Regulation of human RNase H1 by phosphorylation

Dissertation

Zur Erlangung des Grades

“Doktor der Naturwissenschaften”

(Dr. rer. nat.)

am Fachbereich Biologie

der Johannes Gutenberg-Universität Mainz

Ivan Mikicic

geboren am 04.08.1993 in Zagreb, Kroatien

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

Mainz, 15. Juli 2024

Dekan: Prof. Dr. Eckhard Thines

1. Berichterstatter: Redacted

2. Berichterstatter: Redacted

Tag der mündlichen Prüfung: 21.08.2024

Declaration

I, Ivan Mikicic, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that it has been indicated in the thesis. Throughout the thesis I have received support regarding experimental design by my supervisor. I have also received experimental support from the colleagues in the lab, some of whom I have supervised for their thesis research. Western blot in the Figure 5B, as well as the optimization of the bioinformatics workflows for phosphoproteomics analysis were performed by redacted. Data for the following figures were contributed by redacted: Figure 7C,D,E; Figure 9B,C,D,E,F; Figure 12G. Data for the Figure 8G were provided by redacted. The Western blot experiment for the Figure 12H,I was performed by redacted.

Abstract

R-loops, structures consisting of an RNA:DNA hybrid and a displaced DNA strand, are increasingly investigated as a cell-intrinsic and cancer-relevant source of replication stress (RS). On the other hand, RS stabilizes R-loops in actively replicating and transcribing regions. This suggests potential unexplored mechanisms through which cells mitigate toxic accumulation of RS-borne R-loops.

To investigate the relationship between the replication stress and R-loops, we first stratified RS by severity in HCT116 cells, reasoning that mild RS might have a distinct signaling signature in comparison to moderate or strong RS. Next, we employed multiplexed mass spectrometry-based phosphoproteomics to investigate the regulation of R-loop resolving enzymes upon induction of different levels of RS. Comparisons of aphidicolin- and oncogene-induced RS revealed similarities at the phosphoproteome level, and a common loss of phosphorylation on RNA processing factors. Interestingly, we found that RS reduces phosphorylation of a major R-loop nuclease, RNase H1, at three residues (S74, S86 and S80) adjacent to its hybrid-binding domain.

The kinase search revealed that DYRK1A phosphorylates RNase H1 at S74. Phosphoproteomics analyses revealed that the reduction in DYRK1A signaling is a driver of RS-mediated loss of phosphorylation on RNA processing factors. Finally, we explored functions of RNase H1 phosphorylation. In *in vitro* assays, the triple phospho-mimicking (S3xD) mutant of RNase H1 had decreased binding to RNA:DNA hybrids in comparison to wild-type or phospho-dead (S3xA) mutant, implying that phosphorylation directly impacts binding to hybrids. In cells, the S3xD mutant showed decreased separation between nucleoli and the nucleoplasm in comparison to S3xA. Finally, interactome analyses with the phospho-mutants and after DYRK1A inhibition revealed phosphorylation-based interactors of RNase H1 and implicated GSK3β in the regulation of RNase H1 stability.

In summary, this work presents major advancements in the understanding of DYRK1A signaling, and of the regulation of activity and stability of RNase H1 by phosphorylation. This could impact the use of RNase H1 as a tool in R-loop biology and advance the understanding of DYRK1A function in disease.

Zusammenfassung

R-Loops, Strukturen bestehend aus einem RNA:DNA-Hybrid und einem verdrängten DNA-Strang, werden zunehmend als zellintrinsische und krebsrelevante Quