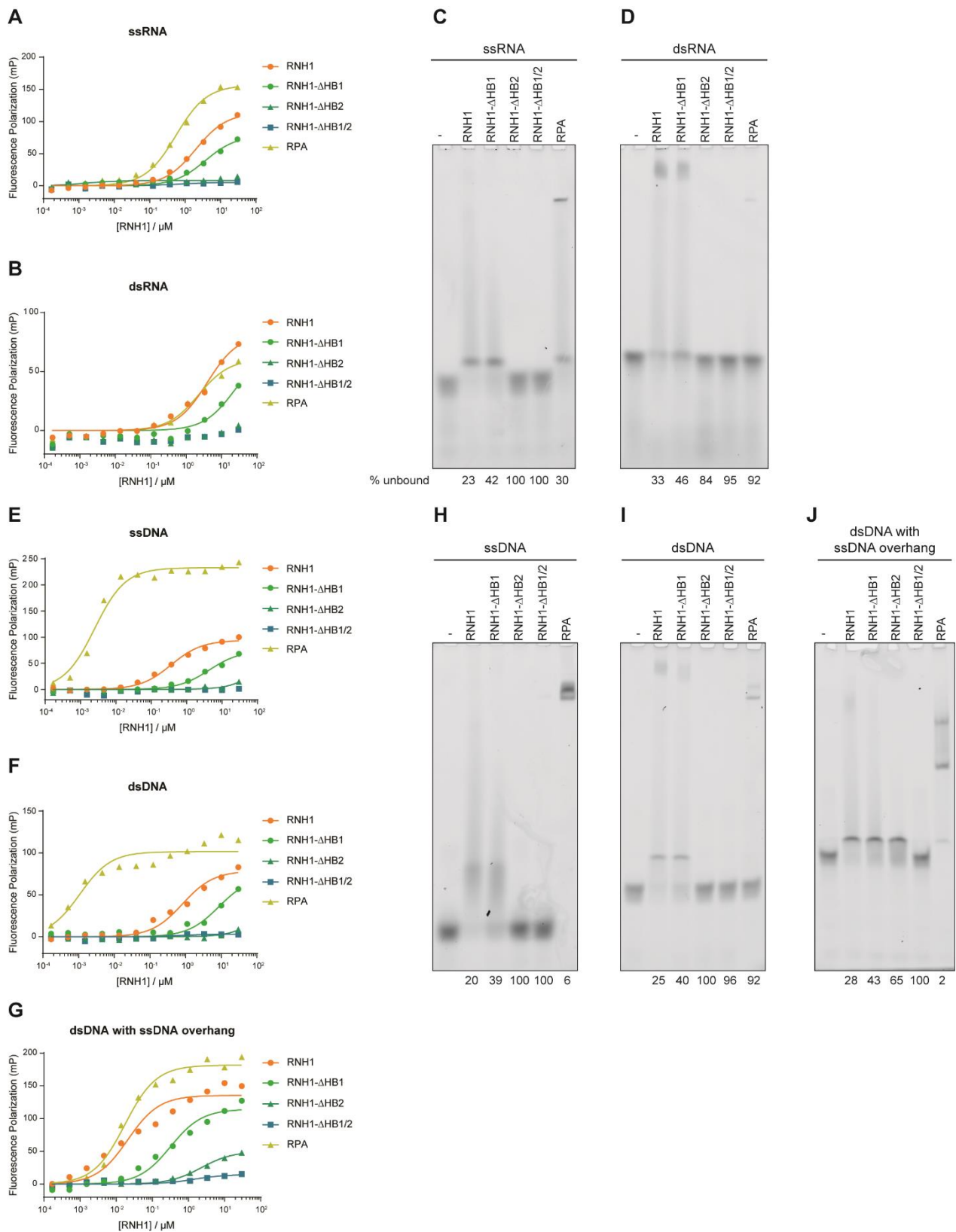
Figure 24 – Second Hybrid-binding of RNase H1 promotes binding to unspecific ssRNA, dsRNA, ssDNA and dsDNA. Fluorescent polarization assay of RNase H1 versions and RPA with 6-FAM-

conjugated ssRNA **(A)**,dsRNA **(B)**, ssDNA **(E)**, dsDNA **(F)** dsDNA with 3'ssDNA overhang **(G)** substrate (5nM). Substrate was incubated with serial dilutions of indicate proteins. Data are represented as mean values (n=3). Colored lines represent Michaelis-Menten fits. **(C, D, H-J)** Indicated substrates (25nM) were incubated with RPA or RNase H1 versions (375 nM in **C, D, H,** and **I**; 125nM in **J**). Binding to substrates was detected on a native polyacrylamide gel.

	RNH1	RNH1-ΔHB1	RNH1-ΔHB2	RNH1-ΔHB1/2	RPA
R-loop	9.9 nM ±1.7	77.3 nM ±14.9	603 nM ±130	14.4 µM ±16	181 nM ±18.8
R-loop with 5' ssRNA overhang	6.3 nM ±0.6	52.6 nM ±13.9	598 nM ±136	7.1 µM ±3.3	98.3 nM ±12.9
RNA:DNA hybrid	20.6 nM ±5.9	232 nM ±71.8	1.1 µM ±0.32	-	1.1 µM ±0.85
RNA:DNA hybrid with 5' ssRNA overhang	11.8 nM ±2.6	95.3 nM ±25.6	623 nM ±120	-	824 nM ±603
RNA:DNA hybrid with 3' ssDNA overhang	7.6 nM ±1.2	81.4 nM ±12.9	1.1 µM ±0.17	-	63.3 nM ±8.6
ssRNA	2 µM ±0.26	3.7 µM ±0.96	-	-	552 nM ±48.9
dsRNA	3.9 µM ±0.7	23.4 µM ±15.9	-	-	2.3 µM ±0.55
ssDNA	0.36 µM ±0.07	4.4 µM ±1.17	-	-	2.5 nM ±0.3
dsDNA	0.8 µM ±0.13	8.6 µM ±2.7	-	-	1 nM ±0.3
dsDNA with ssDNA overhang	20.9 nM ±4.1	321 nM ±47.2	2.3 µM ±0.43	1.4 µM ±1.8	17.4 nM ±2.1

Table 1 – K$_M$ of RNase H1 mutants and RPA measured by FP assay. K$_M$ calculated using a Michaelis-Menten fit for FP assays. Mean values ±SD (n=3).
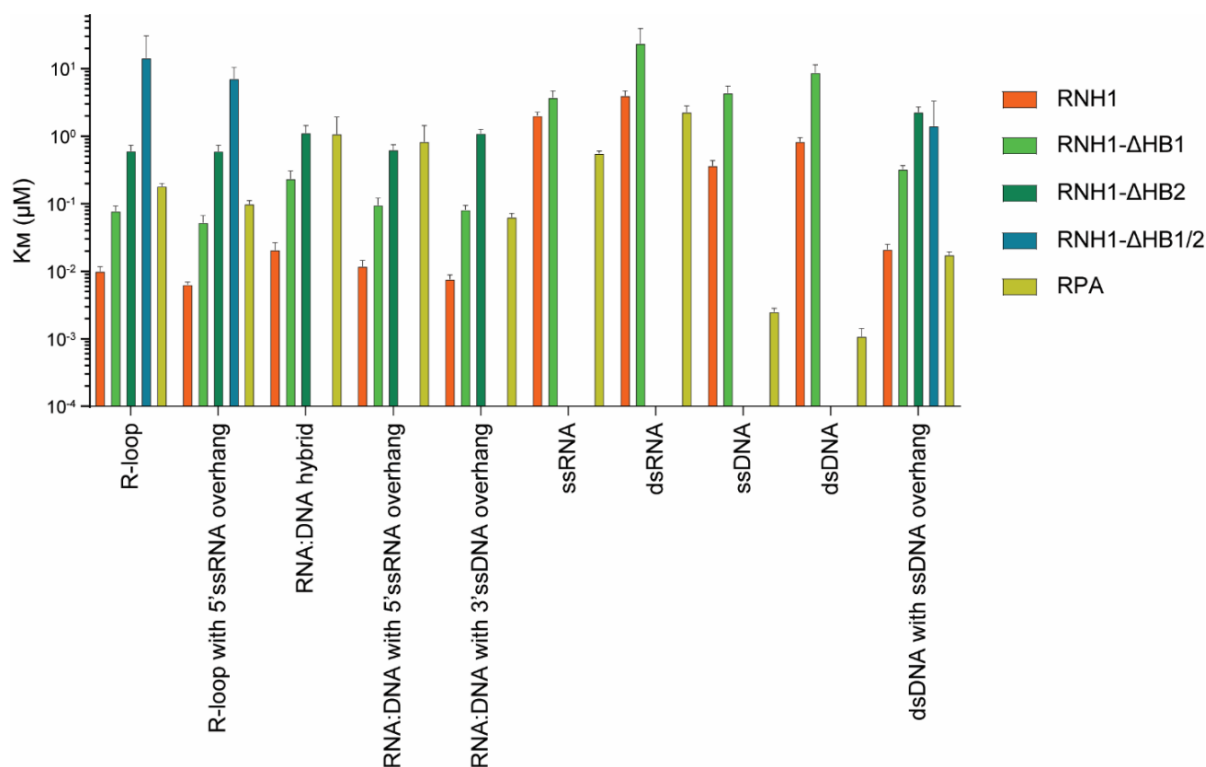
Figure 25 – K_M of RNase H1 mutants and RPA. Graphical representation of K_M values in Table 1. Mean values ±SD (n=3).

Next, we wanted to test the effects of RPA on RNase H1 binding to hybrids, since it is known that in human cells RPA promotes RNase H1 binding to hybrids and degradation of them [235]. For this, we incubated RPA with R-loop or RNA:DNA hybrid substrates and added RPA-substrate mix to a serial dilution of RNase H1 wild-type or truncation mutants (Figure 26; Figure 27). We visualized RNase H1 binding to substrates by EMSA and calculated the percentage of unbound substrate. As shown before, RPA binds strongly to the displaced strand of the R-loop substrate, but in a weaker way to RNA:DNA hybrids. Hence, unbound substrate in RNase H1 samples with RPA were normalized to the amount of unbound substrate in the presence of RPA (Figure 26A, second lane), while samples without RPA were normalized to substrate loaded without RPA (Figure 26A, first lane). Similar to FP assays, full length RNase H1 strongly binds to R-loops and loss of one HB domain reduces its binding (Figure 26B). Also, RNase H1-ΔHB2 shows weaker binding to R-loops than RNase H1-ΔHB1 (Figure 26B). We found that R-loop incubation with RPA before adding RNase H1 did not affect RNase H1 binding to R-loops, suggesting that yeast RPA does not promote RNase H1 binding to R-loops. Moreover, we observe that RPA does not affect binding of RNase H1 without HB1 or HB2 domain. As expected, loss of both HB domains abolishes RNase H1 binding to R-loops and cannot be rescued by RPA addition (Figure 26B).

Since RPA does not bind efficiently to RNA:DNA hybrids, we used higher amounts of RPA and tested RNase H1 binding to RNA:DNA hybrids. As shown previously, RNase H1 binds RNA:DNA hybrids, while lack of one HB domain decreases binding affinity and RNase H1 without both HB domains does not bind to the substrate (Figure 27). RNase H1-ΔHB1 and RNase H1-ΔHB2 bind similar to RNA:DNA hybrid substrate (Figure 27). Addition of RPA to RNA:DNA hybrids decreases amount of unbound substrate faster (Figure 27B), indicating that RPA increases RNase H1 affinity to RNA:DNA hybrids. Surprisingly, RPA also increased RNase H1-ΔHB1 binding to RNA:DNA hybrids, but did not affect binding of RNase H1 without HB2 domain to hybrids (Figure 27B). RNase H1-ΔHB1/2 does not bind to RNA:DNA hybrids and RPA does not increase binding of this mutant (Figure 27B). Together these results show that RPA promotes RNase H1 binding to RNA:DNA hybrids, although it does not bind efficiently to RNA:DNA hybrids. On the other hand, RPA binds to the displaced ssDNA of R-loops, but does not promote RNase H1 binding to R-loops. This may suggest that RPA bound to R-loop or ssDNA does not impact RNase H1 recruitment to R-loops, as described in human cells.

Interestingly, visualization of RNase H1 binding to R-loops and RNA:DNA hybrids with EMSA allows to detect multiple bands of bound RNase H1 to hybrids with increasing concentrations of RNase H1 (Figure 26A; Figure 27A). In both R-loop and RNA:DNA hybrids substrates, a second band is visible upon addition of 60nM RNase H1, which is 2.4 times more than the amount of the substrate in this reaction (Figure 26A; Figure 27A). The multiple sized bands reveal a multimeric binding of RNase H1 to hybrids, meaning that multiple RNase H1 proteins can bind to one substrate molecule. Moreover, incubation of more than 180nM RNase H1 caused majority of substrate to not enter the gel (Figure 26A; Figure 27A), which may suggest that the protein is forming oligomers or aggregates. This was visible mainly with full length RNase H1 and not with truncation mutants, which implies that RNase H1 requires both HB domains to form such oligomers or aggregates.
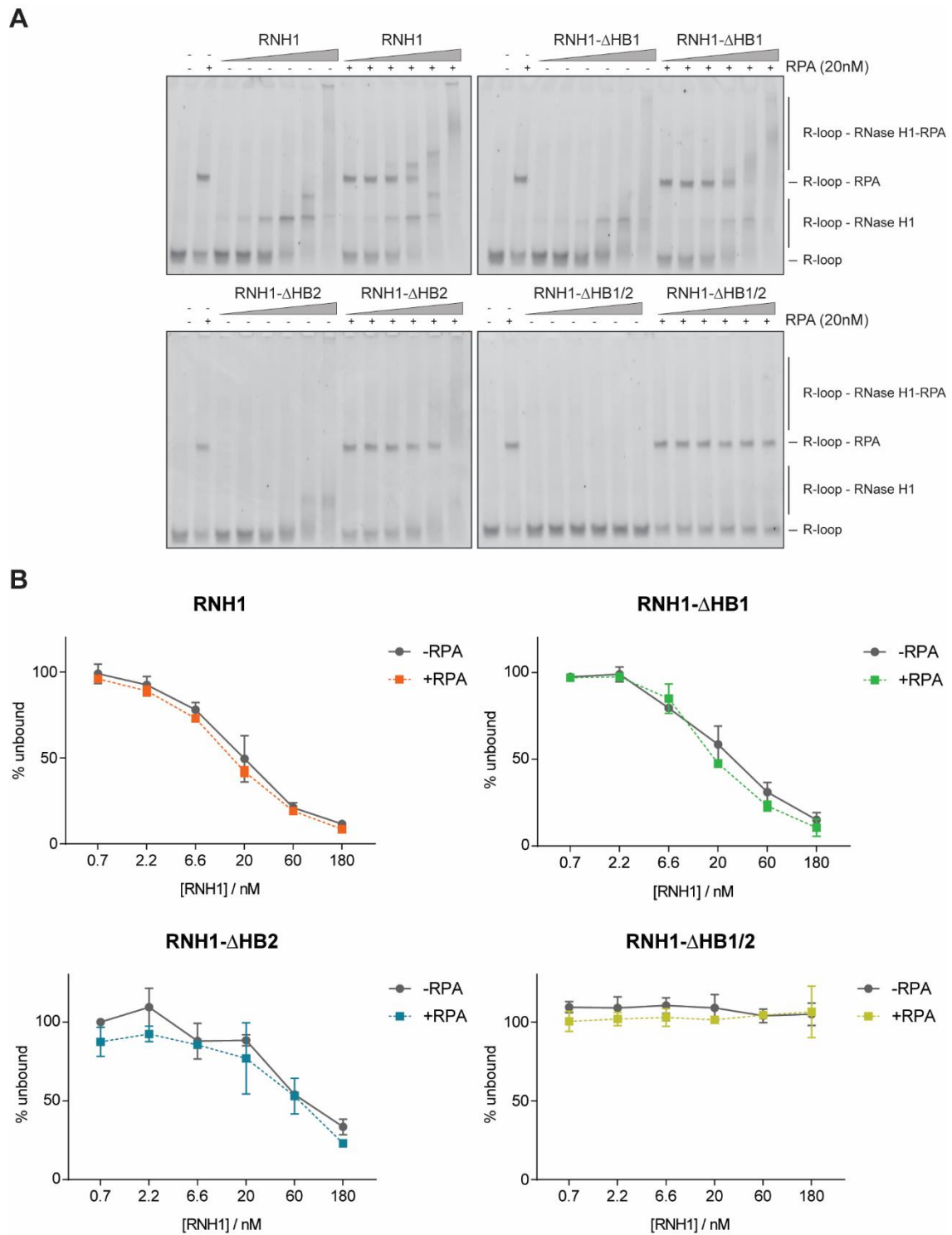
Figure 26 – RPA does not affect RNase H1 binding to R-loops. (A) R-loop substrate (25nM) was incubated with or without RPA (20nM) and with increasing concentrations of RNase H1 versions (0.7, 2.2, 6.6, 20, 60 and 180 nM). Binding to R-loop was detected on a native polyacrylamide gel. **(B)** Quantification of unbound substrate detected in A. Unbound substrate was quantified with ImageJ. Data is presented as mean ±SD (n=2).
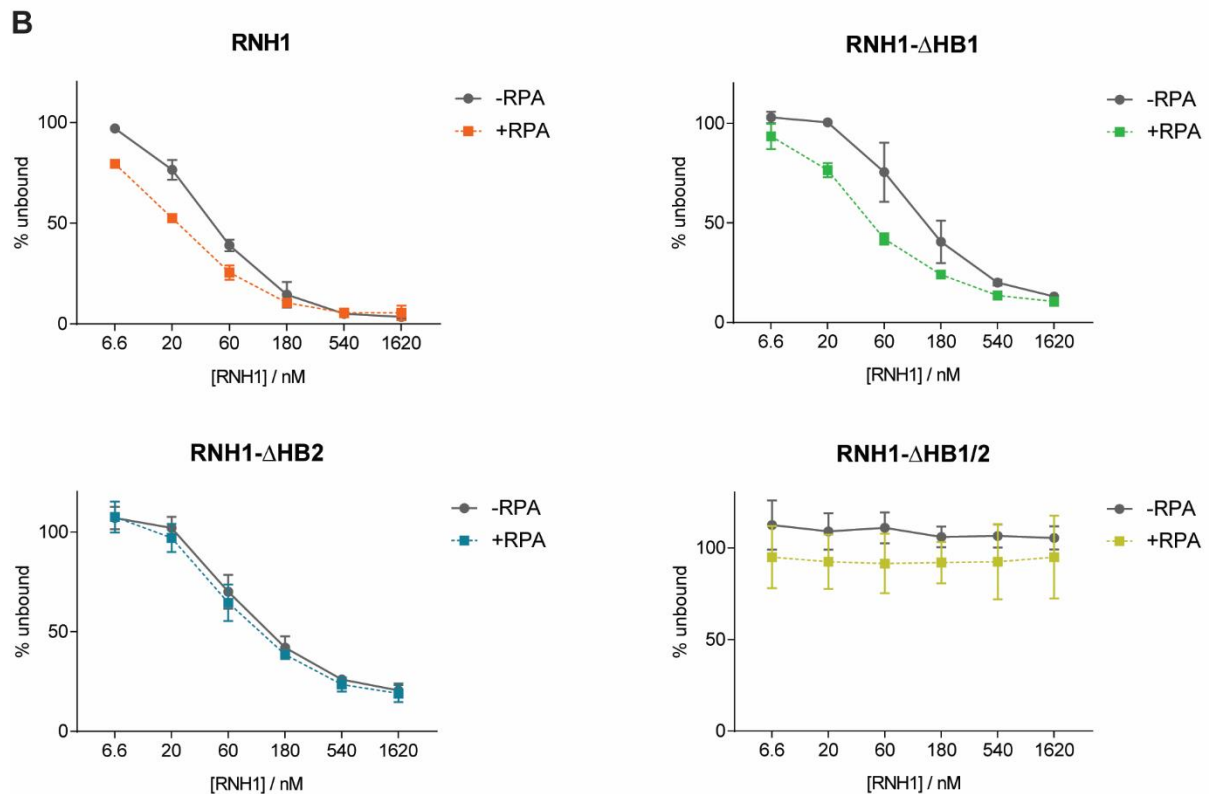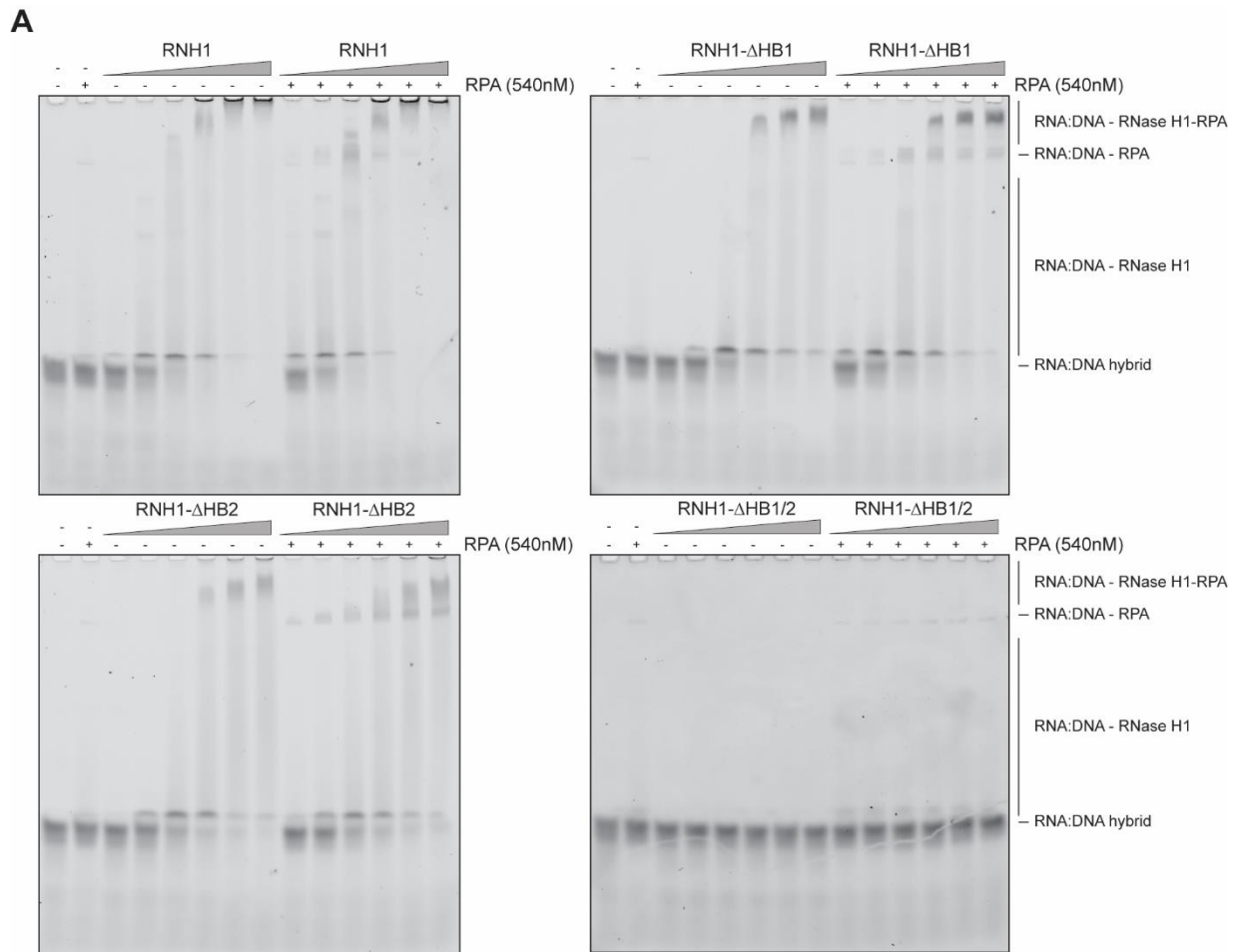
Figure 27 – RPA promotes RNase H1 binding to RNA-DNA hybrids. (A) RNA-DNA hybrid substrate (25nM) was incubated with or without RPA (540nM) and with increasing concentrations of RNase H1

versions (6.6, 20, 60, 180, 540 and 1620 nM). Binding to R-loop was detected on a native polyacrylamide gel. **(B)** Quantification of unbound substrate detected in A. Unbound substrate was quantified with ImageJ. Data is presented as mean ±SD (n=2).

2.6 Identification of RNase H1 regulators

As shown before, *rnh1 rnh201* double mutants are sensitive to conditions that increase R-loop levels, such as in the presence of MMS. Deletion or overexpression of a protein that affects RNase H1 function together with Rnh2 deletion would confer sensitivity to MMS. Similarly, deletion or overexpression of proteins that affect RNase H2 function should be sensitive to MMS in an *rnh1* mutant. Thus, to identify proteins that impact RNase H enzymes activity in removing R-loops, we took advantage of the synthetic genetic array (SGA) procedure to cross *rnh1* or *rnh201* mutants with the complete viable gene deletion library or gene overexpression library of *S. cerevisiae* (Figure 28A).

As outlined in Figure 28B the query strains (*rnh1*, *rnh201* or control ura3 locus marked with a nourseothricin (clonNAT)- resistance marker) were mated with 5154 deletion mutants that were arranged in quadruplicate in a 1536-format using an automated pinning procedure. Each strain in the knock-out collection is marked with a second selective marker, the kanamycin-resistance gene, which confers resistance to the antibiotic G418. The mated cells were further pinned on plates containing both clonNAT and G418 to select for diploids, before pinning on poor sporulation medium to induce sporulation. After efficient sporulation, a two-step selection was performed to obtain haploid double mutants. The query SGA starter strain has the genotype can1Δ::STE2pr-Sp_his5 lyp1Δ: STE2, which encodes the receptor for α-factor. Thus the STE2 promoter is only activated in MATa cells, which then produce histidine from the *S. pombe* His5 gene and can therefore grow on medium lacking histidine. STE2pr-Sp_his5 replaces the CAN1 gene, which encodes an arginine permease; therefore, in the absence of Can1, cells cannot take up resulting in resistance to the toxic arginine analog canavanine. In addition, LYP1 is deleted, which encodes a lysine permease, thus conferring resistance to the toxic lysine analog thialysine (S-(2-Aminoethyl)-L-cysteine hydrochloride). As diploid cells contain one copy of both the CAN1 and LYP1 genes, they are sensitive to canavanine and thialysine treatment. In the first step, haploid MATa cells were selected on medium lacking histidine and containing canavanine and thialysine; arginine and lysine were excluded from the medium in order increase the toxicity of canavanine and thialysine. In the second step, the haploid MATa selection was maintained and double mutants were selected for by the addition of clonNAT and G418. Upon obtaining double mutant haploids, we pinned these haploids in media containing 0.02% MMS and incubated at 30°C. Plates were imaged after 48h to access colony size. Colony size of *rnh1* or *rnh201* query strains were normalized

to colony size of control query strains, which only contained a ORF deletion from KO library. After, each colony was normalized to its background, we compared growth between different conditions, medium with 0.02% MMS and medium without genotoxic drugs. A colony size 20% smaller or more in the presence of MMS was identified as a hit (Figure 28B).

To dissect genes that may prevent RNase H enzymes function, the same query strains were mated with cells containing overexpression plasmids of 4981 proteins under their endogenous promoters (Figure 28A). Each plasmid contains the LEU2 gene, which allows these cells to grow in media lacking leucine. Followed the diploid selection in media lacking leucine and with clonNAT, the same procedure was performed as before (Figure 28A, B).

To identify RNase H1 regulators that promote RNase H1 activity, we quantified colony size of *rnh201* mutants crossed with KO library in the presence of MMS or not. We identified 17 ORFs that, when deleted, showed significantly smaller colony size in the presence of MMS (Figure 28C; Table 2). Almost all genes identified were mitochondrial genes. Further visual inspection of plates was not possible to identify genes that strongly impaired growth, which could suggest that identified genes could not lead to a phenotype similar to *rnh1 rnh201* double mutants in MMS. Therefore, these genes may not affect RNase H1 activity.

As shown previously, Sen1-AID* *rnh1 rnh201* are highly sensitive to low amounts of auxin due to R-loop accumulation (Figure 6). Thus, we used Sen1-AID* rnh201 crossed with a KO library to identify regulators that stimulate RNase H1 similar to before. We identified 29 and 27 genes that caused smaller colony size in the presence of 100µM and 500µM auxin, respectively (Figure 28G, H; Table 6; Table 7). We found that 18 genes were common between the two auxin concentrations tested (Figure 28I), but none of these ORFs were common with the genes identified previously in the presence of MMS. Although this screen revealed 18 potential RNase H1 regulators, colony size differences were subtle, suggesting that all ORFs may be false positive.
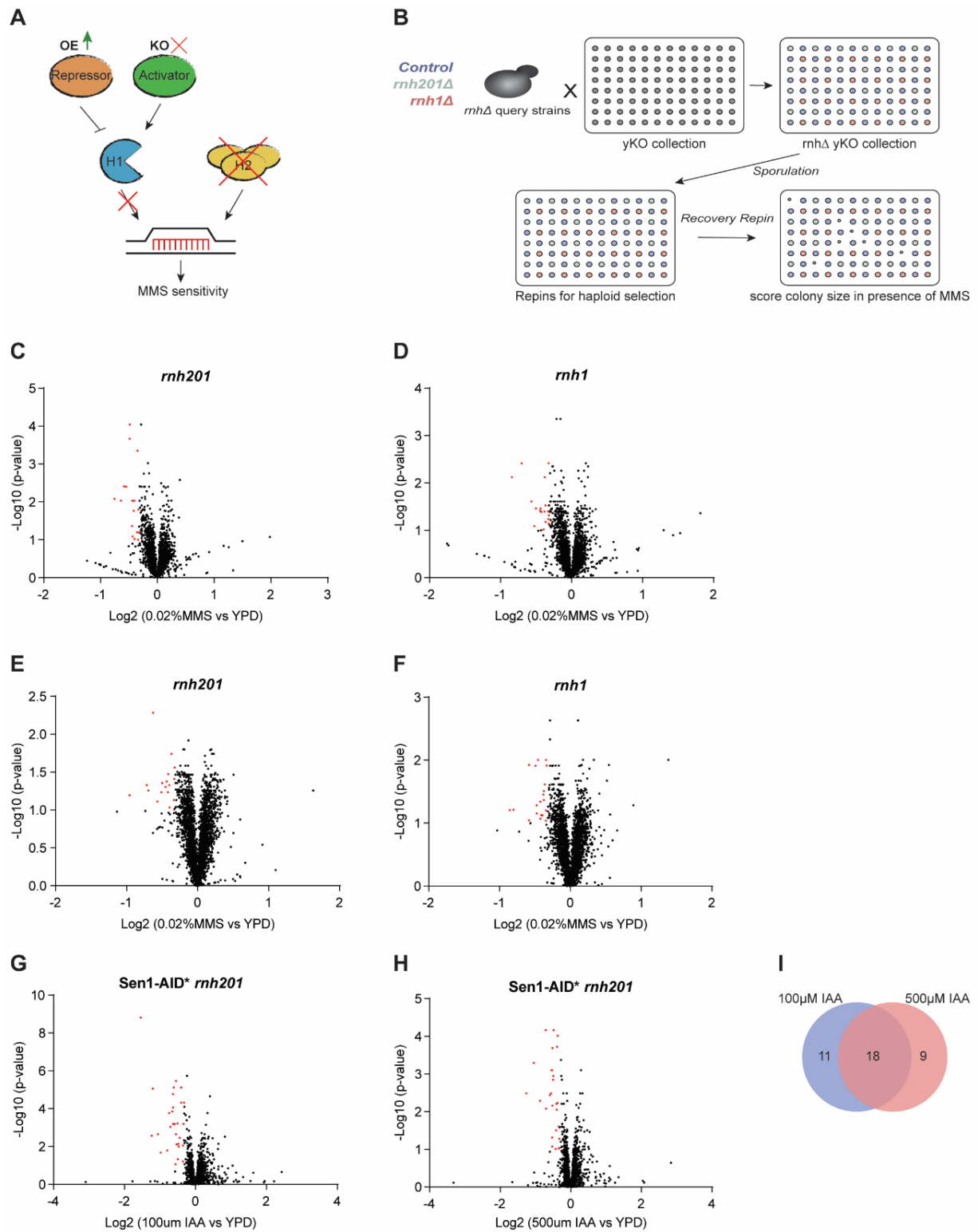
Figure 28 – Genetic screen to identify RNase H regulators. (A) Outline of experiment to identify RNase H1 regulators. Overexpression of repressors or knock-out of activators of RNase H1 would reduce its activity on R-loops. Thus, in rnh201 mutant, this may lead to accumulation of R-loops and high sensitivity to MMS, as observed in rnh1 rnh201 double mutants. **(B)** Outline of the SGA procedure to construct a library of double mutants as described in text. **(C)** Volcano plot of rnh201 mutants crossed with KO library to identify activators of RNase H1. Red dots are ORFs with 20% or more smaller colony size in the presence of 0.02%MMS. **(D)** Volcano plot of rnh1 mutants crossed with KO library to identify activators of RNase H2. Red dots are ORFs with 20% or more smaller colony size in the presence of 0.02%MMS. **(E)** Volcano plot of rnh201 mutants crossed with Overexpression collection to identify

68

repressors of RNase H1. Red dots are ORFs with 20% or more smaller colony size in the presence of 0.02%MMS. **(F)** Volcano plot of rnh1 mutants crossed with overexpression collection to identify activators of RNase H2. Red dots are ORFs with 20% or more smaller colony size in the presence of 0.02%MMS. **(G)** Volcano plot of Sen1-AID* rnh201 mutants crossed with KO library to identify activators of RNase H1. Red dots are ORFs with 20% or more smaller colony size in the presence of 100μM IAA. **(H)** Volcano plot of Sen1-AID* rnh201 mutants crossed with KO library to identify activators of RNase H1. Red dots are ORFs with 20% or more smaller colony size in the presence of 500μM IAA. **(I)** Venn diagram of ORFs from **G** and **H**.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YDR369C	-0.4838	4.04184
YLR069C	-0.4882	3.67074
YER150W	-0.3498	3.35199
YNL005C	-0.5842	2.41168
YER087W	-0.5491	2.40349
YDR418W	-0.7534	2.08035
YOR187W	-0.6407	2.03494
YDR065W	-0.4413	2.03054
YPL024W	-0.404	2.03054
YCL022C	-0.334	1.8
YGR076C	-0.4155	1.76699
YDR115W	-0.4445	1.36403
YIL101C	-0.3572	1.19993
YDR405W	-0.3236	1.19255
YPR024W	-0.4337	1.09051
YKL155C	-0.4017	1.02457
YLR190W	-0.3372	1.00306

Table 2 – SGA hits of rnh201 mutants crossed with KO library (Figure 28C). ORFs with at least 20% smaller, colony size in the absence of RNase H2 upon addition of 0.02%MMS.
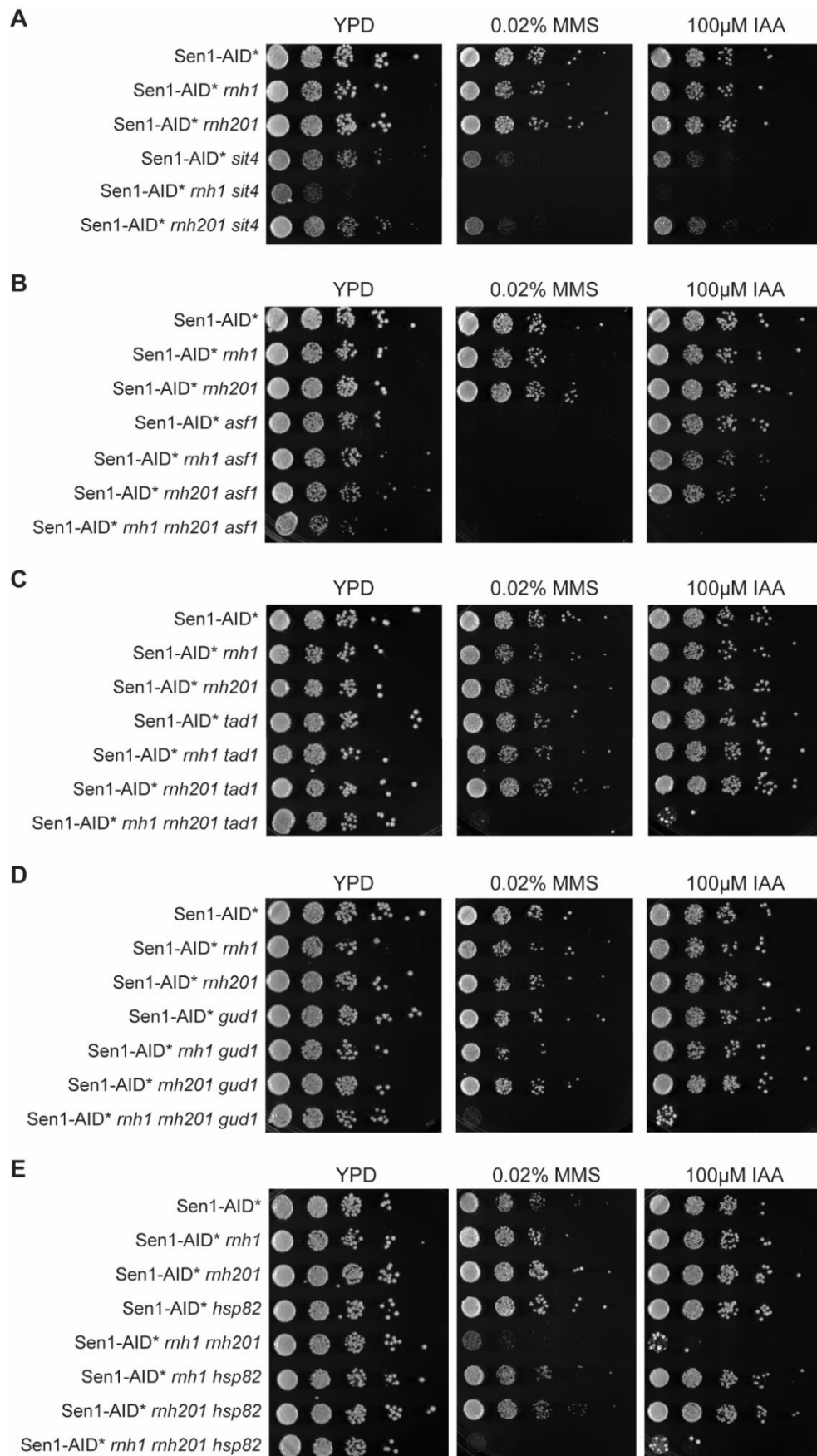
Figure 29 – Confirmation of SGA hits. Cells of the indicated genotypes were spotted in serial dilutions onto YPD, MMS-containing YPD and IAA-containing YPD. The plates were imaged after 48h of incubation at 30°C.

To find proteins that may inhibit RNase H1 we analyzed growth of *rnh201* mutants crossed with an overexpression collection in the presence of MMS (Figure 28E). We show that overexpression of 18 proteins lead to slower growth of *rnh201* mutants in the presence of MMS compared with medium without drugs (Figure 28E; Table 4). Several genes were metabolic-related genes or genes involved in other processes that are not correlated with gene regulation. Similar with previous genetic screen, growth defect was difficult to detect by visual inspection of plates, which may indicate that these genes are false positive.

Next, we crossed the KO library or overexpression collection with *rnh1* mutants to identify proteins that promote or repress RNase H2 activity, respectively (Figure 28A). Using the KO library, we found 20 potential genes with growth defect in MMS (Figure 28D; Table 3). Similar to *rnh201* screen in MMS (Figure 28C; Table 2), majority of genes were mitochondrial genes. Overexpression of 20 genes led to slower growth in *rnh1* mutants upon MMs treatment (Figure 28F; Table 5). In both genetic screens, colony size differences were very small and difficult to visualize.

Since the potential interactors identified in the screens showed small colony size differences compared to control conditions, we confirmed a few genes with spotting assay (Figure 29). Based on our observations, deletion of Tad1, Gud1 or Hsp82 do not affect cell growth of Sen1-AID* *rnh1* or Sen1-AID* *rnh201* in the presence of MMS or auxin (Figure 29C, D, E). Furthermore, deletion of Asf1 causes cell death in the presence of MMS, but removal of Sen1 in *rnh1* or *rnh201* mutants does not affect cell growth (Figure 29B), implying that these genes do not impact RNase H enzymes activity as well. Strikingly, *sit4 rnh1* double mutants are synthetic sick and deletion of Sit4 causes slower cell growth in Sen1-AID* *rnh1* mutants with MMS or auxin (Figure 29A), suggesting that Sit4 impacts RNase H2 activity. However, we did not identify any gene that decreases cell viability of *rnh201* mutants in the presence of high R-loop levels. Collectively, these data lead to the conclusion that further experiments and different approaches are required to dissect RNase H1 regulators.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YJR113C	-0.7022	2.41631
YMR267W	-0.3243	2.41631
YAL065C	-0.8412	2.12317
YDR347W	-0.3768	2.12317
YDR034C	-0.5657	1.60719
YMR097C	-0.3677	1.52585
YMR190C	-0.3238	1.52585
YJL180C	-0.502	1.46322
YNL284C	-0.4366	1.46322
YJR102C	-0.443	1.41449
YLR069C	-0.3293	1.39947
YDR384C	-0.3814	1.39471
YJL096W	-0.4279	1.38707
YNL045W	-0.321	1.31471
YLR265C	-0.3284	1.23656
YIL071C	-0.3658	1.17702
YIL157C	-0.3198	1.14102
YOR158W	-0.3233	1.13707
YDR470C	-0.5257	1.08697
YDL118W	-0.3941	1.0169

Table 3 – SGA hits of rnh1 mutants crossed with KO library (Figure 28D). ORFs with at least 20% smaller colony size in the absence of RNase H1 upon addition of 0.02%MMS.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YEL027W	-0.6282	2.28203
YLL062C	-0.3679	1.74039
YLR042C	-0.3227	1.56039
YIR023W	-0.4149	1.47401
YNL206C	-0.3296	1.40981
YEL017C-A	-0.4363	1.37633
YBR046C	-0.4986	1.35506
YGR254W	-0.7176	1.32793
YMR220W	-0.4	1.32793
YDL122W	-0.4531	1.30351
YFR049W	-0.6941	1.25773
YFR044C	-0.5095	1.23456
YCL050C	-0.4181	1.23395
YDR221W	-0.9598	1.19459
YNL333W	-0.3349	1.12965
YBR028C	-0.5688	1.11053
YFL013C	-0.3964	1.03306
YMR257C	-0.333	1.00494

Table 4 – SGA hits of rnh201 mutants crossed with OE collection (Figure 28E). ORFs with at least 20% smaller colony size in the absence of RNase H2 upon addition of 0.02%MMS.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YNL206C	-0.4558	2.00416
YNL333W	-0.3439	2.00416
YEL027W	-0.583	1.92481
YPR003C	-0.3354	1.9137
YBR046C	-0.4916	1.9119
YNL077W	-0.362	1.61374
YKL222C	-0.3744	1.5093
YLR246W	-0.3811	1.45297
YLR080W	-0.3735	1.36583
YLL019C	-0.4225	1.3419
YIR023W	-0.4769	1.28343
YDR218C	-0.804	1.21087
YFR032C-A	-0.8572	1.20529
YEL017C-A	-0.3436	1.20069
YOR074C	-0.4744	1.1539
YEL042W	-0.4098	1.12662
YKR041W	-0.3943	1.12366
YDR213W	-0.4224	1.0708
YDR179W-A	-0.5844	1.04315
YMR124W	-0.3455	1.03899

Table 5 – SGA hits of rnh1 mutants crossed with OE collection (Figure 28F). ORFs with at least 20% smaller colony size in the absence of RNase H1 upon addition of 0.02%MMS.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YEL036C	-1.5427	8.813
YEL033W	-0.551	5.46589
YEL025C	-0.6133	5.12179
YEL030W	-0.4042	5.12179
YEL023C	-1.2005	5.05393
YPL069C	-0.6348	4.77346
YEL016C	-0.4008	4.3209
YEL006W	-0.3241	4.3209
YEL013W	-0.6313	4.06586
YEL031W	-0.6502	3.85474
YEL020C	-0.7382	3.77119
YEL028W	-0.5099	3.22451
YEL037C	-0.634	3.19478
YEL012W	-0.3344	3.19478
YEL017C-A	-0.5901	3.19111
YMR234W	-0.7093	3.03822
YEL046C	-1.0568	2.65405
YOL036W	-0.5469	2.65405
YEL027W	-1.2335	2.57492
YEL024W	-0.445	2.45905
YER005W	-0.4796	2.1441
YEL038W	-0.5299	2.11238
YEL040W	-0.3522	2.05252
YEL045C	-0.4753	2.00101
YDR364C	-0.7966	1.79328
YLR226W	-0.9795	1.68126
YJL200C	-0.4935	1.3327
YEL047C	-0.3532	1.21464
YLR427W	-0.5617	1.07383

Table 6 – SGA hits of Sen1-AID* rnh201 mutants crossed with KO library (Figure 28G). ORFs with at least 20% smaller colony size in the absence of RNase H2 upon addition removal of Sen1 with 100µM IAA.

ORF	Log$_2$ (0.02%MMS/YPD)	-Log$_{10}$ (p-value)
YEL023C	-0.7108	4.16665
YEL033W	-0.4851	4.16665
YEL042W	-0.3725	4.0137
YEL016C	-0.3904	3.72365
YEL013W	-0.5165	3.68926
YEL036C	-1.0471	3.29448
YEL031W	-0.5452	3.10387
YEL028W	-0.5261	3.10387
YLR096W	-0.4955	2.94983
YEL037C	-0.5003	2.84833
YEL027W	-1.2594	2.48381
YLR382C	-0.5596	2.48381
YOL036W	-0.5937	2.4382
YPL069C	-0.8685	2.28836
YEL024W	-0.3775	2.2328
YER005W	-0.4085	2.20996
YEL038W	-0.5205	2.15874
YEL046C	-0.6982	2.08086
YDR101C	-0.3988	2.05577
YBR266C	-0.3415	1.60698
YEL045C	-0.4164	1.50786
YJL200C	-0.5276	1.31258
YGR205W	-0.3238	1.28052
YMR097C	-0.5285	1.07582
YHR051W	-0.3412	1.03486
YNL254C	-0.33	1.02253
YLR427W	-0.4271	1.01042

Table 7 – SGA hits of Sen1-AID* rnh201 mutants crossed with KO library (Figure 28H). ORFs with at least 20% smaller colony size in the absence of RNase H2 upon addition removal of Sen1 with 500µM IAA.

3 Discussion

3.1 RNase H1 activity: where and when?

3.1.1 RNase H1 binds to accumulated R-loops

R-loops are formed during transcription and can be detected along transcribed genes bodies. However, catalytic inactive RNase H1-based methods of RNA-DNA hybrid capture from previous studies identified R-loops over GC-rich and GC-skewed promoter-proximal pause regions of transcribed genes[20,236]. Thus, it has been proposed that there are two different two classes of R-loops: the promoter-paused R-loops, Class I, and the elongation-associated R-loops, Class II[236]. Furthermore, genome wide approaches suggested that RNase H1 specially associates with Class I R-loops[20]. Mapping of RNA:DNA hybrids in the absence of RNase H enzymes allowed further dissection of potential RNase H1 target loci. Such loci include tRNAs and other RNAPIII-transcribed genes, retrotransposons, mitochondrial genes, other highly transcribed and particularly long genes[20,27,33].

In the present study, we confirm that R-loops accumulate in the absence of RNase H enzymes, specifically in tRNAs (Figure 5). Furthermore, RNase H1 strongly binds to RNPIII-transcribed genes in *rnh1 rnh201* double mutants (Figure 4). Although RNase H1 is recruited to 5S rDNA locus in *rnh1 rnh201* mutants (Figure 4), RNA:DNA hybrid levels measured by DRIP do not increase compared to wild-type cells (Figure 5). Yet, previous studies confirmed that *rnh1 rnh201* double mutants have higher R-loops levels in RNAPIII-transcribed genes, including the 5S rDNA locus[33]. Two possible explanations may account for this difference. In this thesis, DRIP was performed with cross-linked samples, which reduces signal of elongation-associated R-loops[59]. DRIP in cross-linked samples may be favoring capture of promoter-paused R-loops, similar to R-ChIP. R-loops in the 5S rDNA locus were quantified in the gene body, which excluded capture of hybrids present in the promoter region. In order to confirm RNase H1 binding and R-loop levels, it is required to investigate different regions of each locus, such as, the promoter, gene body and terminator regions. Alternatively, DRIP-seq and R-ChIP-seq would allow us to understand genome-wide where RNase H1 binds and its correlation with R-loop levels in a more detailed manner.

During optimization of R-ChIP and DRIP protocols we observed that stabilization of R-loops by catalytic inactive RNase H1 is only captured by DRIP in cross-linked samples and not in native DNA (data not shown), indicating that accumulated R-loops are unstable hybrids with a very high turnover rate, one of the characteristic of promoter-paused R-loops[236].

Catalytic inactive RNase H1 overexpression leads to accumulation of R-loops only in *rnh1 rnh201* double mutants in RNAPII and RNAPIII-transcribed genes (Figure 5), suggesting that RNase H1 targets these genes. However, such R-loop stabilization does not occur in wild-type cells. This might be due to the presence of active RNase H2, which has been described to provide the majority of the RNase H activity in the cell[205]. Since R-loops levels increase only when both RNase H enzymes are absent, it is consider that RNase H1 and RNase H2 have redundant functions and compete to the same subset of R-loops, including tRNA genes[33,237]. This hypothesis is supported by our data demonstrating that wild-type cells do not accumulate R-loops (Figure 5). Quantification of stabilized R-loops by defective RNase H1 in *rnh1* and *rn201* single mutants could identify to which extent each enzyme is capable of suppressing R-loops individually.

Although RNase H1 is overexpressed at lower levels in G1-arrested cells compared to cycling cells (Figure 7), loci with accumulated R-loops in G1-phase show increased RNase H1 binding compared to cells in exponential phase (Figure 9). These results indicate that RNase H1 binds to these loci upon R-loop accumulation, independently of its expression levels and also independently of DNA replication. Loci that accumulate R-loops in G1-arrested cells, such as SUF2 and RPL15a genes, show R-loop stabilization by catalytic inactive RNase H1 (Figure 9). However, the actin locus, which does not have increased R-loop levels in G1-phase, does not accumulate R-loops when RNase H1 (D193N) is overexpressed (Figure 9), supporting the hypothesis that RNase H1 preferentially binds to some loci and specifically when R-loops accumulate.

Overall, we confirm that RNase H1 binds to different loci with different affinities (Figure 4)[27,33] and endogenous levels of RNase H1 maintain R-loop homeostasis (Figure 14). Moreover, cells in G1-phase show strong RNase H1 binding to accumulated R-loops, while RNase H1 is less expressed. Therefore, we propose a model in which RNase H1 acts in a dosage-independent manner. Strikingly, when analyzing *rnh1 rnh201* double mutants' viability in toxic conditions, RNase H1 or its truncation mutants exhibit a dosage-dependent activity to remove R-loops (Figure 21) or to stabilize R-loops and cause toxicity (Figure 22). These disparities might be explain by the conditions used and RNase H1 affinity for hybrids. In detail, RNase H1 has a strong affinity towards RNA:DNA hybrids (Figure 25)[72,219,221], which might explain its recruitment in conditions of accumulated R-loops even when protein levels are low. However, the same affinity will lead to R-loop binding in a dosage-dependent manner and other mechanisms might be required to limit its activity. In the presence of toxic conditions, such as MMS, higher R-loop loads require more RNase H1, supporting further the dosage-dependent activity of RNase H1. Based on a previous study, ectopically expressed RNase H1 associates with multiple R-loops, but it does not actively process the majority of them[205].

Consistently, our findings validate this observation, in which RNase H1 is preferentially recruited to loci with R-loops, but its binding affinity and activity seem to differ in different loci.

3.1.2 RNase H1 targets tRNA genes

The unusual high rates of transcription initiation by RNAPI and RNAPIII may be facilitated by negative supercoiling and strand separation over the promoter regions[33,216]. In the absence of Top1 and Top2, negative supercoiling accumulates in these promoter regions, which leads to increased transcription initiation of tRNA genes due to promoter opening[33]. However, accumulation of hybrids in the absence of RNase H1 offsets the increase of pre-tRNA levels by impaired elongation due to stable R-loops[33]. Here we confirm that RNase H1 binds to tRNA genes to maintain proper tRNA transcription levels. Although not confirmed, overexpression of RNase H1 likely leads to increased tRNA levels, similar to deletion of Top1 and Top2.

Transcription termination is a multistep process consisting of the recruitment of termination factors, the recognition of sequence motifs, RNAP pausing, and release of the RNAP and the transcript from the DNA. RNAPI and RNAPII require extrinsic protein factors to terminate transcription, such as Rat1 exonuclease and the helicase Sen1[238]. In case of RNAPIII, efficient termination is important for its rapid recycling for new cycles of transcription to occur. Thus, multiple mechanisms cooperate to promote the termination of RNAPIII transcription. T-tracts are strictly required for RNAPIII termination, but adjacent RNA structures are important auxiliary elements when the length of the T-tracts falls outside of the optimal range[188]. In some instances, RNA hairpins can form within the RNA exit channel of RNAPIII and increase efficiency of the primary termination, while Sen1 preferentially functions at downstream regions to remove read-through polymerases, independently of transcription-replication conflicts[188]. Although RNase H1 targets both tRNA genes, SUF2 and SUF11 (Figure 4), it is recruited more efficiently to SUF2 and loss of RNase H enzymes causes more R-loop accumulation compared to wild-type cells in the SUF2 locus. Since all tRNA genes terminators show different features, like different T-tracts lengths and spacing, SUF2 and SUF11 might show different transcription termination mechanism. For example, it is possible that SUF2 does not contain efficient primary termination and requires Sen1, while SUF11 does not. Hence, understanding the correlation between tRNA terminators and RNase H1 recruitment would allow us to uncover deeper insights of RNase H1 regulation.

When RNA polymerases backtrack[239], anterior R-loops form ahead of the backtracked RNA polymerase[43]. Restarting transcription requires realigning the 3'end of the RNA with the active site, which can be achieved by either 'sliding' of the enzyme along DNA (1D diffusion)

or endonucleolytic cleavage of the backtracked RNA to generate a new 3' end aligned with the active site. In the case of RNAPIII, its subunit Rpc11 does RNA 3' cleavage to resume transcription after backtracking. While RNA 3' cleavage activity is nonessential, RNAPIII termination requires backtrack on the template to form an RNA hairpin or other form of RNA duplex in the transcript, which is then cleaved[240,241]. Therefore, Rpc11 can provide termination rescue activity and facilitates recycling by resolving arrested RNAPIII. It is possible that RNase H1 and RPA bind effectively to anterior R-loops, particularly in tRNAs and other RNAPIII-transcribed genes.

Overall, RNase H1 targets R-loops in tRNA genes very efficiently, but its effects on RNAPIII transcription initiation and termination remains to be uncovered. Thus, it would be important to understand RNase H1 interaction with Sen1 and RNAPIII, in particular with the subunits that promote RNA cleavage of backtracked polymerases (Figure 30).
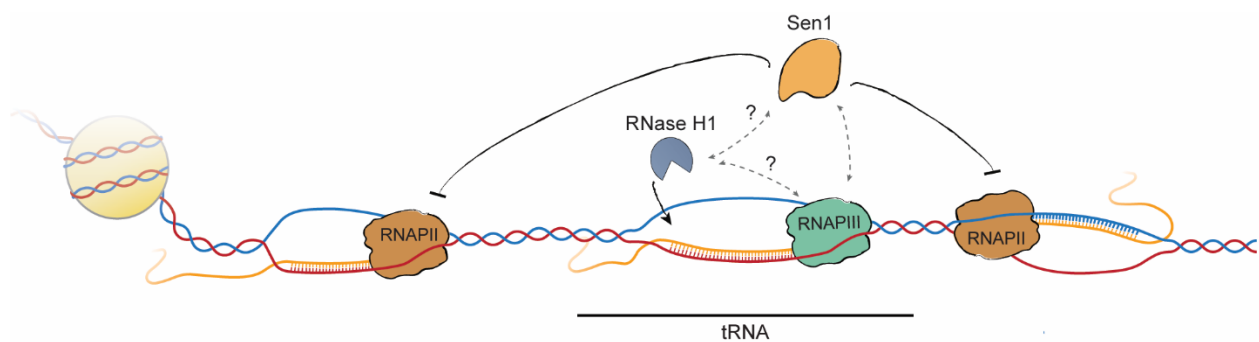
Figure 30 – RNase H1 targets R-loops in tRNA genes. RNase H1 binds and removes accumulated R-loops in tRNAs. Sen1 interacts with RNAPIII to promote efficient transcription termination as well as preventing conflicts between RNAPII and RNAPIII. RNase H1 effect on RNAPIII transcription termination remains unclear.

3.1.3 Does RNase H1 remove R-loops in rRNA genes?

Although RNAPI and RNAPIII are the only enzymes that directly mediate rRNA expression, a recent study in human cells has found that RNAPII localizes near rRNA-encoding genes to drive their expression[242]. RNAPII binds to intergenic spacers between rRNA genes and generates an R-loop shield that prevents intergenic transcription by RNAPI transcription[242]. However, in yeast, intergenic transcription does not regulate rRNA[238] and Sen1 limits ncRNAs from intergenic regions by enforcing epigenetic silencing and transcript turnover[243]. Aiello et al. have shown that yeast Sen1 is important to resolve transcription-replication conflicts at the rDNA replication fork barrier (rRFB) and release RNAPII at the rDNA loci[237]. Using RNase H enzymes to map R-loops, R-loop formation is detected upstream of

RNAPII in the rRFB whereas antisense R-loops are formed upstream 5S rRNA and in spacer regions between other rRNA genes. These R-loops are associated with RNAPII in these loci and deletion of both RNase H enzymes leads to accumulation of such R-loops. Strikingly, RNase H enzymes and Sen1 cooperate to limit RNAPII transcription and promote its release from the rDNA loci[237]. Probably, RNase H digestion would allow Sen1 or Rat1 to access the nascent RNA close to a stalled RNAPII and induce termination[237,244].

In the 18S rRNA gene, we show that elongation-associated R-loop levels are higher in wild-type cells than *rnh1 rnh201* double mutants (Figure 5). Moreover, overexpression of RNase H1 (D193N) in wild-type cells seem to decrease R-loop levels (Figure 5). Reduced R-loop levels in conditions that are expected to accumulate them imply that R-loop stabilization leads to destabilization of elongation-associated R-loops in rRNA genes. One possible explanation would be that RNase H enzymes target R-loops in rRNA promoters and rDNA spacer regions (Figure 31), which are not studied in this thesis. Indeed, similar to tRNA transcription[33], elevated R-loop levels at rDNA promoters lead to reduced rRNA transcription in human cells[245]. R-loop resolution is required to induce RNAPII termination in rDNA loci and to prevent conflicts with RNAPI[237]. RNAPII elongation in the rDNA loci would impact RNAPI transcription and, consequently, the formation of elongation-associated R-loops as we observed (Figure 5). Hence, it remains to be understood in which other regions of the rDNA locus RNase H1 is binding and actively removing R-loops.

Genome-wide mapping of RNA:DNA hybrids shows that Sen1 mutants have increased R-loop levels in the rDNA region compared to wild-type cells, *hpr1* and *rnh1 rnh201* mutants[27]. In yeast, Sen1 might be the main enzyme that regulates R-loop levels in rDNA loci by preventing their formation, while RNase H1 binds specifically to non-coding rDNA regions and 5S rRNA gene[237]. Even though we detect high R-loop levels in the 18S rDNA locus (Figure 5), RNase H1 is recruited similarly to other loci that present less R-loops measured by DRIP, such as SUF11 and 5S rDNA (Figure 4), indicating that RNase H1 targets this gene less efficiently. It is important to note that we measured R-loop levels and RNase H1 recruitment only in the gene body of the 18S rRNA gene. Thus, additional experiments are required to further understand the role of RNase H1 in the rDNA loci, such as RNase H1 recruitment to other regions of the rDNA loci and interaction with Sen1 in this context (Figure 31).

Loss of Top1 is known to slow down or impede RNAPI elongation, which is also associated with increased formation of R-loops in the nucleolus[33,246]. In human cells, SETX inhibition is another factor that can lead to R-loop accumulation. Moreover, RPA co-localizes with R-loops that are accumulated at rDNA promoter regions due to Top1 or SETX inhibition[245]. These stabilized R-loops generate ssDNA that allow RPA to bind, which in turn recruits and activates RNase H1, facilitating R-loop resolution[235]. Taken together, it is suggested that

RNase H1 is only recruited to rDNA in the absence of other R-loop-resolving pathways and may act as a backup mechanism to process R-loops at rRNA genes. However, as it was mentioned above recent findings have found RNase H1 activity to be essential to limit RNAPII accumulation in the rDNA[237]. Nevertheless, it is crucial to consider RNase H1 activity in rDNA regions in the absence of other R-loop-resolving proteins.
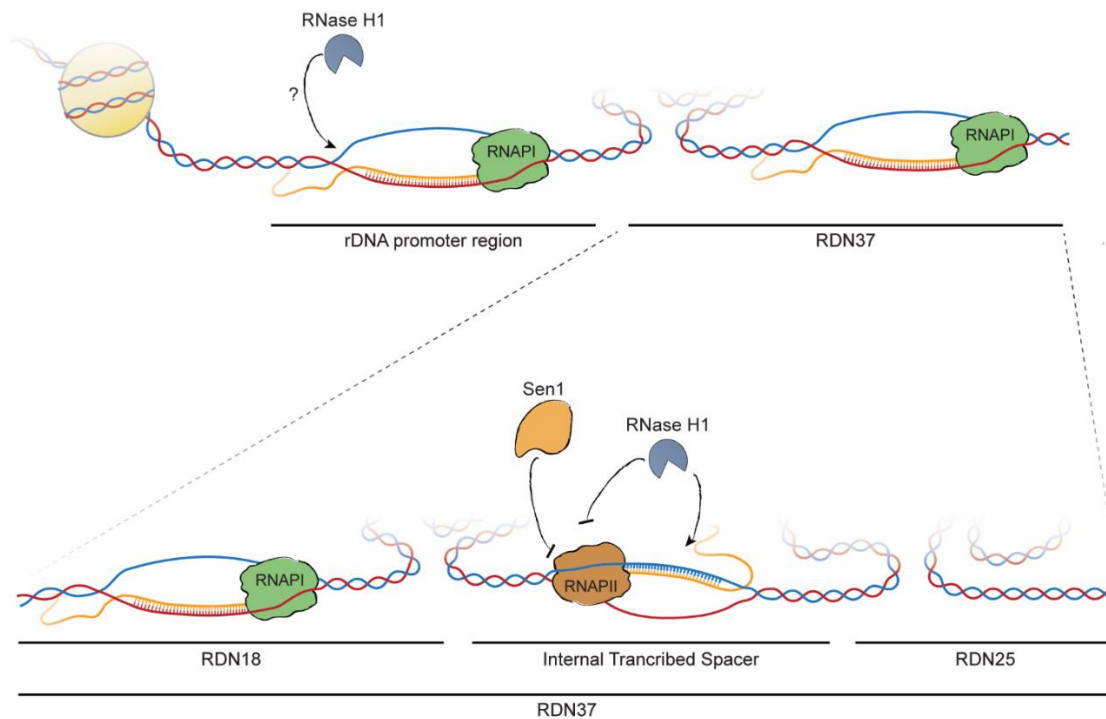
Figure 31 – RNase H1 removes antisense R-loops between rRNA genes. RDN37 represents the primary 35S transcripts, which contains 25S (RDN25), 18S (RDN18) and 5.8S (RDN58, not represented in this scheme) rRNAs. Between 25S, 18S and 5.8S rRNA genes, an internal transcribed spacer is transcribed as part of the 25S rRNA precursor transcript. Antisense R-loops are formed in internal transcribed spacers and can be removed by RNase H1, which helps Sen1 promoting RNAPII release from rDNA loci .It is still unknown if RNase H1 targets promoter-paused R-loops in the rDNA locus.

3.1.4 RNase H1 activity is independent of replication

In the absence of Sen1, RNase H1 expressed only in G1-, S- or G2/M-phase rescues the synthetic lethality of *rnh1 rnh201* (Figure 6), indicating that RNase H1 can act irrespective of cell cycle stage when R-loops accumulate. Although S and G2 alleles of RNase H1 can rescue *rnh1 rnh201* mutants sensitivity to MMS, RNase H1 expressed only in G1-phase does not fully rescue growth in MMS-containing medium (Figure 6). Treatment with MMS leads to increased R-loop levels[76], but when MMS-mediated damage occurs causing R-loop accumulation might explain the slower growth of G1 allele of RNase H1. Previous studies

propose that alkylation damage caused by MMS induces fork stalling at yeast replication origins[247–249]. Hence, cells respond to MMS-induced DNA damage within S-phase, when G1 allele of RNase H1 is not expressed. G1 allele uses a Sic1 promoter, which promotes low expression levels in S- and G2-phase, allowing the partial rescue observed in MMS. Collectively, RNase H1 activity is important for cell survival when replication fork stalling occurs due to MMS treatment.

Cell cycle alleles of RNase H1 are expressed more than RNase H1 under endogenous promoter[76]. It is possible that this overexpression can lead to removal of R-loops due to high expression levels that avoid temporal limitations imposed by the cell cycle alleles. In order to confirm the presence of RNase H1 activity in all cell phases, additional experiments are required, such as DRIP and R-ChIP of synchronized cells in different cell cycle stages.

In G1-arrested cells, we show that some loci, SUF2 and RPL15a, accumulate R-loops in comparison with non-synchronized *rnh1 rnh201* mutant cultures. Similarly, R-loop levels increase in the 18S rRNA genes in G1-arrested wild-type cells. A previous work about other R-loop-resolving proteins supports that THO transcription complex prevents R-loop formation in G1- and S-phase, whereas Sen1 helicase prevents them only in S-phase[250]. Furthermore, cells with Sen1 temperature sensitive allele accumulate hybrids in rRNA, tRNA genes and other highly transcribed genes, suggesting that Sen1 targets R-loops at these genes[27]. Accordingly, the absence of Sen1 activity in G1-phase may cause accumulation of R-loops observed in G1-arrested cells (Figure 9, Figure 11). Nevertheless, RNase H1 overexpression decreases R-loop levels in all loci detected with accumulated hybrids during G1-phase (Figure 9, Figure 11). Altogether, we confirm that RNase H1 is recruited to and removes accumulated R-loops outside of a DNA replication context.

3.1.5 RNase H enzymes maintain R-loop homeostasis

RNase H1 is a very low expressed protein[251]. Because the overexpression of RNase H1 promotes R-loop degradation, it has been proposed that a repressor may need to be inhibited in order to allow RNase H1 activity and that this repressor is titrated out upon overexpression[205]. Currently, little is known about recruitment of endogenous RNase H1 to RNA:DNA hybrids.

Cells with endogenous defective RNase H1 shows sensitivity to MMS and high R-loop levels only in the absence of RNase H2, indicating that defective RNase H1 behaves very similarly to RNase H1 deletion (Figure 12B). While endogenous catalytic inactive RNase H1 does not cause R-loop accumulation, loss of RNase H2 leads to high R-loop loads in SUF2

locus compared to wild-type cells (Figure 14), implying that RNase H2 is the main enzyme responsible to maintain low R-loop levels in the cell. However, it has previously been reported that RNase H1 deletion also results in increased R-loop levels at tRNA genes[33], suggesting that the endogenous catalytic-dead RNase H1 may be similar but not identical to RNase H1 deletion. Furthermore, we provide evidence that deletion of both RNase H enzymes causes higher R-loop accumulation than RNase H single mutants in RNAPIII-transcribed genes (Figure 14), showing that both enzymes play contribute to R-loop homeostasis. Consistently, RNase H1 binds to loci in an R-loop-dependent manner (Figure 13).

Similar to ectopically expressed RNase H1, endogenous RNase H1 is less expressed in G1-phase (Figure 12A). However, endogenous RNase H1 is still recruited in G1-phase to loci with higher R-loop levels, such as SUF2 and SUF11 (Figure 13, Figure 14), confirming that RNase H1 binds to R-loop independently of its expression levels. As it has been demonstrated before, G1-arrested cells have more R-loops than cycling cells, further supporting that other R-loop removal pathways are involved in R-loop resolution in these loci. Overall, we show that endogenous RNase H1 maintains reduced R-loop levels, including outside of S-phase.

3.2 RNase H1 interacts with RPA

In human cells, RNase H1 was identified as an interactor of ssDNA-coated RPA[232]. Following that, another study showed that RPA stimulates RNase H1 activity *in vitro* and enhances the association of RNase H1 with RNA:DNA hybrids. Indeed, RPA binding-defective RNase H1 fails to recognize and suppress R-loops *in vivo*[235]. In yeast, an interactome analysis has identified that RNase H1 interacts with Rfa1 and Rfa3, two subunits of RPA. Consistent with a previous study[252], we also confirm that RNase H1 interacts with RPA in a DNA-dependent manner (Figure 13). This might suggest that RPA needs to associate with ssDNA, such as the displaced strand of R-loops, to interact with RNase H1. Previous studies demonstrated that RPA has different arrangements when binding to ssDNA depending in part by the conformation of the bound ssDNA[253,254]. These variations caused by different ssDNA conformations changes the accessibility of a major protein-protein interaction site, especially if the binding involves a structural domain of the binding partner[253]. Moreover, RPA binds to ssDNA with different modes and induces changes in ssDNA depending on its length[255]. Indeed, RPA binds to the DNA fork differently than it binds to the duplex DNA with 5' ssDNA. In particular, RPA binding to the fork structure is highly dynamic when the 3'ssDNA is between 10 and 30 nt[255]. Thus, the substrate that RPA binds might affect the interaction with other proteins, including RNase H1. It is possible that the length or structures formed in the displaced strand of the R-loop impact this interaction. To test this hypothesis, we need to confirm if RPA-

RNase H1 interaction occurs through the displaced ssDNA of R-loops. For this, Co-IP of RPA has to be performed upon treatment of S1 nuclease, which cuts the ssDNA displaced strand. Additionally, to verify that RPA-RNase H1 interaction is not dependent on RNA, we need to prove that all RNA is degraded upon RNase A addition in the Co-IP assay.

The loop between β1 and β2 of the RNase H1 HB domain (Figure 2) has a different position in the crystal of the HB domain complex with RNA-DNA compared to the structure of the HB domain alone, which implies that it may change conformation when interacting with RNA:DNA hybrids[72,191]. Since RNase H1 and RPA still interact upon degradation of RNA:DNA hybrids by bacterial RNase H (Figure 15), we can probably exclude that such rearrangements of the HB domain of RNase H1 promotes interaction with RPA.

In human cells, RNase H1 interacts with RPA through its HB domain, specifically through R57 residue, since the mutation of this residue abolishes the interaction with RPA[233]. In yeast, we show that RPA interacts with the first HB domain of RNase H1 (Figure 18). However, cells with two HB2 domains display a weak interaction, which might occur partially due to the conformation of RNase H1 when it has two HB domains or RPA interacts with the linker domain between HB domains. The residue R57 in human RNase H1 is not conserved in *S. cerevisiae*. In fact many of the positively charged residues adjacent to the RNA:DNA-binding pocket of the HB domain are not conserved, except for R11 and K12 (R32 and R33 in human) that were shown to not affect RPA-RNase H1 interaction in human cells. However, considering the structure of the HB domain in yeast (Figure 2), R11, K12, R14 residues are potential targets for studying RPA-RNase H1 interaction. Additionally, mutation of the FKKF motif in the positively charged groove, which interacts with the DNA strand of hybrid, might provide additional information of how RNase H1 interacts with DNA and RPA.

Regarding this interaction, we also show that RNase H1 interacts with RPA in all cell cycle phases (Figure 16). Together with the hypothesis that RNase H1 and RPA interact through common substrate, our data support our previous observations on which RNase H1 binds to and removes R-loops in all cell cycle phases.

3.3 Characterization of RNase H1

3.3.1 RNase H1 requires one HB domain to remove R-loops

A single copy of the N-terminal HB domain is commonly found on eukaryotic RNase H1. However, it is still unclear why *S. cerevisiae* RNase H1 contains two HB domains. The first copy of HB domain was shown to be essential to bind to RNA:DNA hybrids and dsRNA, while

loss of the second HB domain slightly reduces affinity to substrates[218,219]. These studies suggest that HB2 domain does not bind to hybrids partly due to the insertion of three amino acids that are absent from the first HB domain[218]. Notably, experiments with only RNase H1 second HB domain were never shown in order to access directly its function. In the present study, we address directly the effect of both HB domains in RNase H1 affinity to RNA:DNA hybrids *in vivo* and *in vitro*. We show that removal of one HB domain leads to weaker binding to tRNA genes compared to full length RNase H1 and removal of both HB domains abolishes RNase H1 enrichment in the SUF2 tRNA (Figure 19B). Although truncation mutants do not bind efficiently to tRNAs genes, R-loop removal remains sufficient (Figure 19C). On the other hand, catalytic-dead truncation mutants with weak binding to hybrids do not stabilize R-loops (Figure 19D). Together, these results suggest that RNase H1 requires both HB domains to bind and process efficiently R-loops.

RNase H1-2xHB1 overexpressed with a Gal promoter leads to strong binding and stabilization of R-loops similar to wild-type RNase H1 (Figure 19B, D), which is not observed when this mutant is overexpressed with the GPD promoter (Figure 20B, D). While all truncation mutants were expressed to similar levels under a GPD promoter, RNase H1-2xHB1 overexpressed with a GAL promoter led to much higher protein levels than any other construct (Figure 17B), indicating that R-loop binding and stabilization observed likely was due to the presence of high protein levels in this mutant. Indeed, stabilization of R-loops by RNase H1 (D193N) leads to viability loss in *rnh1 rnh201* mutants in a dosage dependent-manner (Figure 22C, D). Therefore, RNase H1 (D196N)-2xHB1 expressed by GAL promoter does not reflect a true portrait of this mutant's action on R-loop binding and stabilization.

In the presence of two HB1 or HB2 domains (RNase H1-2xHB1 or RNase H1-2xHB2 respectively), RNase H1 has less affinity for accumulated R-loops in tRNA genes compared to full length RNase H1, which leads to lack of over stabilization of R-loops when overexpressing the catalytic-dead version of these mutants (Figure 19D). However, similar to HB domain truncation mutants, overexpression of catalytic-active versions reduces R-loop levels (Figure 19C). Overall, these results indicate that both HB domains contribute in a unique way for R-loop binding and removal. Likely, the first and second HB domains seem to have different functions.

Enrichment of catalytic-active RNase H1 cannot be used to detect R-loops and RNase H1 binding due to active degradation of the hybrid[256]. Thus, ChIP of RNase H1 shows low enrichment even in loci that RNase H1 binds, like tRNA genes (Figure 20A). Interestingly, RNase H1 mutants without HB2 domain, such as RNase H1-ΔHB2 and RNase H1-2xHB1, are enriched more than full length RNase H1 or other mutants. Although RNase H1-2xHB1 reduces R-loop levels similarly to RNase H1-2xHB2, RNase H1-ΔHB2 does not remove R-

loops like full length RNase H1 and other truncation mutants (Figure 20C). Indeed, overexpression of RNase H1-ΔHB2 does not rescue *rnh1 rnh201* double mutant growth in the presence of MMS in a dosage-dependent manner (Figure 21). Thus, RNase H1 without a second HB domain may not remove R-loops efficiently, allowing detection of its catalytic active version. Another possible explanation is that RNase H1 without HB2 domain stays longer bound to SUF2 locus and HB2 domain is required for RNase H1 dissociation from the locus after removing R-loops. In *Drosophila melanogaster*, RNase H1 contains two HB domains, in which the first HB domain is important for RNA:DNA hybrid binding, while the second HB domain enhances RNase H activity and appears to support an efficient turnover rate[257]. Detection of catalytic active RNase H1-ΔHB2 (Figure 20A) might demonstrate a decreased rate of dissociation of the enzyme-substrate complex, thus implying that the second HB domain may be relevant for effective substrate release. Given that RNase H1 has to progress along the substrate, the binding properties of HB2 domain might be relevant to the translocation process[257]. Similar to *D. melanogaster*, yeast HB2 domain might be important to modulate RNase H1 activity by enhancing RNase H activity and dissociation from the substrate after catalysis.

As described before, MMS leads to increased R-loop levels[76] and damage related to S-phase[247–249]. Dosage-dependent RNase H1-ΔHB2 activity in rescuing *rnh1 rnh201* upon MMS conditions could suggest that the second HB domain is also required for RNase H1 to act in S-phase. Further experiments with RNase H1 truncations in cells arrested in S-phase are required to confirm this hypothesis.

3.3.2 RPA does not promote RNase H1 activity in yeast

Human RPA interacts with RNase H1 HB domain and promotes its recruitment and activity on R-loops. Moreover, interaction with RPA is essential for RNase H1 to detect and remove R-loops *in vivo*[233]. In yeast, RNase H1 interacts with RPA mainly through its first HB domain, but RNase H1 lacking interaction with RPA is recruited to and removes R-loops equally to other RNase H1 truncation mutants that still interact with RPA (Figure 19B, C). Therefore, in yeast, RPA is not essential for RNase H1 activity when the latter is overexpressed. Since RNase H1 has a strong affinity to RNA:DNA hybrids[72,211] and its overexpression strongly promotes R-loop degradation, it would be valuable to understand if endogenous levels of RNase H1 truncation mutants could maintain low R-loop levels similar to wild-type cells and if RPA interaction impacts endogenous RNase H1. Additionally, R-ChIP experiments involving removal of RPA would confirm the interaction effects on RNase H1 activity *in vivo*.

When RNase H1 truncation mutants are overexpressed, RPA is still recruited in conditions with elevated R-loop levels (Figure 19E, F), implying that RPA binding to R-loops is not affected by RNase H1 mutants lacking interaction with RPA, such as RNase H1-ΔHB1. This would be in agreement with both models in which RPA does not affect RNase H1 recruitment or RPA would be recruited to R-loops first and in turn impact on RNase H1 recruitment.

3.3.3 RNase H1 binds specifically to RNA:DNA hybrids

As previously reported, we confirm that RNase H1 binds strongly to R-loops and RNA:DNA hybrids compared to other substrates *in vitro*. HB domain of RNase H1 has been described to bind to hybrids and dsRNA[72,218,219]. Human RNase H1 has 25-fold preference for RNA:DNA compared to the same sequence dsRNA[72]. In yeast, we show that RNase H1 binds to hybrids over 100-fold better than to dsRNA (Figure 25). Indeed, we could not find a preference for dsRNA over other non-hybrid substrates, like ssRNA, ssDNA and dsDNA (Figure 25). A previous study has identified *S. cerevisiae* RNase H1 interaction with RNA:DNA duplexes, dsRNA, ssRNA, ssDNA and dsDNA *in vitro*, in which the preference for binding is dsRNA > RNA:DNA hybrids > ssDNA > ssRNA > dsDNA[219]. The differences observed between existing data and this thesis may be explained by the technique used, as the FP assay offers much higher sensitivity than the qualitative technique used before[218,219]. Moreover, differences in length and sequence of the used substrates might lead to formation of secondary structures that may affect the results observed.

RNase H1 HB domain was initially characterized as a dsRNA-binding domain[219]. Both HB domain and dsRNA binding domain (dsRBD) include three β-sheet and two α-helices, but the topology is different[222,258]. As previously identified[219], we could not observe a preference of dsRNA compared to ssRNA, ssDNA or dsDNA. This might be due to the differences in methods and buffers used for the *in vitro* assays. Thus, further tests with other salt concentrations would be valuable to confirm the differences observed between studies. Alternatively, differences in the purification method used to isolate RNase H1 might confer slight differences to RNase H1 conformation or to HB domain topology, impacting the binding to dsRNA.

Nucleic acid binding by HB domains is independent of catalysis and metal ions[72]. However, it is important to note that high salt concentration is required to overcome non-specific charge-charge interactions of HB domain and nucleic acids[72,219,221]. Indeed, high amounts of NaCl or Mg^{2+} ions reduces RNase H1 binding to ssDNA, dsDNA, ssRNA and

dsRNA[72,219], which we also observed during the optimization of the FP assay (data not shown). Although our NaCl concentrations are similar to physiological levels, it is possible that interactions with ssDNA, dsDNA, ssRNA and dsRNA in different substrates are non-specific charge-charge interactions.

RNase H1 binds stronger to R-loops than linear RNA:DNA hybrids (Figure 25), suggesting that RNase H1 might interact with other elements of the R-loop, such as the displaced ssDNA. Indeed, substrates with an ssDNA or ssRNA overhang always show a stronger binding compared to the same substrates without such overhangs (Figure 25). It is enticing to speculate that RPA occupancy[254] or structures, such as G-quadruplexes, formed in the displaced strand of R-loops might affect RNase H1 activity or recruitment to R-loops (Figure 32). Moreover, these differences in the binding properties of RNase H1 to hybrids with or without ssDNA displaced strand may suggest different functions *in vivo*, such as degradation of stable R-loops that form during transcription, instead of removal of DNA replication intermediates. Alternatively, it is possible that RNase H1 interaction with the displaced ssDNA is non-specific as discussed before[72,219,221]. To confirm this, it would be required to compare the affinity to different hybrids in more NaCl and $MgCl_2$ concentrations. In addition, it would be informative to study the catalysis of the multiple RNA:DNA substrates tested in this thesis.

As observed previously[221,257], EMSA assays show more than one mobility shift of R-loop and RNA:DNA hybrid substrates at high RNase H1 concentrations (Figure 26A, Figure 27A), demonstrating the formation of higher-order complexes, likely more than one protein moieties with one substrate. Overall, these results confirm the RNase H1 preference for RNA:DNA hybrid substrates.

Although previous studies suggested that RNase H1 second HB domain does not bind to RNA:DNA hybrids[218,219], our *in vivo* data indicates that RNase H1 requires both HB domains to bind and remove R-loops efficiently (Figure 20B,C), confirming that the second HB domain plays a major role in RNase H1 activity. Consistently, deletion of one HB domain decreases RNase H1 affinity for RNA:DNA hybrid substrates *in vitro* (Figure 23). Furthermore, we observe less "super-shifting" compared to full length RNase H1 by EMSA analysis (Figure 26A, Figure 27A), which indicates that multiple RNase H1-ΔHB1 or RNase H1-ΔHB2 monomers do not associate with a single substrate at the same time, unlike wild-type RNase H1. In turn, this may lead to decreased processivity of the RNase H1 mutants[221,257], although the catalytic activity of these mutants must be confirmed *in vitro*. Altogether, our data confirms that two copies of HB domain bind to nucleic acids better than one, even though one is sufficient for binding[218]. However, *S. cerevisiae* HB domains are not equal. In particular, removal of HB2 domain impairs more RNase H1 interaction with hybrids than lack of HB1 domain. This supports the *in vivo* observations, based on which RNase H1-ΔHB2 removes R-loops less

efficiently (Figure 20C) and does not completely rescue *rnh1 rnh201* double mutant growth in MMS (Figure 21A). In *D. melanogaster*, the second HB domain promotes dissociation from the substrate, but is required for RNase H activity[257]. Thus, it has been proposed that HB2 may be involved in shuttling substrate from HB1 domain to the catalytic domain, as a part of the processivity mechanism[257]. Further experiments that detail association and dissociation rates of protein-substrate complex are required to confirm this model in yeast.

Based on the literature, the lack of HB2 domain binding to RNA:DNA hybrids is attributed to the fact that it contains 3 additional amino acids, which are not present in HB1 domain[218] (PNI in Figure 2). Nevertheless, in this study we observed that removal of HB2 domain is more detrimental for RNase H1 recruitment and activity across *in vivo* and *in vitro* experiments, similar to *D. melanogaster* RNase H1[257]. These differences might be caused by the use of the recombinant proteins and/or specific techniques applied in our and prior studies. Previously, HB domain binding to RNA:DNA hybrids was performed with purified RNase H1 lacking the catalytic domain and mutations on this recombinant protein[218], while we purified RNase H1 mutants that maintain the catalytic domain. Thus, former studies could not exclude the possibility that HB2 properties are hindered due to conformational changes when the catalytic domain is absent[218]. Indeed, they show that HB1 domain requires the linker domain in order to maintain its binding to substrates, possibly, due to recombinant HB1 domain without linker domain suffering of conformational changes. Additionally, a mutant containing only HB2 domain to confirm its function in the absence of HB1 domain has never been generated. Here, we show that mutants with HB2 domain, such as RNase H1-ΔHB1, can bind to its substrates, confirming that HB2 domain interacts with RNA:DNA hybrids (Figure 23). At the same time, we cannot exclude that RNase H1-ΔHB2 mutant does not have conformational changes in its HB1 domain due to the deletion of the internal region that contains the linker domain and HB2 domain, and causing RNase H1-ΔHB2 to show weaker binding than RNase H1-ΔHB1, which maintains the linker domain (Figure 17A). Although *in vivo* removal of the linker domain (RNase H1-ΔHB1/LD) does not affect the binding and activity of RNase H1 containing only HB2 domain (Figure 20B,C), further *in vitro* experiments are required to understand the effect of the linker domain in RNase H1 binding to substrates and possibly its impact on the conformation of the recombinant proteins. Furthermore, purification of HB1 and HB2 domains would allow to confirm results obtained with RNase-ΔHB1 and RNase H1- ΔHB2.

Specificity towards RNA:DNA hybrids involves both binding to a specific RNA:DNA heteroduplex and not binding to non-hybrid substrates. Besides strong affinity towards hybrids, full length RNase H1 shows strong specificity to all hybrid substrates compared to ssDNA, dsDNA, ssRNA and dsRNA. Indeed, it has a 17, 4, 97 or 189-fold preference to RNA:DNA hybrids compared to the same ssDNA, dsDNA, ssRNA or dsRNA sequence, respectively

(Figure 25). Loss of HB1 domain leads to higher RNase H1 preference to RNA:DNA hybrids over ssDNA and dsDNA (19 and 37-fold, respectively, Figure 25). This may be due to the reduction of RNase H1-ΔHB1 affinity to DNA substrates, suggesting that HB1 domain recognizes ssDNA and dsDNA. On the other hand, the lack of HB1 domain decreases the preference of RNA:DNA hybrids over ssRNA and dsRNA to 15 and 100-fold, respectively (Figure 25). Although deletion of HB1 domain causes lower specificity towards RNA:DNA hybrids compared to RNA substrates, it leads to reduced affinity towards ssRNA and dsRNA. This shows that HB1 domain promotes a weak interaction with RNA and it does not decrease RNase H1 binding to ssRNA and dsRNA (Figure 25). Thus, these results indicate that HB1 domain modulates RNase H1 specificity mainly by increasing affinity towards RNA:DNA hybrids and not by decreasing affinity to RNA. These observations are also detected with R-loop substrates (Figure 25). Due to RNase H1-ΔHB2 lack of affinity to RNA and DNA substrates, we cannot postulate to what extend HB2 domain impacts RNase H1 specificity to RNA:DNA hybrids. Since RNase H1-ΔHB2 shows weaker binding to all substrates compared to full length RNase H1 and RNase H1-ΔHB1, it is possible that HB2 domain increases RNase H1 specificity by promoting affinity towards hybrids instead of decreasing affinity to DNA and RNA substrates. Overall, our data offers another example that combination of multiple binding domains lead to increased specificity[259]. Strikingly, reduction of specific interaction reduces RNase H activity[219,221]. Although RNase H1-ΔHB1 decreases hybrids specificity compared to RNA substrates *in vitro*, its overexpression *in vivo* reduces accumulated R-loops (Figure 20C). One possible explanation is that high protein levels of this mutant are capable of compensating for the loss of specificity. Hence, it would require to use an endogenous truncation mutant to address this issue.

In vitro RNase H1 without both HB domains does not bind to hybrids (Figure 23). This confirms that deficiency in R-loop removal by RNase H1-ΔHB1/2 *in vivo* (Figure 20C) is due to its lack of binding to RNA:DNA hybrids through HB domains, as described previously *in vitro*[72,221]. In human cells, RNase H1 contains only one HB domain which can anchor the enzyme, or facilitate its re-association, to the substrate to perform consecutive rounds of cleavage[72]. Since *S. cerevisiae* RNase H1 has two HB domains that confer strong affinity and specificity to RNA:DNA hybrids, likely RNase H1 binds to R-loops first through interaction of both HB domains with the RNA:DNA heteroduplex and, consecutively, the catalytic domain can bind and process the hybrid (Figure 32, left R-loop scheme). The binding of two HB domains causes deformation of the double helix of RNA:DNA hybrids[72]. It is possible that the catalytic domain requires deformation of the hybrid for its efficient processing.

In yeast, the linker domain between HB domains is longer than the second linker domain connection HB2 domain to the catalytic domain (Figure 2A, B). Likely, the first linker domain

confers a lot of flexibility and mobility to HB domains, allowing them to cover a wide range of the hybrid or bind multiple elements of the R-loop. Considering the different properties of both HB domains, it is tempting to speculate that RNase H1 recruitment to R-loops starts with the interaction of HB1 domain with RPA and/or the RNA:DNA hybrid and HB2 binding to the nascent RNA (Figure 32, right R-loop). Other possibility is that upon degradation of the RNA moiety by RNase H1, HB2 domain would bind to nascent RNA or the displaced ssDNA in order to keep a strong binding to the hybrid while it gets shorter due to RNase H activity.
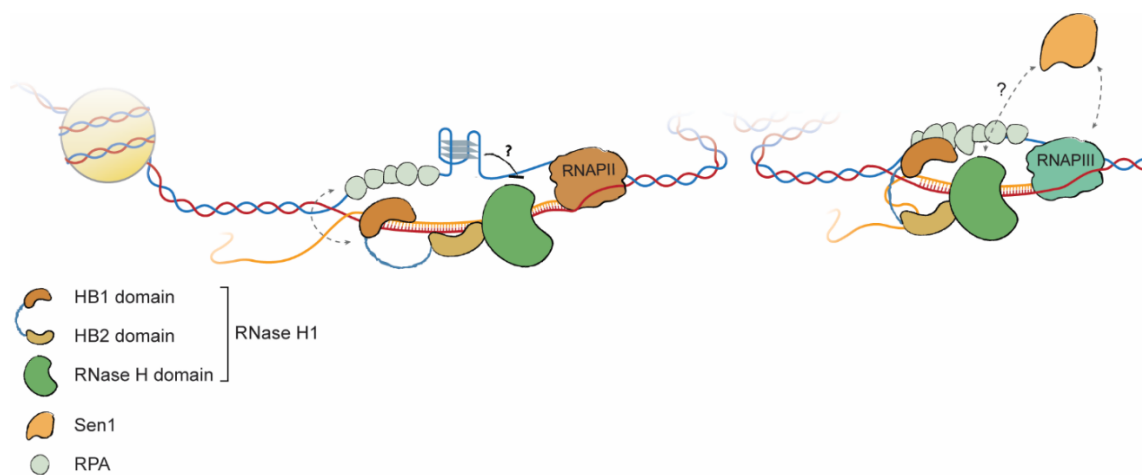
Figure 32 – RNase H1 binding to RNA:DNA hybrids model. RNase H1 requires both HB domains to have a strong and efficient binding to R-loops. HB domains have strong affinity to RNA:DNA hybrids and possibly both HB domains first interact with the hybrid to bring the catalytic domain to the substrate and start catalysis (left R-loop illustration). Alternatively, it is possible that HB2 domain interacts with other components of the R-loop structure, such as the nascent RNA, and the HB1 domain interacts with RPA and/or RNA:DNA hybrid (right R-loop illustration). RPA interacts with HB1 domain, but does not promote RNase H1 binding to R-loops. Other factors present in the R-loop are not currently known to affect RNase H1 recruitment, such G-quadruplexes (left) or Sen1 (right).

3.3.4 RPA does not promote RNase H1 binding to R-loops *in vitro*

RPA is a major single-stranded DNA-binding protein in eukaryotic cells. Here, we confirm that RPA has strong affinity for substrates that contain ssDNA, including R-loops (Figure 25). A recent study found that human RPA binds to ssRNA and can promote R-loop formation with homologous dsDNA[234]. Surprisingly RPA binding to ssRNA is 300-400-fold stronger than the homologous dsDNA and only 30-60-fold weaker than a ssDNA of identical sequence[234]. Indeed, a proteome study has identified RPA among RNA-interacting proteins[260]. We provide evidence that RPA strongly associates with ssRNA compared to other substrates without exposed ssDNA (Figure 25). RPA-ssRNA interaction is 220-fold weaker than RPA binding to ssDNA with identical sequence (Figure 25). Although we cannot compare binding to dsDNA,

RPA shows a 2-fold or 4-fold preference of ssRNA compared to RNA:DNA hybrid or dsRNA with the same sequence, respectively (Figure 25). These results suggest that yeast RPA does not have a high specificity to ssRNA compared to other substrates that do not contain ssDNA, unlike human RPA[234]. It is important to note that in that case we have used shorter oligomers than the previous study[234], which might affect the observed results observed concerning the affinity and specificity of RPA towards ssRNA.

RPA binds to R-loops *in vitro* (Figure 23A), but our *in vivo* data implies that RPA is not required for RNase H1 binding to R-loops and their removal (Figure 20B, C). Similarly, RPA does not promote RNase H1 binding to R-loops *in vitro* (Figure 26). In human cells, increasing amounts of RPA promote R-loop degradation by RNase H1[233]. Nevertheless, the same study does not show if RPA affects RNase H1 binding to R-loops *in vitro*. Since RNase H1 activity is associated with its binding capacity to its substrate[72,221], it would be expected that RPA influences RNase H1 recruitment to R-loops, as it was previously shown for RNA:DNA hybrids[233]. Surprisingly, even though RPA does not promote RNase H1 binding to R-loop, incubation of high amounts of RPA increases RNase H1 interaction with RNA:DNA hybrids *in vitro* (Figure 27). One possible explanation that might account for this difference between RNA:DNA hybrids and R-loops is the amount of RPA used for each condition. Since RPA weakly interacts with RNA:DNA hybrids, we used a higher RPA amount with RNA:DNA hybrids than with R-loops. Thus, RPA might affect RNase H1 binding to hybrids only when highly concentrated. In order to confirm these results, the same assay must be repeated with higher and lower amounts of RPA incubated with R-loops or RNA:DNA hybrids, respectively.

Understanding if RPA promotes RNase H1 binding to hybrids in a dosage-dependent manner, it would allow us to suggest that yeast RPA induces RNase H1 interaction with hybrids by direct interaction with RNase H1. However, we show that RNase H1-ΔHB1 binding to RNA:DNA hybrids is enhanced by RPA, while RNase H1-ΔHB2 is not (Figure 27). *In vivo*, RNase H1 interacts with RPA through its first HB domain (Figure 18). Thus, RNase H1 binding to RNA:DNA hybrids cannot be promoted due to a direct interaction with RPA through its HB1 domain. Since full length RNase H1 and RNase H1-ΔHB1 are the only recombinant proteins containing the linker domain between HB1 and HB2 domains, it is probable that the higher RNase H1 affinity to RNA:DNA hybrids is caused by direct interaction of RPA with the linker domain. Indeed, RNase H1-2xHB2, which contains the linker domain, but not the HB1 domain (Figure 17A), does interact with RPA *in vivo* (Figure 18). Moreover, it has been speculated that the linker domain might interact with other proteins or complexes, due to its differences in length and composition between organisms[191].

Although RNase H1 interacts with RPA *in vivo* (Figure 18), we could not reproduce the same result using the purified proteins. Since RNase H1-RPA interaction is DNA-dependent,

we tested multiple size ssDNA, but it was not possible to observe the interaction. Potentially, the purified proteins present conformational changes to *in vivo* proteins or other factors may affect RPA binding to RNase H1 besides ssDNA, such post-translational modifications.

RPA binding to ssDNA is regulated by its phosphorylation[254]. S178 phosphorylation reduces affinity of RPA to ssDNA by increasing the affinity between adjacent RPA. Consequently, the reduced coverage of ssDNA by RPA exposes enough ssDNA to allow the action of recombinases and nucleases[254]. This model suggests that the cooperative binding behavior of RPA could be 'switched on' by its phosphorylation and this might be important for its efficient assembly and its subsequent recruitment and exchange with other ssDNA-binding factors, such as Rad51[254]. It is possible that other proteins affecting RNase H1 activity might be recruited when RPA exhibits this cooperative behavior. Since it is also known that S178 phosphorylation leads to rearrangements of the domains of RPA, it is questionable whether a phosphorylation of RPA could modulate its binding to RNase H1. Testing this hypothesis would be critical in order to clarify if RPA-RNase H1 interaction can be reproduced *in vitro*.

3.4   Identification of RNase H1 regulators

Since RNase H1 overexpression promotes R-loop degradation, it has been suggested that a repressor must be titrated out upon overexpression and inhibited in order to unleash RNase H1 activity[205]. In our study, a gene overexpression library in *rnh201* mutants was used to identify proteins that may inhibit RNase H1 and by looking for candidates that cause growth defect in conditions with high R-loop loads, similar to *rnh1 rnh201* double mutants (Figure 28A). We identified 18 potential repressors of RNase H1 and cause *rnh201* single mutants to grow less in MMS when overexpressed (Figure 28E, Table 4). However, all the identified proteins show minor growth differences compared to wild-type cells that overexpress the same candidates. Indeed, upon visual inspection of the plates, these differences were not noticeable. Moreover, most of these proteins have mitochondrial and metabolic functions. Interestingly, Rtt106 (YNL206C) is identified as potential repressor of RNase H1. Rtt106 is a histone chaperone involved in regulation of chromatin structure of both transcribed and silenced chromosomal regions[261–263]. Thus, it would be very interesting to explore whether this candidate constitutes an important connection between RNase H1 activity and chromatin state. However, a recent work showed that Rtt106 is an essential protein for drug resistance in yeast[264]. Therefore, growth defects in cells overexpressing Rtt106 might be associated with a defective response to MMS, which is used in our screen. Nevertheless, the interplay between chromatin remodelers and RNase H1 activity would be a valuable subject to further investigate R-loop regulation.

To identify genes that might promote RNase H1 activity, we crossed the yeast KO library with *rnh201* single mutants and selected genes that lead to reduced growth in conditions with accumulated R-loops. We found 17 and 29 potential genes that promote RNase H1 activity when R-loop levels increase due to MMS addition or Sen1 absence, respectively (Figure 28C, G; Table 2, Table 6). As expected RNase H1 deletion lead to growth defect of *rnh201* mutants in MMS or Sen1 absence, showing that these conditions lead to R-loop accumulation. Once again, most of the identified genes are mitochondrial genes. Two possibilities may account for enrichment of metabolic-related genes. On one hand, RNase H1 is the only RNase H enzyme recruited to the mitochondria to regulate R-loop levels. Accumulation of R-loop in the mitochondria might affect mtDNA replication and cause metabolic stress[161]. On the other hand, it is probable that the conditions used to induce high R-loop levels, in particular MMS, might trigger stress response. However, removing Sen1 as an alternative way to induce R-loop accumulation does not seem to alter the detection of mitochondrial genes in our genetic screen, suggesting that such enrichment can be due to RNase H1 function specifically in the mitochondria.

Also, after identifying multiple candidates by spotting assay, we confirmed that most of the potential genes that might affect RNase H do not show any growth defect (Figure 29). However, Sit4 deletion is synthetic sick in the absence of RNase H1 and conditions with high R-loop loads are lethal for *sit4 rnh1* double mutants (Figure 29A). This suggests that Sit4 has an impact on RNase H2 activity resulting in a similar phenotype as *rnh1 rnh201* double mutants. On the other hand, the *sit4 rnh1* double mutant growth defect might indicate that Sit4 is required to respond to stress induced by RNase H1 depletion and not RNase H2. Sit4 is a protein phosphatase required for a proper cell cycle progression[265] and is important to maintain mitochondrial function[266]. Hence, Sit4 might be essential to deal with mitochondrial stress accumulated in *rnh1* mutants.

Different conditions and further analyses of the genetic screens could not identify a clear candidate that affect the activity of RNase H enzymes. The limitations of this screen might count on the complexity we added in order to identify candidates that have an impact on R-loop accumulation. An alternative genetic screen using *rnh1 rnh201* double mutants crossed with yeast KO or OE collections would allow us to identify genes important for R-loop accumulation response. However, such screen would not detect proteins that would affect RNase H1 regulation particularly.

3.5 Conclusion

In this study, RNase H1 overexpression was employed to understand its regulation and influence on R-loops. The data obtained confirm RNase H1 binds and removes accumulated R-loops in order to keep R-loop homeostasis. Furthermore, RNase H1 has a preference towards RNAPIII-transcribed genes, which might suggest a possible interaction with RNAPIII. However, why and how such preference occurs remains unclear and further experiments are required. RNase H1 removes R-loops outside of a replication-transcription context. Unlike other R-loop resolving proteins[250], the cell cycle-independent activity allows RNase H1 to be responsible to keep R-loop in check in all conditions.

Finally we confirm that RNase H1 interaction with RPA is DNA-dependent. The fact that we could not reconstitute this interaction *in vitro* with purified proteins and ssDNA might indicate that such interaction requires other conditions to happen, such as post-translational modifications. Unlike human cells, RNase H1 does not require interaction with RPA to bind and remove R-loops. Further investigation of RNase H1 domains shows that both HB domains are important for R-loop binding *in vivo* and likely both HB domains have different functions in RNase H1 binding to substrates and activity. Although the genetic screen did not identify possible regulators of RNase H enzymes, additional experiments are required to confirm that RNase H enzymes are not regulated by other proteins regulating their activity. Conclusively, the results of this study provide deeper insights of how RNase H1 is regulated and may provide a basis to further understand mechanisms of R-loop regulation.

4 Materials and Methods

4.1 Materials

4.1.1 Yeast strains

All yeast strains used in this work are described in Table 1 and were derived from the commonly used *Saccharomyces cerevisiae* strain BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*)[267]. As indicated, many of the strains were taken or derived from the "*Saccharomyces* Genome Deletion Project*"[268–270], supplied via the purchased collections from Dharmacon/Horizon Discovery. Correct gene deletion, or tagging, of strains obtained from the collection was verified by PCR and if possible by Western blot.

Strain	Genotype	Source
YBL7	*MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 lys2Δ0*	Euroscarf
yFB307	*MATα can1Δ::STE2pr-his5 lyp1Δ::STE3prLEU2 ura3Δ0 leu2Δ0 his3Δ1 met15Δ0 rnh201::NAT*	This study
yFB312	*MATα can1Δ::STE2pr-his5 lyp1Δ::STE3prLEU2 ura3::NAT leu2Δ0 his3Δ1 met15Δ0*	This study
yFB324	*MATα can1Δ::STE2pr-his5 lyp1Δ::STE3prLEU2 ura3Δ0 leu2Δ0 his3Δ1 met15Δ0 rnh1::NAT*	This study
yFB340	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG*	This study
yFB1029	*MATα his3Δ1; leu2Δ0; ura3Δ0; asf1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1030	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN asf1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1031	*MATa his3Δ1; leu2Δ0; ura3Δ0; asf1::KAN rnh201::HYG sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1032	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN rnh201::HYG asf1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1033	*MATa his3Δ1; leu2Δ0; ura3Δ0; sit4::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1034	*MATα his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN sit4::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1035	*MATa his3Δ1; leu2Δ0; ura3Δ0; sit4::KAN rnh201::HYG sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study

yFB1036	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN rnh201::HYG sit4::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1037	*MATα his3Δ1; leu2Δ0; ura3Δ0; tad1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1038	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN tad1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1039	*MATα his3Δ1; leu2Δ0; ura3Δ0; tad1::KAN rnh201::HYG sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1040	*MATα his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN rnh201::HYG tad1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1041	*MATa his3Δ1; leu2Δ0; ura3Δ0; gud1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1042	*MATα his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN gud1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1043	*MATα his3Δ1; leu2Δ0; ura3Δ0; gud1::KAN rnh201::HYG sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1044	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN rnh201::HYG gud1::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1061	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN rnh201::HYG hsp82::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1062	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh1::KAN hsp82::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1063	*MATa his3Δ1; leu2Δ0; ura3Δ0; rnh201::HYG hsp82::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1064	*MATa his3Δ1; leu2Δ0; ura3Δ0; hsp82::KAN sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2)*	This study
yFB1071	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 RNH1(D913N)-6xHA::KAN*	This study
yFB1091	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 RNH1-6xHA::KAN*	This study
yFB1121	*MATa his3Δ1; leu2Δ0; ura3Δ0; RNH1(D193N)-6xHA::KAN rnh201::HYG*	This study
yFB1123	*MATα his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1(D193N)-6xHA::KAN*	This study
yFB1125	*MATα his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1(D193N)-6xHA::KAN rnh201::HYG*	This study
yFB1127	*MATa his3Δ1; leu2Δ0; ura3Δ0;RNH1-6xHA::KAN rnh201::HYG*	This study
yFB1129	*MATa his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1-6xHA::KAN*	This study

yFB1131	*MATa his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1-6xHA::KAN rnh201::HYG*	This study
yFB1137	*MATa his3Δ1; leu2Δ0; ura3Δ0; RNH1(D193N)-6xHA::KAN rnh201::HYG*	This study
yFB1139	*MATa his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1(D193N)-6xHA::KAN*	This study
yFB1141	*MATa his3Δ1; leu2Δ0; ura3Δ0; sen1::SEN1-AID*-9MYC-HIS3MX6 LEU2::pGPD-AFB2 (LEU2) RNH1(D193N)-6xHA::KAN rnh201::HYG*	This study
yFB1175	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL906*	This study
yFB1178	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL797*	This study
yFB1179	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL837*	This study
yFB1241	*MATa his3Δ1;leu2Δ0; ura3Δ0; rnh201::HYG NAT-S-RNH1-TAP-HIS*	This study
yFB1242	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-S-RNH1-TAP-HIS*	This study
yFB1243	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU rnh201::HYG NAT-S-RNH1-TAP-HIS*	This study
yFB1244	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; rnh201::HYG NAT-G2-RNH1-TAP-HIS*	This study
yFB1245	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-G2-RNH1-TAP-HIS*	This study
yFB1246	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU rnh201::HYG NAT-G2-RNH1-TAP-HIS*	This study
yFB1247	*MATα his3Δ1;leu2Δ0 ; ura3Δ0; rnh201::HYG NAT-G1-RNH1-TAP-HIS*	This study
yFB1248	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-G1-RNH1-TAP-HIS*	This study
yFB1249	*MATα his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU rnh201::HYG NAT-G1-RNH1-TAP-HIS*	This study
yFB1258	*MATα his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-S-RNH1-TAP-HIS*	This study
yFB1259	*MATa his3Δ1;leu2Δ0 ; ura3Δ0; SEN1-AID::HIS AFB2::LEU rnh201::HYG NAT-S-RNH1-TAP-HIS*	This study
yFB1260	*MATa his3Δ1;leu2Δ0; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-G2-RNH1-TAP-HIS*	This study
yFB1261	*MATa his3Δ1;leu2Δ0; ura3Δ0; SEN1-AID::HIS AFB2::LEU rnh201::HYG NAT-G2-RNH1-TAP-HIS*	This study
yFB1262	*MATα his3Δ1;leu2Δ0; ura3Δ0; SEN1-AID::HIS AFB2::LEU NAT-G1-RNH1-TAP-HIS*	This study
yFB1356	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL797*	This study

yFB1357	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL799*	This study
yFB1359	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL800*	This study
yFB1361	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL802*	This study
yFB1364	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL804*	This study
yFB1366	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL806*	This study
yFB1367	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL808*	This study
yFB1416	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL833*	This study
yFB1418	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL835*	This study
yFB1424	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 /pBL906*	This study
yFB1426	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 /pBL797*	This study
yFB1428	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 /pBL837*	This study
yFB1494	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL872*	This study
yFB1496	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL873*	This study
yFB1498	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL874*	This study
yFB1500	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL875*	This study
yFB1514	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 AFB2::LEU Sen1-AID*::HIS G1-delta52-RNH1::HIS/NAT*	This study
yFB1516	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh201::HYG AFB2::LEU Sen1-AID*::HIS G1-delta52-RNH1::HIS/NAT*	This study
yFB1518	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 AFB2::LEU Sen1-AID*::HIS S-delta52-RNH1::HIS/NAT*	This study
yFB1519	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh201::HYG AFB2::LEU Sen1-AID*::HIS S-delta52-RNH1::HIS/NAT*	This study
yFB1521	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 AFB2::LEU Sen1-AID*::HIS G2-delta52-RNH1::HIS/NAT*	This study
yFB1523	*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh201::HYG AFB2::LEU Sen1-AID*::HIS G2-delta52-RNH1::HIS/NAT*	This study
yFB1543	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL907*	This study
yFB1545	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL908*	This study
yFB1547	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL909*	This study
yFB1549	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL910*	This study
yFB1551	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL911*	This study

yFB1553	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL912*	This study
yFB1555	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL913*	This study
yFB1557	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL914*	This study
yFB1559	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL923*	This study
yFB1561	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL924*	This study
yFB1563	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL925*	This study
yFB1565	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL926*	This study
yFB1567	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL927*	This study
yFB1569	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL928*	This study
yFB1571	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL929*	This study
yFB1573	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL930*	This study
yFB1575	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL931*	This study
yFB1577	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL932*	This study
yFB1579	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL933*	This study
yFB1581	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL934*	This study
yFB1583	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL935*	This study
yFB1585	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL936*	This study
yFB1587	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL937*	This study
yFB1589	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL938*	This study
yFB1591	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL939*	This study
yFB1593	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL940*	This study
yFB1595	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL941*	This study
yFB1597	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL942*	This study
yFB1599	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL943*	This study
yFB1601	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL944*	This study
yFB1709	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL959*	This study
yFB1711	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL960*	This study
yFB1713	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL961*	This study
yFB1715	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL962*	This study

yFB1717	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL963*	This study
yFB1719	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL964*	This study
yFB1721	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL965*	This study
yFB1723	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL966*	This study
yFB1725	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL967*	This study
yFB1727	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL968*	This study
yFB1729	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL969*	This study
yFB1731	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL970*	This study
yFB1733	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL971*	This study
yFB1735	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL972*	This study
yFB1737	*MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rnh1::KAN rnh201::HYG /pBL973*	This study

4.1.2 Plasmids

Plasmid	Description	Source
pBL58	*pENTR1A*	Alberti, Lindquist, 2007[271]
pBL797	*pRS416-pGAL1-RNH1-3xHA*	This study
pBL799	*pRS416-pGAL1-RNH1-ΔHB1-3xHA*	This study
pBL800	*pRS416- pGAL1-RNH1-ΔHB1/2-3xHA*	This study
pBL802	*pRS416- pGAL1-RNH1-ΔHB2-3xHA*	This study
pBL804	*pRS416- pGAL1-RNH1(D193N)-ΔHB1-3xHA*	This study
pBL806	*pRS416- pGAL1-RNH1(D193N)-ΔHB1/2-3xHA*	This study
pBL808	*pRS416- pGAL1-RNH1(D193N)-ΔHB2-3xHA*	This study
pBL811	*pGEX-GST-His6-3C-RNH1(D193N)*	This study
pBL813	*pGEX-GST-His6-3C-RNH1(D193N)-ΔHB1*	This study
pBL815	*pGEX-GST-His6-3C-RNH1(D193N)-ΔHB1/2*	This study
pBL817	*pGEX-GST-His6-3C-RNH1(D193N)-ΔHB2*	This study
pBL829	*pDEST-MBP-His6-RPA*	This study
pBL833	*pRS416- pGAL1-RNH1-2xHB2-3xHA*	This study

pBL835	*pRS416- pGAL1-RNH1(D193N)-2xHB2-3xHA*	This study
pBL837	*pRS416- pGAL1-RNH1(D193N)-3xHA*	This study
pBL872	*pRS416- pGAL1-RNH1-2HB1-3xHA*	This study
pBL873	*pRS416- pGAL1-RNH1(D193N)-2HB1-3xHA*	This study
pBL874	*pRS416- pGAL1-RNH1 -ΔHB1/LD-3xHA*	This study
pBL875	*pRS416- pGAL1-RNH1 (D193N)-ΔHB1/LD-3xHA*	This study
pBL906	*pRS416- pGAL1-EV-3xHA*	This study
pBL907	*pRS416-GPD-EV-3xHA*	This study
pBL908	*pRS416-GPD-RNH1-3xHA*	This study
pBL909	*pRS416-GPD-RNH1-ΔHB1-3xHA*	This study
pBL910	*pRS416-GPD-RNH1-ΔHB1/2-3xHA*	This study
pBL911	*pRS416-GPD-RNH1-ΔHB2-3xHA*	This study
pBL912	*pRS416-GPD-RNH1 -ΔHB1/LD-3xHA*	This study
pBL913	*pRS416-GPD-RNH1-2HB1-3xHA*	This study
pBL914	*pRS416-GPD-RNH1-2xHB2-3xHA*	This study
pBL923	*pRS416-GPD-RNH1(D193N)-3xHA*	This study
pBL924	*pRS416-GPD-RNH1(D193N)-ΔHB1-3xHA*	This study
pBL925	*pRS416-GPD-RNH1(D193N)-ΔHB1/2-3xHA*	This study
pBL926	*pRS416-GPD-RNH1(D193N)-ΔHB2-3xHA*	This study
pBL927	*pRS416-GPD-RNH1(D193N)-2xHB2-3xHA*	This study
pBL928	*pRS416-GPD-RNH1(D193N)-2HB1-3xHA*	This study
pBL929	*pRS416-GPD-RNH1 (D193N)-ΔHB1/LD-3xHA*	This study
pBL930	*pRS426-GPD-EV-3xHA*	This study
pBL931	*pRS426-GPD-RNH1-3xHA*	This study
pBL932	*pRS426-GPD-RNH1-ΔHB1-3xHA*	This study
pBL933	*pRS426-GPD-RNH1-ΔHB1/2-3xHA*	This study
pBL934	*pRS426-GPD-RNH1-ΔHB2-3xHA*	This study
pBL935	*pRS426-GPD-RNH1-2xHB2-3xHA*	This study
pBL936	*pRS426-GPD-RNH1-2HB1-3xHA*	This study

pBL937	*pRS426-GPD-RNH1 -ΔHB1/LD-3xHA*	This study
pBL938	*pRS426-GPD-RNH1(D193N)-3xHA*	This study
pBL939	*pRS426-GPD-RNH1(D193N)-ΔHB1-3xHA*	This study
pBL940	*pRS426-GPD-RNH1(D193N)-ΔHB1/2-3xHA*	This study
pBL941	*pRS426-GPD-RNH1(D193N)-ΔHB2-3xHA*	This study
pBL942	*pRS426-GPD-RNH1(D193N)-2xHB2-3xHA*	This study
pBL943	*pRS426-GPD-RNH1(D193N)-2HB1-3xHA*	This study
pBL944	*pRS426-GPD-RNH1 (D193N)-ΔHB1/LD-3xHA*	This study
pBL959	*pRS426- pGAL1-EV-3xHA*	This study
pBL960	*pRS426- pGAL1-RNH1-3xHA*	This study
pBL961	*pRS426- pGAL1-RNH1-ΔHB1-3xHA*	This study
pBL962	*pRS426- pGAL1-RNH1-ΔHB1/2-3xHA*	This study
pBL963	*pRS426- pGAL1-RNH1-ΔHB2-3xHA*	This study
pBL964	*pRS426- pGAL1-RNH1-2xHB2-3xHA*	This study
pBL965	*pRS426- pGAL1-RNH1-2HB1-3xHA*	This study
pBL966	*pRS426- pGAL1-RNH1 -ΔHB1/LD-3xHA*	This study
pBL967	*pRS426- pGAL1-RNH1(D193N)-3xHA*	This study
pBL968	*pRS426- pGAL1-RNH1(D193N)-ΔHB1-3xHA*	This study
pBL969	*pRS426- pGAL1-RNH1(D193N)-ΔHB1/2-3xHA*	This study
pBL970	*pRS426- pGAL1-RNH1(D193N)-ΔHB2-3xHA*	This study
pBL971	*pRS426- pGAL1-RNH1(D193N)-2xHB2-3xHA*	This study
pBL972	*pRS426- pGAL1-RNH1(D193N)-2HB1-3xHA*	This study
pBL973	*pRS426- pGAL1-RNH1 (D193N)-ΔHB1/LD-3xHA*	This study

4.1.3 Oligonucleotides

Oligo	Use	Sequence (5'-3')
oFB256	Oligonucleotide for *in vitro* assays. RNA oligo (with FAM)[233]	GCAGCUGGCACGACAGGUAUGAAUC-36-FAM
oFB258	Oligonucleotide for *in vitro* assays. DNA oligo 1[233]	GCCAGGGACGAGGTGAACCTGCAGGTGGGCGGCTACTACTTAGATGTCATCCGAGGCTTATTGGTAGAATTCGGCAGCGTCATGCGACGGC
oFB259	Oligonucleotide for *in vitro* assays. DNA oligo 2[233]	GCCGTCGCATGACGCTGCCGAATTCTACCACGCGATTCATACCTGTCGTGCCAGCTGCTTTGCCCACCTGCAGGTTCACCTCGTCCCTGGC
oFB260	Oligonucleotide for *in vitro* assays. DNA oligo 3[233]	GCAGCTGGCACGACAGGTATGAATC-36-FAM
oFB261	Oligonucleotide for *in vitro* assays. DNA oligo 4[233]	GATTCATACCTGTCGTGCCAGCTGC
oFB262	Oligonucleotide for *in vitro* assays. DNA oligo 5[233]	GCAGTAGCATGACGCTGCTGAATTCTACCACGCTATGCTCTCGTCTAGGTTCACTCCGTCCCTGCGATTCATACCTGTCGTGCCAGCTGC
oFB270	Oligonucleotide for *in vitro* assays. RNA oligo (with FAM)	CUACCACGCUAUGCUGCAGCUGGCACGACAGGUAUGAAUC-36-FAM
oFB171	qPCR - SUF2	TATGATTCTCGCTTAGGGTGCGGGAGG
oFB172	qPCR - SUF2	CATTAACATTGGTCTTCTCCAGCTTACTC
oFB181	qPCR - RPL15a	ACC GCT GAA GAA AGA GTT GG
oFB182	qPCR - RPL15a	TGT TGA GGG TCG ACC AAG AT
oFB197	qPCR - SUF11	TCTTAACAACAAGTTAACAAGGGCG
oFB198	qPCR - SUF11	CTGTTACCCGACCAATAGGAAATAA
oSM16	qPCR - 5SrDNA	GGTTGCGGCCATATCTACCA
oSM17	qPCR - 5SrDNA	ACCTGAGTTTCGCGTATGGT
oAM47	qPCR - 18SrDNA	TCC AAT TGT TCC TCG TTA AG
oAM48	qPCR - 18SrDNA	ATT CAG GGA GGT AGT GAC AA
oBL292	qPCR - Actin	CCCAGGTATTGCCGAAAGAATGC
oBL293	qPCR - Actin	TTTGTTGGAAGGTAGTCAAAGAAGCC

4.1.4 Liquid media

Medium	Composition
Luria broth (LB) medium +/- carbenicillin	10 g/l NaCl 10 g/l bacto tryptone 5 g/l bacto yeast extract +/- 100 µg/mL carbenicillin
Sporulation (SPO) medium	5 mg/l zinc acetate 10 g/l potassium acetate
Synthetic complete (SC) medium without amino acid	1.92 g/l yeast synthetic dropout medium without amino acid 6.7 g/l yeast nitrogen base without amino acids 20 g/l glucose/ galactose/ raffinose
YPD medium	20 g/l peptone 10 g/l bacto yeast extract 20 g/l glucose
YPGal/Raf (2 % galactose/2 % raffinose)	20 g/l peptone 10 g/l bacto yeast extract 20 g/l raffinose x 5 H_2O 20 g/l galactose

4.1.5 Agar plates

Plate	Composition
LB plates	10 g/l NaCl 10 g/l bacto tryptone 5 g/l yeast extract 15 g/l agar
Presporulation plates	30 g/l standard nutrient broth 10 g/l yeast extract 20 g/l agar 50 g/l glucose
SC plates (SGA)	2 g/l yeast synthetic dropout medium without amino acid 1.7 g/l yeast nitrogen base without amino acids and ammonium sulfate 1 g/l monosodium glutamic acid

	24 g/l bacto agar
	20 g/l glucose
SC plates without amino acid	1.92 g/l yeast synthetic dropout medium w/o amino acid
	6.7 g/l yeast nitrogen base without amino acids
	24 g/l agar
	20 g/l glucose
Sporulation plates (SGA)	10 g/l yeast extract
	10 g/l potassium acetate
	24 g/l bacto agar
YPD plates	65 g/l YPD agar
YPD plates (SGA)	10 g/l yeast extract
	20 g/l peptone
	24 g/l bacto-agar
	20 g/l glucose

4.1.6 Media supplements

Antibiotic/non-proteinogenic amino acid	Concentration
Canavanine	50 µg/mL
Carbenicillin	100 µg/mL
G418 disulfate solution (Kanamycin)	250 µg/mL (200 µg/mL in SGA plates)
Hygromycin B	300 µg/mL
Nourseothricin-dihydrogen sulfate (ClonNaT)	100 µg/mL
Thialysine	50 µg/mL

4.1.7 Buffers and solutions

Buffer/Solution	Composition
10x alkaline running buffer	0.5 mM NaOH
	10 mM EDTA
10x blotting buffer	25 mM Tris
	192 mM glycine

10x PBS	1.37 M NaCl 27 mM KCl 100 mM Na2HPO4 18 mM KH2PO4 adjusted to pH 7.4 with HCl, autoclaved
1 x PBST	1 x PBS 0.1 % Tween-20
10x SDS running buffer	25 mM Tris glycine 0.1 % SDS adjusted to pH 8.3, autoclaved
10x TBE	0.89 M Tris base 0.89 M boric acid 0.02 M EDTA pH 8.0, autoclaved
10x TE	M Tris-HCl pH 7.5 10 mM EDTA pH 8.0
20x SSC	0.3M Sodium citrate tribasic dihydrate 3M NaCl pH adjusted to 7.0 with HCl, autoclaved
6x alkaline loading buffer	300 mM KOH 6 mM EDTA 18% (w/v) Ficoll 0.15% (w/v) bromocresol green 0.25% (w/v) xylene cyanol
6x DNA orange loading dye	15 % Ficoll 10 mM EDTA pH 8.0 Orange G
6x neutral loading buffer	30% (v/v) glycerol in TE 0.25% (w/v) bromphenol blue 0.25% (w/v) xylene cyanol
AE-Buffer	50mM NaAc (pH 5.3) 10mM EDTA
Blocking buffer	5 % (w/v) skim milk powder in 1 x PBST

Buffer III	250mM LiCl 10mM Tris-HCl (pH 8.0) 1% NP-40 1mM EDTA (pH 8.0) 24mM Sodium deoxycholate
Dialysis buffer	33mM Tris pH7.5 800mM NaCl 5% Glycerol 1mM DTT
EDTA pH 8.0	0.5 M disodium EDTA x 2H2O adjusted to pH 8.0 with NaOH
Elution Buffer	1% SDS 50mM Tris-HCl (pH 7.5) 10mM EDTA (pH 8.0)
Extraction buffer	50 mM Tris-HCl pH 7.5 100 mM KCl 2.5 mM MgCl2
FA Lysis Buffer -SOD	140mM NaCl 50mM HEPES (pH 7.5) 1% Triton X-100 1mM EDTA (pH 8.0)
FA Lysis Buffer + SOD	140mM NaCl 50mM HEPES (pH 7.5) 1% Triton X-100 1mM EDTA (pH 8.0) 2.4mM Sodium deoxycholate
FA Lysis Buffer 500	0.5M NaCl 50mM HEPES (pH 7.5) 1% Triton X-100 1mM EDTA (pH 8.0) 2.4mM Sodium deoxycholate
FP Buffer	200mM NaCl 20mM HEPES (pH 7.5) 5% Glycerol 0.025% Triton X-100 0.5mM EDTA (pH 8.0)

Gel filtration buffer	30mM Na-HEPEs pH7.4 300mM NaCl 10% Glycerol
GST wash buffer	50mM Tris pH7.5 800mM NaCl 5% Glycerol
GST elution buffer	GST wash buffer 15mM Glutathion, pH7.5
Hybridization buffer	90mM Tris pH7.5 10mM $MgCl_2$ 50mM NaCl
IP buffer +/- NP40	50 mM Tris pH 7.5 150 mM NaCl 5 mM $MgCl_2$ +/- 0.2%(v/v) NP40
Laemmelli Buffer (2x)	125mM Tris pH6.8 4% SDS 20% Glycerol 0.004% Bromophenol Blue 10% β-mercaptoethanol
LiAc mix	0.1 M lithium acetate 1x TE
Lysis buffer	50 mM HEPES pH 7.5 140 mM NaCl, 1 mM EDTA pH 8.0 1 % Triton X-100
Neutralization solution	700 mM Tris-HCl pH 8 1.5 M NaCl
PEG mix	40 % (w/v) PEG 400 dissolved in LiAc mix, sterile filtered
Solution 1	1.09M β-mercaptoethanol 1.85M NaOH
Solution 2	50% Trichloroacetic acid
Solution 3	100% Acetone

	1 M sorbitol
Spheroblasting buffer	50 mM potassium phosphate buffer pH 7.4
	1 mM 1M DTT
	62.5 mM Tris pH 6.8
Stripping buffer	2 % SDS
	0.7 % (v/v) ß-mercaptoethanol
Sucrose solution	30 % (w/v) sucrose in H2O
	120 mM Tris-HCl pH 6.8
	5 % glycerol
	8 M urea
Urea buffer	143 mM β-mercaptoethanol
	8 % SDS
	bromophenol blue

4.1.8 Antibodies

Antibody	Dilution	Source	Identifier
Goat monoclonal anti-mouse (HRP conjugate)	1:3000	BioRad	Cat# 170-5047
Goat monoclonal anti-rabbit (HRP conjugate)	1: 3000	BioRad	Cat# 170-5046
Mouse monoclonal anti-actin	1:2000	Millipore	Cat# MAB1501R
Mouse monoclonal anti-GFP	1:1000	Roche	Cat# 11814460001
Mouse monoclonal anti-HA	1:2000	IMB CF	N/A
Mouse monoclonal anti-Myc (clone 9B11)	1:1000	Cell Signaling /NEB	Cat# 2276S
Mouse monoclonal anti-PGK1	1:20000	Invitrogen	Cat# 459250
Mouse monoclonal anti-Rad53 (clone EI7E1)	1:1000	Abcam	Cat# ab166859
Rabbit monoclonal peroxidase anti-peroxidase (PAP)	1:3000	Sigma Aldrich	Cat# P1291
Rabbit polyclonal anti-Clb2	1:1000	Santa Cruz	Cat# y-180
Rabbit polyclonal anti-H3	1:2000	Abcam	Cat# ab1791
Rabbit polyclonal anti-RPA	1:20000	Agrisera	Cat#AS07214
Rabbit polyclonal anti-sheep (HRP conjugate)	1:5000	Abcam	Cat# ab97130
Rabbit polyclonal anti-Sic 1	1:2000	Santa Cruz	Cat# sc-50441
Rat monoclonal anti-HA High-affinity (clone 3F10)		Roche	Cat#11867423001

4.1.9 Reagents, enzymes and commercially available kits

Reagent/Resource	Source	Identifier
1 kb DNA ladder	New England Biolabs	Cat# N3232L

100 bp DNA ladder	New England Biolabs	Cat# N3231L
2-Propanol	Roth	Cat# 9866
2-Propanol	Sigma Aldrich	Cat# 34959
4–15% Mini-PROTEAN™ TGX Stain-Free™ Protein Gels	Bio-Rad	Cat# 4568086
7.5% Mini-PROTEAN® TGX Stain-Free™ Protein Gels	Bio-Rad	Cat# 4568026
Acetone	Sigma Aldrich	Cat# 1000141000
Acrylamide/Bis-Acrylamide 30%	Sigma-Aldrich	Cat# A3574
Agar	Sigma Aldrich	Cat# 05040
Agarose	Sigma Aldrich	Cat# A9539
Alpha-factor (synthetic peptide)	Zymo	Cat# Y1001
Amilose resin	New England Biolabs	Cat# E8021S
Ammonium Sulfate	Sigma Aldrich	Cat# 31119
Ammoniumperoxodisulfat (APS)	VWR	Cat# SIAL009913
Bacto tryptone	BD Biosciences	Cat# 211705
Bacto yeast extract	BD Biosciences	Cat# 212750
Blue DNA loading dye (6X)	New England Biolabs	Cat# B7024S
Bradford solution	AppliChem	Cat# A6932,0500
Bromophenol Blue	Sigma Aldrich	Cat# B0126
cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail	Sigma Aldrich	Cat# 4693159001
Concanavalin A	Sigma-Aldrich	Cat# L7647
DAPI	Sigma Aldrich	Cat# D9542
DL-Dithiothreitol (DTT)	Sigma-Aldrich	Cat# 43816
DMSO	Sigma Aldrich	Cat# 34943
D-sorbitol	Sigma-Aldrich	Cat# S1876
Dynabeads™ Protein G	Thermo Fisher Scientific	Cat# 10607605
Ethanol Absolute 99.8+%	Thermo Fisher Scientific	Cat# 10437341
Formaldehyde Solution, 37 %	AppliChem	Cat# A0877
Formaldehyde Solution, 37%	Sigma Aldrich	Cat# F8775
Galactose	AppliChem	Cat# A3609
Glucose	AppliChem	Cat# A1422
Glycerol	Thermo Fisher Scientific	Cat# 17904
Guanidine-hydrochloride	Sigma-Aldrich	Cat# 50950
IgG Sepharose	Thermo Fisher Scientific	Cat# 11574955
Imidazole	Sigma-Aldrich	Cat# I5513
Indole-3-acetic acid sodium salt (auxin)	Sigma-Aldrich	Cat# I3750
L-Canavanine sulfate salt ≥99%	Sigma Aldrich	Cat# C9758
L-Canavanine sulfate salt ≥99% (TLC), powder	Sigma Aldrich	Cat# C9758
L-glutamic acid monosodium salt hydrate	Sigma Aldrich	Cat# G1626
Magnesium chloride 5xH2O	AppliChem	Cat# B10
Methyl methanesulfonate	Sigma-Aldrich	Cat# 129925
MG-132	Sigma-Aldrich	Cat# C2211
NEBuffer	New England Biolabs	

Ni-NTA magnetic beads	Thermo Fisher Scientific	Cat# 88831
Nonidet P 40 (NP-40)	Sigma-Aldrich	Cat# 98379
Peptone	AppliChem	Cat# A2210
PhosSTOP, Phosphatase Inhibitor Tablets	Sigma Aldrich	Cat# 4906845001
Phos-tag Acrylamide AAL-107	Wako Chemicals	Cat# 300-93523
Poly(ethylenglycol) 400 (PEG)	Sigma-Aldrich	Cat# 81240
Ponceau S	Sigma-Aldrich	Cat# P7170
Potassium acetate	Sigma Aldrich	Cat# 25059
Potassium Phosphate monobasic (KH_2PO_4)	Sigma-Aldrich	Cat# P0662
Prestained Protein Marker, Broad Range (11 - 190 kDa)	New England Biolabs	Cat# P7706
Protein A Sepharose	Thermo Fisher Scientific	Cat# 11359931
Raffinose	AppliChem	Cat# A6882
RedSafe Nucleic Acid Stain	HISS-Diagnostics	Cat# 21141
S-(2-Aminoethyl)-L-cysteine hydrochloride	Sigma Aldrich	Cat# A2636
SDS - 20 % solution	AppliChem	Cat# A0675
Skim milk powder	Sigma Aldrich	Cat# 70166
Sodium acetate	Sigma Aldrich	Cat# S2889
Sodium chloride	Sigma Aldrich	Cat# 31434
Sodium deoxycholate	Sigma Aldrich	Cat# D6750
Sodium hydroxide	Sigma Aldrich	Cat# 30620
Sodium phosphate dibasic (Na_2HPO_4 x 2 H_2O)	Sigma-Aldrich	Cat# 71643
Sodium thiosulfate ($Na_2S_2O_3$)	Sigma-Aldrich	Cat# 217247
Standard nutrient broth	Sigma Aldrich	Cat# S4681
SuperSignal West DURA Extended Duration Substrate	Thermo Fisher Scientific	Cat# 10220294
SuperSignal West PICO Substrate	Thermo Fisher Scientific	Cat# 15669364
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific	Cat# 10328162
SYTOX Green Nucleic Acid Stain	Thermo Fisher Scientific	Cat# S7020
Trichloroacetic acid	Sigma Aldrich	Cat# 27242
Triton X-100	Sigma Aldrich	Cat# X100
Tween-20	AppliChem	Cat# T1503
Urea	Sigma Aldrich	Cat# U5378
Yeast extract	Thermo Fisher Scientific	Cat# 16259781
Yeast nitrogen base without amino acids	Sigma Aldrich	Cat# Y0626
Yeast synthetic dropout medium without uracil	Sigma Aldrich	Cat# Y1501
Yeastmarker carrier DNA	Takara	Cat# 630440
YPD-agar	Sigma Aldrich	Cat# Y1500
Zinc acetate	Sigma Aldrich	Cat# 383317
β-mercaptoethanol	Sigma Aldrich	Cat# M6250

Enzyme	Source	Identifier
2x Gibson assembly enzyme mix	CF Protein production, IMB	N/A

2x Phusion HF PCR Mastermix	New England Biolabs	Cat# M0531L
2x Q5® High-Fidelity Master Mix	NEB	Cat# M0492L
2x Taq PCR Mastermix	NEB	Cat# M0270L
Benzonase Nuclease	Sigma Aldrich	Cat# E1014
DNase I (1500 Kunitz units)	Qiagen	Cat# 79254
Lyticase	Sigma Aldrich	Cat# L4025
LR Clonase	CF Protein production, IMB	N/A
Proteinase K	myneolab	Cat# 1151ML010
RNase A	Thermo Fisher Scientific	Cat# EN0531
RNase III	CF Protein production, IMB	N/A
RNase T1	Thermo Fisher Scientific	Cat# 10314820
S1 nuclease	Thermo Fisher Scientific	Cat# 10529310
Zymolyase T100	Zymo research	Cat# E1005

Kit	Source	Identifier
Gentra Puregene Yeast/Bact. Kit	Qiagen	Cat# 158567
MinElute PCR Purification Kit	Qiagen	Cat# 28004
QIAprep Spin Miniprep Kit	Qiagen	Cat# 27106X4
QIAquick Gel Extraction Kit	Qiagen	Cat# 28506
QIAquick PCR Purification Kit	Qiagen	Cat# 28106
RNeasy MinElute Cleanup Kit	Qiagen	Cat# 74204
Trans-Blot Turbo RTA Midi NitrocelluloseTransfer Kit	Bio-Rad	Cat# 1704271

4.1.10 Electronic devices and software

Electronic device	Source
BD FACSVerse	Becton Dickinson
BD LSRFortessa SORP	Becton Dickinson
BioRuptor Pico & Water Cooler Minichiller	Diagenode
ChemiDoc Touch Imaging System	BioRad
DeltaVision Elite system	GE Healthcare
Dissection Microscope MSM 400	Singer Instruments
FastPrep-25	MP Biomedicals
Leica DM1000 LED	Leica
NanoDrop 2000	Thermo Fisher Scientific
PhenoBooth+	Singer Instruments
PowerPac Basic	Bio-Rad
Qubit 2.0 Fluorometer	Thermo Fisher Scientific
Singer ROTOR HDA	Singer Instruments
Sonifier 450	Branson
Spark 20M multimode microplate reader	Tecan
Spectrophotometer Ultrospec 2100 pro	Biochrom
Thermal Cycler C1000 Touch	BioRad
Trans-Blot Turbo Transfer System	Bio-Rad
Typhoon FLA9500	GE Healthcare
Widefield microscope AF7000	Leica

Software	Source
Adobe Illustrator	Adobe
FileMaker Pro 10	FileMaker Inc
FlowJo V10	Becton Dickinson
ImageJ Fiji	https://fiji.sc/
ImageLab V5	BioRad
Mendeley Desktop	Elsevier
MS Office 2016	Microsoft
Prism 7.03	GraphPad
SoftWoRx	GE Healthcare

4.2 Methods

4.2.1 Yeast culture and strain generation

All strains were maintained in YPD or synthetic complete (SC) media supplemented with the appropriate amino acids at 30°C if not indicated otherwise. In general, overnight cultures contained 5 mL of the appropriate media inoculated with a single colony from a streak out or a tetrad from a dissection, and incubated for 8-16h at the respective temperature and 250 rpm. For exponential growth, overnight cultures were diluted in a new cultivating vessel with fresh media to an OD_{600}= 0.1-0.2 and incubated at the respective temperature and 200 rpm until reaching a cell density of OD_{600}= 0.8-1.

Strains harboring a plasmid with a gene under the *GAL1* promoter were grown in 2% raffinose-selective media and induced with a final concentration of 2% galactose for the indicated time period.

Gene tagging was carried out with standard PCR-based methods as previously described[272,273] verified by PCR and Western blot. Strains carrying multiple gene mutations, deletions or tags were generated by tetrad dissection of a diploid cross carrying the relevant modifications.

For the synthetic genetic array (SGA) query strain construction, complete *RNH1* or *RNH201* deletion was carried in the background strain (Y8205, Source C. Boone) using standard PCR-based methods containing an antibiotic marker for selection.

4.2.2 Bacterial transformation

The competent Escherichia coli DH5α cells were kept on ice during the transformation process. After adding 100 ng recombinant DNA to the thawed mix and 30 min incubation, cells were heat shocked for 1 min at 42°C and cooled down on ice for 1 min. The transformation mix, mixed with 300µL LB medium, was recovered for 30 min at 37°C and plated on selective plates. Plates were incubated overnight at 37°C.

4.2.3 Yeast transformation

Transformation of buddying yeast with alkali cations was performed as described[274]. Here, a final concentration of 35% (w/v) PEG 400 was used instead of 35% (w/v) PEG3350

and cells were heat shocked for 15 min instead of 40 min. Yeast transformed with plasmids or PCR fragments for integration were recovered for 30 min or five hours, respectively. Cells were plated onto selective media and incubated at 30°C until visible colonies were formed.

4.2.4 Synthetic Genetic Array (SGA) analysis

Overnight cultures of the query strains constructed for SGA analysis (4.2.1) were diluted for exponential growth in 25 mL cultures, respectively and transferred to fill a 96-well culture plate with 150 µL per well. Using the Singer RoToR HDA pinning robot loaded with sterile Singer RePads, each query strain and the entire yeast knockout collection were replicated from a 96-well culture plate to solid medium containing Singer PlusPlate in a 384-colony format array, respectively. Due to greater accessibility of nutrients at the plate border, to minimize spatial effects, a dummy strain was used instead of the query strain in the perimeter wells of the 96-well culture plates. The yeast KO and MOBY2 collections were re-arrayed beforehand to exclude these border wells. These border dummy colonies were excluded from the analysis. After 1 day of growth, each query strain 384-colony format array was replicated in quadruplicate onto a fresh solid medium containing Singer PlusPlate to generate a 1536-colony density array. After 2 days of growth, query strains were mated with yeast KO collection in 1536-well format on fresh YPD plates. After 1 day of growth, the mated strains were replicated to YPD + clonNAT+KAN Singer PlusPlates to select for diploids. After 2 days of growth, diploid-selected colonies were replicated to Singer PlusPlates-containing sporulation medium. After 8 days at 23°C, sporulated colonies were replicated to Singer PlusPlates containing SC – (Arg, Lys, His) + (canavanine, thialysine, KAN) for the first round of haploid selection. After 2 days of growth, colonies were replicated to Singer PlusPlates containing SC – (Arg, Lys, His) + (canavanine, thialysine, KAN, clonNAT) for the second round of haploid selection. After 2 days of growth, selected haploid double mutants were replicated to Singer PlusPlates containing SC – (Arg, Lys, His) + (KAN, clonNAT) plates for recovery. Images of selected haploid double mutants for analysis were taken after 24 h of recovery with the Singer PhenoBooth colony imager.

The medium for the Singer PlusPlates was mixed, autoclaved and poured each day before pinning. Singer RePads were reused by washing in water and sterilization in 70% ethanol prior to pinning.

4.2.5 Spotting/Serial Dilution assay

Yeast overnight cultures were diluted to 0.5 OD600 and spotted in ten-fold serial dilutions onto YPD plates or SC media plates supplemented with the appropriate amino acids. If indicated, the plates contained concentrations of indol acetic acid and/or, methane methyl sulfonate. Plates were incubated for up to 72h. Plate's images were taken with ChemiDoc™ Touch Imaging System and analyzed using Image Lab V5 software.

4.2.6 Cell cycle arrest and release

Overnight cultures of MATa cells were diluted for exponential growth and synchronized in G1 by addition of 4 µg/mL alpha-factor for 2 h. Synchronized cells were released into S phase by centrifugation at 3000g, washing two times with sterile H2O, and resuspension in the indicated medium and temperature. The release was monitored by flow cytometry at the respective time points.

4.2.7 DNA content flow cytometry

Cells were fixed with 70% ethanol and incubated at 4°C overnight. Samples were washed with 50 mM sodium citrate (pH 7.0) and incubated with 200 mg/mL RNase A at 37°C for 3 h. Afterwards, cells were incubated with 400 mg/mL proteinase K at 50°C for 1 h. Samples were sonicated with ultrasound for 2 cycles 30 sec each using Bioruptor Pico (Diagenode) and stained with 0.5 µM Sytox Green. DNA content was analyzed by flow cytometry with a FACSVerse- or a FACSFortessa cytometer, and data was analyzed with the FlowJo v10 software.

4.2.8 Protein extraction, SDS-PAGE and Western blot

Total yeast protein extracts were prepared by trichloroacetic acid (TCA) precipitation as described[275]. Proteins were resolved by SDS-PAGE using 7.5% or 4-15% Mini-PROTEAN® TGX Stain-Free™ precast polyacrylamide gels (Bio-Rad) in 1x SDS running buffer using Mini-PROTEAN® electrophoresis chambers (Bio-Rad). The gels were photoactivated for 45 sec with the ChemiDoc Touch Imaging System for total protein visualization. Proteins were transferred onto 0.45 µm pore size nitrocellulose membrane with the semi-dry Trans-Blot® Turbo™ Transfer System (Bio-Rad) and stained with Ponceau S solution to visualize the transfer. The membrane was blocked for 1h at RT with blocking buffer (5% skim milk in PBST)

and incubated with the relevant primary antibody in blocking buffer overnight at 4°C. Membranes were washed in four washes of PBST before staining with an HRP conjugated secondary antibody for 1 h at RT. Membranes were washed in three washes of PBST and one wash of PBS before signal development with Super Signal West Pico/Dura Chemiluminescent Substrate on a Bio-Rad ChemiDoc Touch Imaging System. Image Lab software was used for quantification. Antibodies are listed in 4.1.8.

If necessary, membranes were stripped of antibodies in two washes of stripping buffer fro 15 min followed by three washes of PBS. Stripped membranes were block again with 5% skim milk in PBST and reprobed with the relevant antibodies.

4.2.9 Co-Immunoprecipitation

Overnight cultures were diluted for exponential growth in fresh 100 mL cultures, grown until 0.8-1 OD600, and harvested by centrifugation. Prior to cell lysis, protease inhibitors were added to the IP buffer. Harvested cells were lysed in 200 µL IP buffer (- NP40) using FastPrep (30 sec at 6.5m/s followed by incubation for 1 min on ice before a second and third run) at 4°C using Lysing Matrix C tubes. Lysates were diluted on ice with 800 µL IP buffer (+ NP40), transferred to a new tube and centrifuged twice at 17000g, 4°C for 5 min collecting the supernatant. 2 mg of protein lysate, measured by Bradford assay, and 4 µL DNase I were added to 50 µL equilibrated IgG Dynabeads in 1 ml final volume of IP buffer (+ NP40). Immunoprecipitation of tagged proteins proceeded for 3 h at 4°C with gentle rotation. Afterwards, beads were washed four times with 1 mL of IP buffer (+NP40), resuspended in 50 µL Laemmelli buffer, boiled at 75°C for 5 min. Input samples were prepared by mixing 116.7 µg of protein lysate with 35 µL of Laemmelli buffer. 20% of the total IP and an input equal to 12.5% of the loaded IP were resolved by SDS-PAGE following immunoblotting as described above.

4.2.10 RNA extraction and clean-ups

Cells from overnight cultures were grown to exponential phase in 15 ml of appropriate medium at the required temperature. After pelleting at 17000g, cells were resuspended in 400 µl AE-buffer and mixed with 20 µl 20% SDS and 500 µl pre-equilibrated phenol. The resulting mix was incubated for 5 min at 65°C and subsequently on ice for 5 min. Afterwards, the samples were centrifuged at 4°C for 3 min at 17000g. The supernatant was mixed with 500 µl phenol-chloroform (1:1) and incubated for 5 min at room temperature. After that, the samples were centrifuged as before and the supernatant combined with 40 µl 3M NaAc and 1 ml 100%

ethanol. The resulting mix was incubated for 30 min at -20°C. Subsequently, the samples were centrifuged as previously described and the pellet was washed with 1 ml 80% ethanol and resuspended in a solution composed of 86 µl H2O, 10 µl RDD buffer and 4 µl DNase I, followed by incubation for 1 hr at 37°C.

To further purify RNA fraction, Qiagen RNeasy Min Elute Cleanup Kit was used and RNA eluted in 30 µl of water. For purification of TERRA RNA, 50 µg RNA were purified 3 consecutive times with the RNeasy kit. Incubation with 87 µl H2O, 10 µl RDD buffer and 3 µl DNase I at 37°C took place between each clean-up step. During the final step, RNA was eluted in 30 µl H2O.

4.2.11 Reverse transcription

For reverse transcription 3 µg RNA were diluted in a final volume of 7 µl and mixed with 0.4 µl dNTPs (25 mM each stock), 2 µl random hexamers (50 ng/µl) and 4.6 µl H2O. Following, the samples were first heated for 1 min at 90°C and then brought with a gradient of ca. 0.8°C/sec to a temperature of 55°C. At this step, a mix composed of 1 µl DTT (0.1M stock), 1 µl SuperScript III Reverse Transcriptase, 1 µl RNase OUT and 4 µl First Strand buffer (5X stock) was added and samples were kept first for 1 hr at 65°C and then for 15 min at 70°C. The same reaction was performed in parallel in absence of reverse transcriptase (1 µl H2O was added instead) to generate the corresponding negative control. Once the cDNA was synthesized, 30 µl H2O were added to both the reverse transcriptase-containing samples and the controls.

4.2.12 RT-qPCR

The qPCR was performed by the CFX384 Touch Real-Time PCR Detection System (Bio-Rad) in a reaction volume of 10 µl using 2 µl cDNA was combined with 1 µl each primer (listed in section 4.1.3), 1 µl water, 5 µl SYBR-Green. The qPCR was performed with the following settings: 10 min at 95 °C; 40 cycles of 15 sec at 95 °C and 1 min at 60 °C; 5 min at 95 °C; 1 min at 65 °C. The melting curve of the amplicon was determined with a gradient from 65 °C to 97 °C with an increase of 0.5 °C/cycle in cycles of 5 sec. The data analysis was performed in CFX ManagerTM software. Ct values were determined automatically in a regression mode. The relative target gene expression level was normalized to actin levels as follows:

Ct(mean) actin = Average Ct value for actin

Ct(mean)target gene = Average of Ct values for target gene

dCt = Ct(mean)target gene – Ct(mean)actin

2(-dCt) → target gene expression relative to actin

4.2.13 Chromatin immunoprecipitation

Exponentially growing cells were cross-linked with 1.2% formaldehyde at room temperature for 10 min. Followed quenching with 115 mM glycine for 5 min and incubation on ice for at least 1 min. The samples were pelleted then at 4°C and washed twice with 20 ml cold 1X PBS. After centrifuging, the pellets were stored at -80°C.

The pellets were thawed on ice and resuspended in 400 µl cold FA lysis buffer -SOD supplemented with protease inhibitor. The samples were transferred then in lysing Matrix C tubes kept on ice and lysed with the FastPrep machine for 3 runs (each of 30 sec) at 4°C and level of 6.5 M/sec, with 1 min on ice between runs. After that, the extracts were recovered by adding 800 µl cold FA lysis buffer + SOD supplemented with protease inhibitor. After mixing, the extracts were centrifuged for 15 min at 4°C and the pellet was resuspended in 1.5 ml cold FA lysis buffer + SOD supplemented with protease inhibitor. 20 µl 20% SDS were added and 750 µl of the resulting mix were combined with 0.4 g beads kept on ice. Sonication took place for 5 cycles of 30 sec on/off at 4°C with Bioruptor Pico. The remaining volume of the mix was combined with 0.4 g beads kept on ice and sonicated in the same manner. The samples were centrifuged then for 15 min at 4°C and the supernatant, which constituted the ChIP extract, was stored at -80°C.

Sonication efficiency was determined by evaluating the average length of the sheared DNA fragments. To accomplish this, 100 µl ChIP extract were combined with 100 µl elution buffer, de-cross-linked overnight at 65°C, subjected to digestion with 7.5 µl Proteinase K and 1 µl RNase A and run on a 1.5% agarose gel for 45 min at 100V.

When pulling down the protein of interest, the protein concentration within the ChIP extract was measured by Bradford and diluted to 1 mg/ml in 2 ml cold FA lysis buffer + SOD supplemented with protease inhibitor. From this mix, 50 µl were stored as 5% input at -20°C. The remaining volume was split in 2: one half was used to pull down the protein in presence of the antibody (+Ab) and the other one as a negative control in absence of the antibody (-Ab). The proper amount of antibody (10 µl Anti-HA high affinity, 2 µl S9.6Ab, antibodies are listed in section 4.1.8) was added to the supernatant and incubation for 30min at 4°C on rotating wheel took place. Subsequently, 50 µl of the IgG Dynabeads, previously washed and

supplemented with 5% BSA, were added and the samples were incubated overnight at 4°C on rotating wheel. For TAP ChIPs, IgG sepharose beads were used, whereas for the rest Dynabeads Protein G were adopted.

The following day, beads were washed with 1 ml cold FA lysis buffer + SOD, 1 ml cold FA lysis buffer 500, 1 ml cold buffer III and 1 ml TE pH 8.0. At this point, the beads were eluted twice in 100 μl elution buffer, vortexed and incubated for 8 min at 65°C. After centrifugation, the eluate was combined with 7.5 μl Proteinase K (20 mg/ml stock) and incubated overnight at 65°C in order to be de-cross-linked. Meanwhile, the input DNA was thawed at room temperature, mixed with 150 μl elution buffer and 7.5 μl Proteinase K and incubated overnight at 65°C. Following, both the immunoprecipitation and the input samples were purified with the Minelute PCR purification kit, eluted in 50 μl water and the DNA amount measured by qPCR, as previously described.

4.2.14 Protein Purification

Various N-terminally GST-tagged RNase H1 proteins (wild-type and mutants) were purified from E. Coli. Plasmids were transformed into BL21 derivative Rosetta (DE3) cells. Cells were grown at 37°C to OD600 = 0.8. Cell pellets were resuspend in lysis buffer. Cell lysates were sonicated (20% duty cycle, 2 x 3 min.), and clarified by centrifugation (40,000 x g for 30 min). The supernatants were collected and NaCl was added (final concentration of 800 mM NaCl). 0.2% PEI was added and stirred for 5 min at 4°C. Then it was spinned down for 20 min in falcons (4000 x g). All recombinant RNase H1 proteins were run through GSTrap (binding at 1ml/min). 3C digestion was done to cut GST tag from RNase H proteins overnight in at 4°C. Digested proteins were run in reverse-GSTrap to remove GST tag. For RPA complex (Rfa1-His6-MBP, Rfa2 and Rfa3 untagged) was run through HisTrap. 3C digestion was done to cut His6-MBP tag from Rfa1 overnight in at 4°C. To remove nucleic acids, RPA proteins were run through a Heparin column. All RNase H1 proteins were concentrated using Amicon spin concentrators (Merck Milipore) and subjected to gel filtration on a Superdex 75 16/60pg (Cytiva), while RPA was concentrated using Superdex 200 16/60 pg. Peak fractions containing the recombinant proteins after gel filtration were pooled and protein concentration was determined by using absorbance spectroscopy and the respective extinction coefficient at 280 nm, before aliquots were flash frozen in liquid nitrogen and stored at −80 °C.

4.2.15 RNA:DNA hybrid substrates

Oligonucleotides used are listed in 4.1.3. Fluorescently labeled oligonucleotides were annealed to a complementary strand without label by heating to 95°C and slow cooling over a long period of time in Hybridization buffer. The substrates (1μM final concentration) were stored at -20°C and later used for *in vitro* assays. For the RNA:DNA hybrid substrate, oFB261 was annealed with oFB256. oFB258, oFB259 and oFB256 were used to make R-loop. For R:D hybrid with ssDNA overhang, oFB262 was annealed with oFB256. For the R-loop with 5' ssRNA overhang substrate, oFB270 was annealed with oFB258 and oFB259. oFB261 and oFB270 were used to anneal R:D hybrid with RNA overhang. Additionally, dsDNA was annealed using oFB260 and oFB261. oFB260 and oFB262 were used to make R:D hybrid with ssDNA overhang. oFB256, oFB260 and oFB270 were used as ssRNA or ssDNA controls.

4.2.16 EMSA

RNA:DNA hybrid substrates containing 3'-FAM-labelled RNA as well respective control (dsDNA, ssDNA and ssRNA) were incubated with various amounts of RNaseH1 and/or RPA at 25°C for 15 min in FP buffer. The resulting protein-substrate complexes were resolved on 6% or 10% polyacrylamide gels using 1 x TBE buffer[276]. Gels were imaged with Typhoon FLA9500 and analyzed with ImageJ.

4.2.17 Fluorescent polarization

Substrates containing 3'-FAM-labelled RNA or DNA were incubated with 12 serial dilutions of RNaseH1, RNase H1 truncations or RPA at 25°C for 15 min in FP buffer. Final concentration of substrates were 5nM. Fluorescent polarization was measured using Spark 20M multimode microplate reader. Graphs were plotted and analyzed with Graphpad Prism 7.

4.2.18 *In vitro* pulldown

RNase H (wild type or mutants) and/or RPA were mixed with the described concentration at 25°C for 15 min in IP buffer (+NP40). 10μL of amylose beads or anti-MBP magnetic beads were added and pulldown of RPA proceeded for 3 h at 4°C with gentle rotation. After that, beads were washed four times with 1 mL of IP buffer (+NP40), resuspended in 25 μL Laemmelli buffer, boiled at 75°C for 5 min. Input samples were prepared by mixing 10μL of protein samples with Laemmelli buffer and IP buffer+NP40 to a final volume of 25 μL. 30% of

the total IP and an input equal to 3% of the loaded IP were resolved by SDS-PAGE following Coomassie staining.

5 Abbreviations

5-FOA	5-fluoroorotic acid
A	Adenine
aa	Amino acid
AFM	Atomic force microscopy
AID	Auxin-inducible degron
bp	Base pairs
C	Cytosine
CAT	Catalytic domain
CEN	Centromeric
ChIP	Chromatin immunoprecipitation
Co-IP	Co-immunoprecipitation
CSR	Class switch recombination
D-loop	DNA loop
DIP	DNA immunoprecipitation
DNA	Deoxyribonucleic acid
DNMT	DNA (cytosine-5)-methyltransferase
DRIP	DNA-RNA immunoprecipitation
ds	Double stranded
DSB	Double strand break
E. coli	Escherichia coli
EV	Empty vector
FACS	Fluorescence-activated cell sorting
FP	Fluorescent polarization
G	Guanine
G4	G-quadruplex
GFP	Green fluorescent protein
HA	Hemaglutinin
HBD	Hybrid-binding domain
HIS	Histidine
HR	Homologous recombination
HU	Hydroxyurea
HYG	Hygromycin
IAA	Indole-3-acetic acid
IgG	Immunoglobulin G
IgH	Immunoglobulin H
KAN	Kanamycin
kb	Kilobases
kDa	KiloDaltons

LEU	Leucine
lncRNA	Long non-coding RNA
mtDNA	Mitochondrial DNA
MMS	Methyl methanesulfonate
MNase	Micrococcal nuclease
MTS	Mitochondrial targeting sequence
NAT	Nourseothricin
ncRNA	Non-coding RNA
NHEJ	Non-homologous end joining
nt	Nucleotides
ORF	Open reading frame
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PIP	PCNA-interacting protein
qPCR	Quantitative PCR
RAD51AP1	RAD51 associated protein 1
Raff	Raffinose
RBP	RNA binding protein
rDNA	Ribosomal DNA
RED	Ribonucleotide excision defective
RER	Ribonucleotide excision repair
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNase H	Ribonuclease H
rNMP	Ribonucleoside monophosphate
rNTP	Ribonucleoside triphosphate
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
RT	Reverse transcription
S. cerevisiae	Saccharomyces cerevisiae
S. pombe	Schizosaccharomyces pombe
SC	Synthetic complete
SD	Standard deviation
Sen1	Senataxin (yeast homolog)
ss	Single stranded
STX	Senataxin
T	Thymine
TA-HRR	Transcription-associated homologous recombination repair
TC-NER	Transcription-coupled nucleotide excision repair
TERRA	Telomeric repeat containing RNA
tRNA	Transfer RNA

TRP	Tryptophan
TSS	Transcription start site
URA	Uracil

6 References

1. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription Separation of RNA from DNA by RNA polymerase II. *Science (80-.).* **303**, 1014–1016 (2004).

2. Rich, A. A Hybrid Helix Containing Both Deoxyribose and Ribose Polynucleotides and Its Relation To the Transfer of Information Between the Nucleic Acids. *Proc. Natl. Acad. Sci.* **46**, 1044–1053 (1960).

3. Thomas, M., White, R. L. & Davis, R. W. Hybridization of RNA to double stranded DNA: Formation of R loops. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 2294–2298 (1976).

4. García-Muse, T. & Aguilera, A. R Loops: From Physiological to Pathological Roles. *Cell* **179**, 604–618 (2019).

5. Santos-Pereira, J. M. & Aguilera, A. R loops: New modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 (2015).

6. Aguilera, A. & García-Muse, T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* **46**, 115–124 (2012).

7. Sollier, J. & Cimprich, K. A. Breaking bad: R-loops and genome integrity. *Trends Cell Biol.* **25**, 514–522 (2015).

8. Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Mol. Cell* **73**, 398–411 (2019).

9. Niehrs, C. & Luke, B. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.* **21**, 167–178 (2020).

10. Roberts, R. W. & Crothers, D. M. Stability and properties of double and triple helices: Dramatic effects of RNA or DNA backbone composition. *Science (80-.).* **258**, 1463–1466 (1992).

11. Shaw, N. N. & Arya, D. P. Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* **90**, 1026–1039 (2008).

12. Stodola, J. L. & Burgers, P. M. Mechanism of lagging-strand DNA replication in eukaryotes. *Adv. Exp. Med. Biol.* **1042**, 117–133 (2017).

13. Nick McElhinny, S. A. *et al.* Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4949–4954 (2010).

14. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol. Cell* **50**, 437–443 (2013).

15. Ghodgaonkar, M. M. *et al.* Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol. Cell* **50**, 323–332 (2013).

16. Pryor, J. M. *et al.* Ribonucleotide incorporation enables repair of chromosome breaks by nonhomologous end joining. *Science (80-.).* **361**, 1126–1129 (2018).

17. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: Nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489 (2004).

18. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7024–7027 (1987).

19. Roy, D., Zhang, Z., Lu, Z., Hsieh, C.-L. & Lieber, M. R. Competition between the RNA Transcript and the Nontemplate DNA Strand during R-Loop Formation In Vitro: a Nick Can Serve as a Strong R-Loop Initiation Site. *Mol. Cell. Biol.* **30**, 146–159 (2010).

20. Chen, L. *et al.* R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol. Cell* **68**, 745-757.e5 (2017).

21. Roy, D., Yu, K. & Lieber, M. R. Mechanism of R-Loop Formation at Immunoglobulin Class Switch Sequences. *Mol. Cell. Biol.* **28**, 50–60 (2008).

22. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711–721 (2003).

23. Domínguez-Sánchez, M. S., Barroso, S., Gómez-González, B., Luna, R. & Aguilera, A. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* **7**, 19–22 (2011).

24. Mischo, H. E. *et al.* Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* **41**, 21–32 (2011).

25. Wahba, L., Gore, S. K. & Koshland, D. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *Elife* **2013**, 1–20 (2013).

26. Hage, A. El, French, S. L., Beyer, A. L. & Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. 1546–1558 (2010). doi:10.1101/gad.573310.Freely

27. Chan, Y. A. *et al.* Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip. *PLoS Genet.* **10**, (2014).

28. Ouyang, J. *et al.* RNA transcripts stimulate homologous recombination by forming DR-loops. *Nature* **594**, 283–288 (2021).

29. Kasahara, M., Clikeman, J. A., Bates, D. B. & Kogoma, T. RecA protein-dependent R-loop formation in vitro. *Genes Dev.* 360–365 (2000).

30. Zaitsev, E. N. & Kowalczykowski, S. C. A novel pairing process promoted by Escherichia coli

RecA protein: Inverse DNA and RNA strand exchange. *Genes Dev.* **14**, 740–749 (2000).

31. Xiao, Y. *et al.* Structure Basis for Directional R-loop Formation and Substrate Handover Mechanisms in Type I CRISPR-Cas System. *Cell* **170**, 48-60.e11 (2017).

32. Stolz, R. *et al.* Interplay between DNA sequence and negative superhelicity drives R-loop structures. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 6260–6269 (2019).

33. El Hage, A., Webb, S., Kerr, A. & Tollervey, D. Genome-Wide Distribution of RNA-DNA Hybrids Identifies RNase H Targets in tRNA Genes, Retrotransposons and Mitochondria. *PLoS Genet.* **10**, (2014).

34. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chédin, F. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Mol. Cell* **45**, 814–825 (2012).

35. Xu, W. *et al.* The R-loop is a common chromatin feature of the Arabidopsis genome. *Nat. Plants* **3**, 704–714 (2017).

36. Arora, R. *et al.* RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* **5**, 5220 (2014).

37. Nadel, J. *et al.* RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics and Chromatin* **8**, 1–19 (2015).

38. Dumelie, J. G. & Jaffrey, S. R. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. *Elife* **6**, 1–39 (2017).

39. Crossley, M. P., Bocek, M. J., Hamperl, S., Swigut, T. & Cimprich, K. A. QDRIP: A method to quantitatively assess RNA-DNA hybrid formation genome-wide. *Nucleic Acids Res.* **48**, (2020).

40. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Mol. Cell* **42**, 794–805 (2011).

41. Yanling Zhao, D. *et al.* SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature* **529**, 48–53 (2016).

42. Brickner, J. R., Garzon, J. L. & Cimprich, K. A. Walking a tightrope: The complex balancing act of R-loops in genome stability. *Mol. Cell* (2022). doi:10.1016/j.molcel.2022.04.014

43. Zatreanu, D. *et al.* Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. *Mol. Cell* **76**, 57-69.e9 (2019).

44. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zonj, G. & Wilson, W. D. Sequence Specific Thermodynamic and Structural Properties for DNA* *RNA Duplexes. *Biochemistry* 5298–5304 (1994).

45. De Magis, A. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 816–825 (2019).

46. Kuznetsov, V. A., Bondarenko, V., Wongsurawat, T., Yenamandra, S. P. & Jenjaroenpun, P. Toward predictive R-loop computational biology: Genome-scale prediction of R-loops reveals their association with complex promoter structures, G-quadruplexes and transcriptionally active enhancers. *Nucleic Acids Res.* **46**, 7566–7585 (2018).

47. Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F. & Maizels, N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618–1629 (2004).

48. Carrasco-Salas, Y. *et al.* The extruded non-template strand determines the architecture of R-loops. *Nucleic Acids Res.* **47**, 6783–6795 (2019).

49. Wahba, L., Amon, J. D., Koshland, D. & Vuica-Ross, M. RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability. *Mol. Cell* **44**, 978–988 (2011).

50. Sanz, L. A. *et al.* Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol. Cell* **63**, 167–178 (2016).

51. Groh, M. & Gromak, N. Out of Balance: R-loops in Human Disease. *PLoS Genet.* **10**, (2014).

52. Richard, P. & Manley, J. L. R Loops and Links to Human Disease. *J. Mol. Biol.* **429**, 3168–3180 (2017).

53. Marabitti, V. *et al.* ATM pathway activation limits R-loop-associated genomic instability in Werner syndrome cells. *Nucleic Acids Res.* **47**, 3485–3502 (2019).

54. Barnhoorn, S. *et al.* Cell-Autonomous Progeroid Changes in Conditional Mouse Models for Repair Endonuclease XPG Deficiency. *PLoS Genet.* **10**, 7–9 (2014).

55. Trego, K. S. *et al.* The DNA repair endonuclease XPG interacts directly and functionally with the WRN helicase defective in Werner syndrome. *Cell Cycle* **10**, 1998–2007 (2011).

56. Niedernhofer, L. J. *et al.* A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444**, 1038–1043 (2006).

57. Pohjoismäki, J. L. O. *et al.* Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of RNA/DNA hybrid. *J. Mol. Biol.* **397**, 1144–1155 (2010).

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

59. Chédin, F., Hartono, S. R., Sanz, L. A. & Vanoosthuyse, V. Best practices for the visualization, mapping, and manipulation of R-loops. *EMBO J.* **40**, 1–13 (2021).

60. Halász, L. *et al.* RNA-DNA hybrid (R-loop) immunoprecipitation mapping: An analytical workflow to evaluate inherent biases. *Genome Res.* **27**, 1063–1073 (2017).

61. Chen, P. B., Chen, H. V., Acharya, D., Rando, O. J. & Fazzio, T. G. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.* **22**, 999–1007 (2015).

62. Phillips, D. D. *et al.* The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. *J. Mol. Recognit.* **26**, 376–381 (2013).

63. König, F., Schubert, T. & Längst, G. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. *PLoS One* **12**, 1–13 (2017).

64. Sanz, L. A. & Chédin, F. High-resolution, strand-specific R-loop mapping via S9.6-based DNA–RNA immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* **14**, 1734–1755 (2019).

65. Hartono, S. R. *et al.* The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *J. Mol. Biol.* **430**, 272–284 (2018).

66. Wahba, L., Costantino, L., Tan, F. J., Zimmer, A. & Koshland, D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* **30**, 1327–1338 (2016).

67. Wu, H., Lima, W. F. & Crooke, S. T. Investigating the Structure of Human RNase H1 by Site-directed Mutagenesis. *J. Biol. Chem.* **276**, 23547–23553 (2001).

68. Yan, Q., Shields, E. J., Bonasio, R. & Sarma, K. Mapping Native R-Loops Genome-wide Using a Targeted Nuclease Approach. *Cell Rep.* **29**, 1369-1380.e5 (2019).

69. Wang, K. *et al.* Genomic profiling of native R loops with a DNA-RNA hybrid recognition sensor. *Sci. Adv.* **7**, 1–18 (2021).

70. Stork, C. T. *et al.* Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *Elife* **5**, 1–21 (2016).

71. Vanoosthuyse, V. & Vincent. Strengths and Weaknesses of the Current Strategies to Map and Characterize R-Loops. *Non-Coding RNA* **4**, 9 (2018).

72. Nowotny, M. *et al.* Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD. *EMBO J.* **27**, 1172–1181 (2008).

73. Kabeche, L., Nguyen, H. D., Buisson, R. & Zou, L. A mitosis-specific and R loop–driven ATR pathway promotes faithful chromosome segregation. *Science (80-.).* 108–114 (2018). doi:10.1016/S1369-5274(01)00265-X

74. Li, Y. *et al.* R-loops coordinate with SOX2 in regulating reprogramming to pluripotency. *Sci. Adv.* **6**, (2020).

75. Malig, M., Hartono, S. R., Giafaglione, J. M., Sanz, L. A. & Chedin, F. Ultra-deep Coverage Single-molecule R-loop Footprinting Reveals Principles of R-loop Formation. *J. Mol. Biol.* (2020). doi:10.1016/j.jmb.2020.02.014

76. Lockhart, A. *et al.* RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. *Cell Rep.* **29**, 2890-2900.e5 (2019).

77. Chaudhuri, J. *et al.* Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422**, 726–730 (2003).

78. Powell, W. T. *et al.* R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13938–13943 (2013).

79. Grunseich, C. *et al.* Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. *Mol. Cell* **69**, 426-437.e7 (2018).

80. Arab, K. *et al.* GADD45A binds R-loops and recruits TET1 to CpG island promoters. *Nat. Genet.* **51**, 217–223 (2019).

81. Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N. J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436–439 (2014).

82. Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **52**, 583–590 (2013).

83. Nakama, M., Kawakami, K., Kajitani, T., Urano, T. & Murakami, Y. DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation. *Genes to Cells* **17**, 218–233 (2012).

84. Beckedorff, F. C. *et al.* The Intronic Long Noncoding RNA ANRASSF1 Recruits PRC2 to the RASSF1A Promoter, Reducing the Expression of RASSF1A and Increasing Cell Proliferation. *PLoS Genet.* **9**, (2013).

85. Cloutier, S. C. *et al.* Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation. *Mol. Cell* **61**, 393–404 (2016).

86. Boque-Sastre, R. *et al.* Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5785–5790 (2015).

87. Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J. & Dean, C. R-Loop Stabilization Represses Antisense Transcription at the Arabidopsis FLC Locus. *Science (80-.).* **340**, 619–621 (2013).

88. Belotserkovskii, B. P., Shin, J. H. S. & Hanawalt, P. C. 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 (2017).

89. Bonnet, A. *et al.* Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic

Instability. *Mol. Cell* **67**, 608-621.e6 (2017).

90. Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* **170**, 774-786.e19 (2017).

91. Lang, K. S. *et al.* Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell* **170**, 787-799.e18 (2017).

92. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chédin, F. GC skew at the 59 and 39 ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res.* **23**, 1590–1600 (2013).

93. Huppert, J. L., Bugaut, A., Kumari, S. & Balasubramanian, S. G-quadruplexes: The beginning and end of UTRs. *Nucleic Acids Res.* **36**, 6260–6268 (2008).

94. Belotserkovskii, B. P. *et al.* Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12816–12821 (2010).

95. Kireeva, M. L., Komissarova, N. & Kashlev, M. Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. *J. Mol. Biol.* **299**, 325–335 (2000).

96. Cristini, A., Groh, M., Kristiansen, M. S. & Gromak, N. RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Rep.* **23**, 1891–1905 (2018).

97. Morales, J. C. *et al.* XRN2 Links Transcription Termination to DNA Damage and Replication Stress. *PLoS Genet.* **12**, 1–22 (2016).

98. Mersaoui, S. Y. *et al.* Arginine methylation of the DDX 5 helicase RGG / RG motif by PRMT 5 regulates resolution of RNA:DNA hybrids . *EMBO J.* **38**, 1–20 (2019).

99. Rivosecchi, J. *et al.* Senataxin homologue Sen 1 is required for efficient termination of RNA polymerase III transcription. 1–14 (2019). doi:10.15252/embj.2019101955

100. Hatchi, E. *et al.* BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **57**, 636–647 (2015).

101. Yu, Z. *et al.* DDX5 resolves R-loops at DNA double-strand breaks to promote DNA repair and avoid chromosomal deletions. *NAR Cancer* **2**, 1–19 (2020).

102. Luna, R. *et al.* Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. *Mol. Cell* **18**, 711–722 (2005).

103. Wimberly, H. *et al.* R-loops and nicks initiate DNA breakage and genome instability in non-growing Escherichia coli. *Nat. Commun.* **4**, (2013).

104. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).

105. Beletskii, A. & Bhagwat, A. S. Transcription-induced mutations: Increase in C to T mutations in the nontranscribed strand during transcription in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13919–13924 (1996).

106. Polak, P. & Arndt, P. F. Transcription induces strand-specific mutations at the 5′ end of human genes. *Genome Res.* **18**, 1216–1223 (2008).

107. Chaudhuri, J., Khuong, C. & Alt, F. W. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* **430**, 992–998 (2004).

108. Gómez-González, B. & Aguilera, A. Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8409–8414 (2007).

109. Tan, J., Wang, X., Phoon, L., Yang, H. & Lan, L. Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase. *FEBS Lett.* **594**, 1359–1367 (2020).

110. Tan, J. *et al.* An R-loop-initiated CSB-RAD52-POLD3 pathway suppresses ROS-induced telomeric DNA breaks. *Nucleic Acids Res.* **48**, 1285–1300 (2020).

111. Teng, Y. *et al.* ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB. *Nat. Commun.* **9**, (2018).

112. Kogoma, T. Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* **61**, 212–238 (1997).

113. Stuckey, R., García-Rodríguez, N., Aguilera, A. & Wellinger, R. E. Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5779–5784 (2015).

114. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).

115. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell* **47**, 497–510 (2012).

116. Zhao, H., Zhu, M., Limbo, O. & Russell, P. RNase H eliminates R-loops that disrupt DNA replication but is nonessential for efficient DSB repair. *EMBO Rep.* **19**, 1–10 (2018).

117. Ortega, P., Mérida-Cerro, J. A., Rondón, A. G., Gómez-González, B. & Aguilera, A. DNA-RNA hybrids at DSBs interfere with repair by homologous recombination. *Elife* **10**, 1–22 (2021).

118. Nimonkar, A. V. *et al.* BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.* **25**, 350–362 (2011).

119. Sturzenegger, A. *et al.* DNA2 cooperates with the WRN and BLM RecQ helicases to mediate

long-range DNA end resection in human cells. *J. Biol. Chem.* **289**, 27314–27326 (2014).

120. Alfano, L. *et al.* Depletion of the RNA binding protein HNRNPD impairs homologous recombination by inhibiting DNA-end resection and inducing R-loop accumulation. *Nucleic Acids Res.* **47**, 4068–4085 (2019).

121. Matsui, M. *et al.* USP42 enhances homologous recombination repair by promoting R-loop resolution with a DNA–RNA helicase DHX9. *Oncogenesis* **9**, (2020).

122. Sessa, G. *et al.* BRCA2 promotes DNA-RNA hybrid resolution by DDX5 helicase at DNA breaks to facilitate their repair‡. *EMBO J.* **40**, 1–25 (2021).

123. D'Alessandro, G. *et al.* BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nat. Commun.* **9**, (2018).

124. Cohen, S. *et al.* Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nat. Commun.* **9**, (2018).

125. Daley, J. M. *et al.* Specificity of end resection pathways for double-strand break regions containing ribonucleotides and base lesions. *Nat. Commun.* **11**, 1–12 (2020).

126. Amon, J. D., Koshland, D. & States, U. RNase H enables efficient repair of R-loop induced DNA damage. 1–20 (2016). doi:10.7554/eLife.20533

127. Costantino, L. & Koshland, D. Genome-wide Map of R-Loop-Induced Damage Reveals How a Subset of R-Loops Contributes to Genomic Instability. *Mol. Cell* **0**, 1–11 (2018).

128. Ohle, C. *et al.* Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* **167**, 1001-1013.e7 (2016).

129. Marnef, A. & Legube, G. R-loops as Janus-faced modulators of DNA repair. *Nat. Cell Biol.* **23**, 305–313 (2021).

130. Rawal, C. C. *et al.* Senataxin Ortholog Sen1 Limits DNA:RNA Hybrid Accumulation at DNA Double-Strand Breaks to Control End Resection and Repair Fidelity. *Cell Rep.* **31**, 107603 (2020).

131. Jang, Y. *et al.* Intrinsically disordered protein RBM14 plays a role in generation of RNA:DNA hybrids at double-strand break sites. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 5329–5338 (2020).

132. Schellenbauer, A. *et al.* Phospho-Ku70 induced by DNA damage interacts with RNA Pol II and promotes the formation of phospho-53BP1 foci to ensure optimal cNHEJ. *Nucleic Acids Res.* **49**, 11728–11745 (2021).

133. Simon, N. E., Yuan, M. & Kai, M. RNA-binding protein RBM14 regulates dissociation and association of non-homologous end joining proteins. *Cell Cycle* **16**, 1175–1180 (2017).

134. Domingo-Prim, J. *et al.* EXOSC10 is required for RPA assembly and controlled DNA end resection at DNA double-strand breaks. *Nat. Commun.* **10**, (2019).

135. Li, L. *et al.* DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol. Cell. Biol.* **36**, 2794–2810 (2016).

136. Liu, S. *et al.* RNA polymerase III is required for the repair of DNA double-strand breaks by homologous recombination. *Cell* **184**, 1314-1329.e10 (2021).

137. Yasuhara, T. *et al.* Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. *Cell* **175**, 558-570.e11 (2018).

138. Hamperl, S. & Cimprich, K. A. Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell* **167**, 1455–1467 (2016).

139. Lalonde, M., Trauner, M., Werner, M. & Hamperl, S. Consequences and resolution of transcription–replication conflicts. *Life* **11**, (2021).

140. Petryk, N. *et al.* Replication landscape of the human genome. *Nat. Commun.* **7**, 1–13 (2016).

141. Chen, Y. H. *et al.* Transcription shapes DNA replication initiation and termination in human cells. *Nat. Struct. Mol. Biol.* **26**, 67–77 (2019).

142. Helmrich, A., Ballarino, M. & Tora, L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Mol. Cell* **44**, 966–977 (2011).

143. Zeman, M. K. & Cimprich, K. A. Causes and consequences of replication stress. *Nat. Cell Biol.* **16**, 2–9 (2014).

144. Prado, F. & Aguilera, A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J.* **24**, 1267–1276 (2005).

145. Gan, W. *et al.* R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* **25**, 2041–2056 (2011).

146. García-Muse, T. & Aguilera, A. Transcription-replication conflicts: How they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol.* **17**, 553–563 (2016).

147. Wellinger, R. E., Prado, F. & Aguilera, A. Replication Fork Progression Is Impaired by Transcription in Hyperrecombinant Yeast Cells Lacking a Functional THO Complex. *Mol. Cell. Biol.* **26**, 3327–3334 (2006).

148. Kumar, C., Batra, S., Griffith, J. D. & Remus, D. The interplay of RNA:DNA hybrid structure and G-quadruplexes determines the outcome of R-loop-replisome collisions. *Elife* **10**, 1–34 (2021).

149. Schauer, G. D., Spenkelink, L. M., Lewis, J. S., Yurieva, O. & Mueller, S. H. Replisome bypass of a protein-based R-loop block by Pif1. 1–8 (2020). doi:10.1073/pnas.2020189117

150. Brüning, J. G. & Marians, K. J. Replisome bypass of transcription complexes and R-loops. *Nucleic Acids Res.* **48**, 10353–10367 (2020).

151.   Luke, B. *et al.* The Rat1p 5' to 3' Exonuclease Degrades Telomeric Repeat-Containing RNA and Promotes Telomere Elongation in Saccharomyces cerevisiae. *Mol. Cell* **32**, 465–477 (2008).

152.   Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric Repeat–Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* **318**, 798–801 (2007).

153.   Blasco, M. & Schoeftner, S. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol.* **10**, 228–36 (2008).

154.   Lopez de Silanes, I. *et al.* Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun* **5**, 4723 (2014).

155.   Arora, R. & Azzalin, C. M. Telomere elongation chooses TERRA ALTernatives. *RNA Biol.* **12**, 938–941 (2015).

156.   Montero J.J., López de Silanes I., Graña O., B. M. A. Telomeric RNAs are essential to maintain telomeres. *Nat Commun* **in press**, 1–13 (2016).

157.   Balk, B. *et al.* Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* **20**, 1199–1205 (2013).

158.   Graf, M. *et al.* Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* **170**, 72-85.e14 (2017).

159.   Holt, I. J. The mitochondrial R-loop. *Nucleic Acids Res.* **47**, 5480–5489 (2019).

160.   Lima, W. F. *et al.* Viable RNaseH1 knockout mice show RNaseH1 is essential for R loop processing, mitochondrial and liver function. *Nucleic Acids Res.* **44**, 5299–5312 (2016).

161.   Posse, V., Id, A. A., Uhler, J. P., Id, A. R. C. & Id, A. R. RNase H1 directs origin-specific initiation of DNA replication in human mitochondria. *PLoS Genet.* 1–20 (2019).

162.   Cheng, L., Wang, W., Yao, Y. & Sun, Q. Mitochondrial RNase H1 activity regulates R-loop homeostasis to maintain genome integrity and enable early embryogenesis in Arabidopsis. *PLoS Biol.* **19**, 4–6 (2021).

163.   Renaudin, X., Lee, M., Shehata, M., Surmann, E. M. & Venkitaraman, A. R. BRCA2 deficiency reveals that oxidative stress impairs RNaseH1 function to cripple mitochondrial DNA maintenance. *Cell Rep.* **36**, 109478 (2021).

164.   García-Benítez, F., Gaillard, H. & Aguilera, A. Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. *Proc. Natl. Acad. Sci.* **114**, 201707845 (2017).

165.   Gavaldá, S., Santos-Pereira, J. M., García-Rubio, M. L., Luna, R. & Aguilera, A. Excess of Yra1 RNA-Binding Factor Causes Transcription-Dependent Genome Instability, Replication

Impairment and Telomere Shortening. *PLoS Genet.* **12**, (2016).

166.   Abakir, A. *et al.* N 6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nat. Genet.* **52**, 48–55 (2020).

167.   Chen, H. *et al.* m5C modification of mRNA serves a DNA damage code to promote homologous recombination. *Nat. Commun.* **11**, 3–14 (2020).

168.   Zhang, C. *et al.* METTL3 and N6-Methyladenosine Promote Homologous Recombination-Mediated Repair of DSBs by Modulating DNA-RNA Hybrid Accumulation. *Mol. Cell* **79**, 425-442.e7 (2020).

169.   Salas-Armenteros, I. *et al.* Human THO –Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability . *EMBO J.* **36**, 3532–3547 (2017).

170.   Costantino, L. & Koshland, D. The Yin and Yang of R-loop biology. *Curr. Opin. Cell Biol.* **34**, 39–45 (2015).

171.   Ribeiro de Almeida, C. *et al.* RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol. Cell* **70**, 650-662.e8 (2018).

172.   Schwab, R. A. *et al.* The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol. Cell* **60**, 351–361 (2015).

173.   Tedeschi, F. A., Cloutier, S. C., Tran, E. J. & Jankowsky, E. The DEAD-box protein Dbp2p is linked to noncoding RNAs, the helicase Sen1p, and R-loops. *RNA* **24**, 1693–1705 (2018).

174.   Wang, I. X. *et al.* Human proteins that interact with RNA/DNA hybrids. *Genome Res.* **28**, 1405–1414 (2018).

175.   Song, C., Hotz-Wagenblatt, A., Voit, R. & Grummt, I. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.* **31**, 1370–1381 (2017).

176.   Yun, E. *et al.* RECQ-like helicases Sgs1 and BLM regulate R-loop– associated genome instability. *J. Cell Biol.* **216**, 3991–4005 (2017).

177.   Sollier, J. *et al.* Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Mol. Cell* **56**, 777–785 (2014).

178.   Sridhara, S. C. *et al.* Transcription Dynamics Prevent RNA-Mediated Genomic Instability through SRPK2-Dependent DDX23 Phosphorylation. *Cell Rep.* **18**, 334–343 (2017).

179.   Tran, P. L. T. *et al.* PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. *Nat. Commun.* **8**, 1–10 (2017).

180.   Mosler, T. *et al.* R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat. Commun.* **12**, 1–17 (2021).

181. Hodroj, D. *et al.* An ATR -dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism . *EMBO J.* **36**, 1182–1198 (2017).

182. Lafuente-Barquero, J. *et al.* The Smc5/6 complex regulates the yeast Mph1 helicase at RNA-DNA hybrid-mediated DNA damage. *PLoS Genet.* **13**, 1–25 (2017).

183. Saponaro, M. *et al.* RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* **157**, 1037–1049 (2014).

184. Li, M., Pokharel, S., Wang, J. T., Xu, X. & Liu, Y. RECQ5-dependent SUMOylation of DNA topoisomerase I prevents transcription-associated genome instability. *Nat. Commun.* **6**, 1–13 (2015).

185. Kim, H. D., Choe, J. & Seo, Y. S. The sen1+ gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. *Biochemistry* **38**, 14697–14710 (1999).

186. Steinmetz, E. J., Conrad, N. K., Brow, D. A. & Corden, J. L. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. *Nature* **413**, 327–331 (2001).

187. Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S. & Meinhart, A. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* **15**, 795–804 (2008).

188. Xie, J. *et al.* An integrated model for termination of RNA polymerase III transcription. *Sci. Adv.* **8**, 19–22 (2022).

189. Appanah, R. *et al.* Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability Report Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability. *CellReports* **30**, 2094-2105.e9 (2020).

190. Stein, H. & Hausen, P. Enzyme from Calf Thymus Degrading the RNA Moiety of DNA-RNA Hybrids: Effect on DNA-Dependent RNA Polymerase. *Science (80-.).* 393–396 (1969).

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

192. Hyjek, M., Figiel, M. & Nowotny, M. RNases H_ Structure and mechanism. *DNA Repair (Amst.).* 102672 (2019). doi:10.1016/j.dnarep.2019.102672

193. Nowotny, M. Retroviral integrase superfamily: The structural perspective. *EMBO Rep.* **10**, 144–151 (2009).

194. Pileur, F., Toulme, J. J. & Cazenave, C. Eukaryotic ribonucleases HI and HII generate characteristic hydrolytic patterns on DNA-RNA hybrids: Further evidence that mitochondrial RNase H is an RNase HII. *Nucleic Acids Res.* **28**, 3674–3683 (2000).

195. Wahba, L., Amon, J. D., Koshland, D. & Vuica-Ross, M. RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability. *Mol. Cell* **44**, 978–988 (2011).

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

197. Crow, Y. J. *et al.* Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. *Nat. Genet.* **38**, 910–916 (2006).

198. Rice, G. *et al.* Clinical and molecular phenotype of Aicardi-Goutières syndrome. *Am. J. Hum. Genet.* **81**, 713–725 (2007).

199. Sparks, J. L. *et al.* RNase H2-Initiated Ribonucleotide Excision Repair. *Mol. Cell* **47**, 980–986 (2012).

200. Williams, J. S., Lujan, S. A. & Kunkel, T. A. Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat. Rev. Mol. Cell Biol.* **17**, 350–363 (2016).

201. Huang, S. N., Williams, J. S., Arana, M. E., Kunkel, T. A. & Pommier, Y. Topoisomerase I-mediated cleavage at unrepaired ribonucleotides generates DNA double-strand breaks. *EMBO J.* **36**, 361–373 (2017).

202. Williams, J. S. *et al.* Topoisomerase 1-Mediated Removal of Ribonucleotides from Nascent Leading-Strand DNA. *Mol. Cell* **49**, 1010–1015 (2013).

203. Kim, N. *et al.* Mutagenic Processing of Ribonucleotides in DNA by Yeast Topoisomerase I. *Science (80-.).* **332**, 1561–1564 (2011).

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

205. Zimmer, A. D. & Koshland, D. Differential roles of the RNases H in preventing chromosome instability. *Proc. Natl. Acad. Sci.* **113**, 12220–12225 (2016).

206. Arudchandran, A. *et al.* The absence of ribonuclease H1 or H2 alters the sensitivity of Saccharomyces cerevisiae to hydroxyurea, caffeine and ethyl methanesulphonate: Implications for roles of RNases H in DNA replication and repair. *Genes to Cells* **5**, 789–802 (2000).

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

208. Rychlik, M. P. *et al.* Crystal structures of rnase h2 in complex with nucleic acid reveal the mechanism of RNA-DNA junction recognition and cleavage. *Mol. Cell* **40**, 658–670 (2010).

209. Chon, H. *et al.* Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex. *Nucleic Acids Res.* **37**, 96–110

(2009).

210. Lim, Y. W., Sanz, L. A., Xu, X., Hartono, S. R. & Chédin, F. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome. *Elife* **4**, 1–21 (2015).

211. Cerritelli, S. M. *et al.* Failure to Produce Mitochondrial DNA Results in Embryonic Lethality in Rnaseh1 Null Mice strand mode of DNA synthesis and that the RNA-DNA hybrids are RNA primers of DNA replication. The pres-ence of abundant RNA-DNA hybrids in replicating Laboratory of M. *Mol. Cell* **11**, 807–815 (2003).

212. Williamson, D. The curious history of yeast mitochondrial DNA. *Nat. Rev. Genet.* **3**, 475–481 (2002).

213. Paulsen, R. D. *et al.* A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol. Cell* **35**, 228–239 (2009).

214. Britton, S. *et al.* DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Res.* **42**, 9047–9062 (2014).

215. Tan-Wong, S. M., Dhir, S. & Proudfoot, N. J. R-Loops Promote Antisense Transcription across the Mammalian Genome. *Mol. Cell* **76**, 600-616.e6 (2019).

216. Shen, W. *et al.* Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops. *Nucleic Acids Res.* **45**, 10672–10692 (2017).

217. Skourti-Stathaki, K. *et al.* R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes. *Mol. Cell* **73**, 930-945.e4 (2019).

218. Cerritelli, S. M., Fedoroff, O. Y., Reid, B. R. & Crouch, R. J. A common 40 amino acid motif in eukaryotic RNases H1 and caulimovirus ORF VI proteins binds to duplex RNAs. *Nucleic Acids Res.* **26**, 1834–1840 (1998).

219. Cerritelli, S. M. & Crouch, R. J. The non-RNase H domain of Saccharomyces cerevisiae RNase H1 binds double-stranded RNA: Magnesium modulates the switch between double-stranded RNA binding and RNase H activity. *Rna* **1**, 105–109 (1995).

220. Evans, S. P. & Bycroft, M. NMR structure of the N-terminal domain of Saccharomyces cerevisiae RNase HI reveals a fold with a strong resemblance to the N-terminal domain of ribosomal protein L9. *J. Mol. Biol.* **291**, 661–669 (1999).

221. Gaidamakov, S. A. *et al.* Eukaryotic RNases H1 act processively by interactions through the duplex RNA-binding domain. *Nucleic Acids Res.* **33**, 2166–2175 (2005).

222. Nowotny, M. *et al.* Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Mol. Cell* **28**, 264–276 (2007).

223. Lima, W. F. *et al.* Human RNase H1 discriminates between subtle variations in the structure of

the heteroduplex substrate. *Mol. Pharmacol.* **71**, 83–91 (2007).

224. Lima, W. F. *et al.* The positional influence of the helical geometry of the heteroduplex substrate on human RNase H1 catalysis. *Mol. Pharmacol.* **71**, 73–82 (2007).

225. Ohtani, N. *et al.* 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 (1999).

226. Nowotny, M., Gaidamakov, S. A., Crouch, R. J. & Yang, W. Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis. *Cell* **121**, 1005–1016 (2005).

227. Wu, H., Lima, W. F. & Crooke, S. T. Properties of cloned and expressed human RNase H1. *J. Biol. Chem.* **274**, 28270–28278 (1999).

228. Keck, J. L., Goedken, E. R. & Marqusee, S. Activation/attenuation model for RNase H: A one-metal mechanism with second-metal inhibition. *J. Biol. Chem.* **273**, 34128–34133 (1998).

229. Goedken, E. R. & Marqusee, S. Co-crystal of Escherichia coli RNase H with Mn2+ Ions Reveals Two Divalent Metals Bound in the Active Site. *J. Biol. Chem.* **276**, 7266–7271 (2001).

230. Nowotny, M. & Yang, W. Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release. *EMBO J.* **25**, 1924–1933 (2006).

231. Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).

232. Maréchal, A. *et al.* PRP19 Transforms into a Sensor of RPA-ssDNA after DNA Damage and Drives ATR Activation via a Ubiquitin-Mediated Circuitry. *Mol. Cell* **53**, 235–246 (2014).

233. Nguyen, H. D. *et al.* Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* **65**, 832-847.e4 (2017).

234. Mazina, O. M. *et al.* Replication protein A binds RNA and promotes R-loop formation. *J. Biol. Chem.* **295**, 14203–14213 (2020).

235. Nguyen, H. D. *et al.* Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* **in**-**press**, 832-847.e4 (2017).

236. Castillo-Guzman, D. & Chédin, F. Defining R-loop classes and their contributions to genome instability. *DNA Repair (Amst).* **106**, (2021).

237. Aiello, U. *et al.* Sen1 is a key regulator of transcription-driven conflicts. *Mol. Cell* **82**, 2952-2966.e6 (2022).

238. Kawauchi, J., Mischo, H., Braglia, P., Rondon, A. & Proudfoot, N. J. Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* **22**, 1082–1092 (2008).

239. Sheridan, R. M., Fong, N., D'Alessandro, A. & Bentley, D. L. Widespread Backtracking by RNA Pol II Is a Major Effector of Gene Activation, 5′ Pause Release, Termination, and Transcription Elongation Rate. *Mol. Cell* **73**, 107-118.e4 (2019).

240. Mishra, S., Hasan, S. H., Sakhawala, R. M., Chaudhry, S. & Maraia, R. J. Mechanism of RNA polymerase III termination-associated reinitiation-recycling conferred by the essential function of the N terminal-and-linker domain of the C11 subunit. *Nat. Commun.* **12**, 1–16 (2021).

241. Arimbasseri, A. G., Rijal, K. & Maraia, R. J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5**, 1–13 (2013).

242. Abraham, K. J. *et al.* Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature* **585**, 298–302 (2020).

243. Steinmetz, E. J. *et al.* Genome-Wide Distribution of Yeast RNA Polymerase II and Its Control by Sen1 Helicase. *Mol. Cell* **24**, 735–746 (2006).

244. Lai, F., Damle, S. S., Ling, K. K. & Rigo, F. Directed RNase H Cleavage of Nascent Transcripts Causes Transcription Termination. *Mol. Cell* **77**, 1032-1043.e4 (2020).

245. Feng, S. & Manley, J. L. Replication protein A associates with nucleolar R loops and regulates rRNA transcription and nucleolar morphology. *Genes Dev.* **35**, 1579–1594 (2021).

246. Manzo, S. G. *et al.* DNA Topoisomerase I differentially modulates R-loops across the human genome. *Genome Biol.* **19**, 1–18 (2018).

247. Lundin, C. *et al.* Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res.* **33**, 3799–3811 (2005).

248. Merrick, C. J., Jackson, D. & Diffley, J. F. X. Visualization of Altered Replication Dynamics after DNA Damage in Human Cells. *J. Biol. Chem.* **279**, 20067–20075 (2004).

249. Minca, E. C. & Kowalski, D. Replication fork stalling by bulky DNA damage: Localization at active origins and checkpoint modulation. *Nucleic Acids Res.* **39**, 2610–2623 (2011).

250. San Martin-Alonso, M., Soler-Oliva, M. E., García-Rubio, M., García-Muse, T. & Aguilera, A. Harmful R-loops are prevented via different cell cycle-specific mechanisms. *Nat. Commun.* **12**, 1–14 (2021).

251. Ho, B., Baryshnikova, A. & Brown, G. W. Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome. *Cell Syst.* **6**, 192-205.e3 (2018).

252. Maestroni, L. *et al.* RPA and Pif1 cooperate to remove G-rich structures at both leading and lagging strand. *Cell Stress* **4**, 48–63 (2020).

253. Fan, J. & Pavletich, N. P. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev.* **26**, 2337–2347 (2012).

254. Yates, L. A. *et al.* A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nat. Commun.* **9**, (2018).

255. Wang, Q. M. *et al.* Human replication protein A induces dynamic changes in single-stranded DNA and RNA structures. *J. Biol. Chem.* **294**, 13915–13927 (2019).

256. Chedin, F., Hartono, S. R., Sanz, L. A. & Vanoosthuyse, V. Best practices for the visualization , mapping , and manipulation of R-loops. *EMBO J.* 1–13 (2021). doi:10.15252/embj.2020106394

257. Gonzalez De Cozar, J. M. *et al.* A second hybrid-binding domain modulates the activity of Drosophila ribonuclease H1. *J. Biochem.* **168**, 515–533 (2020).

258. Bycroft, M., Grünert, S., Murzin, A. G., Proctor, M. & St Johnston, D. NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J.* **14**, 3563–3571 (1995).

259. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: Modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* **8**, 479–490 (2007).

260. Mitchell, S. F., Jain, S., She, M. & Parker, R. Global Analysis of Yeast mRNPs Sarah. *Nat. Struct. Mol. Biol.* **20**, 127–133 (2013).

261. Zunder, R. M., Antczak, A. J., Berger, J. M. & Rine, J. Two surfaces on the histone chaperone Rtt106 mediate histone binding, replication, and silencing. *Proc. Natl. Acad. Sci. U. S. A.* **109**, (2012).

262. Huang, S., Zhou, H., Tarara, J. & Zhang, Z. A novel role for histone chaperones CAF-1 and Rtt106p in heterochromatin silencing. *EMBO J.* **26**, 2274–2283 (2007).

263. Fillingham, J. *et al.* Two-Color Cell Array Screen Reveals Interdependent Roles for Histone Chaperones and a Chromatin Boundary Regulator in Histone Gene Repression. *Mol. Cell* **35**, 340–351 (2009).

264. Nikolov, V. N., Malavia, D. & Kubota, T. SWI/SNF and the histone chaperone Rtt106 drive expression of the Pleiotropic Drug Resistance network genes. *Nat. Commun.* **13**, (2022).

265. De La Torre-Ruiz, M. A., Torres, J., Ariño, J. & Herrero, E. Sit4 is required for proper modulation of the biological functions mediated by Pkc1 and the cell integrity pathway in Saccharomyces cerevisiae. *J. Biol. Chem.* **277**, 33468–33476 (2002).

266. Vilaça, R. *et al.* The ceramide activated protein phosphatase Sit4 impairs sphingolipid dynamics, mitochondrial function and lifespan in a yeast model of Niemann-Pick type C1. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1864**, 79–88 (2018).

267. Winston, F., Dollard, C. & Ricupero-Hovasse, S. L. Construction of a set of convenient saccharomyces cerevisiae strains that are isogenic to S288C. *Yeast* **11**, 53–55 (1995).

268. Winzeler, E. A. *et al.* Functional characterization of the S. cerevisiae genome by gene deletion

and parallel analysis. *Science (80-.).* **285**, 901–906 (1999).

269.   Giaever, G., Chu, A. M., Ni, L. & Johnson, M. Functional profiling of the Saccharomyces cerevisiae genome. *Nature* 387–391 (2002).

270.   Wach, A., Brachat, A., Pöhlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast* **10**, 1793–1808 (1994).

271.   Alberti, S., Gitler, A. D. & Lindquist, S. A suite of Gateway® cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. *Yeast* **24**, 913–919 (2007).

272.   Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).

273.   Morawska, M. & Ulrich, H. D. Identi cation of an NADH-dependent 5-hydroxymethylfurfural-reducing alcohol dehydrogenase in. *Yeast* 191–198 (2008). doi:10.1002/yea

274.   Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).

275.   Foiani, M., Marini, F., Gamba, D., Lucchini, G. & Plevani, P. The B subunit of the DNA polymerase alpha-primase complex in Saccharomyces cerevisiae executes an essential function at the initial stage of DNA replication. *Mol. Cell. Biol.* **14**, 923–933 (1994).

276.   Yadav, T. *et al.* Genetic recombination in Bacillus subtilis: A division of labor between two single-strand DNA-binding proteins. *Nucleic Acids Res.* **40**, 5546–5559 (2012).

7 Acknowledgements

8 Curriculum Vitae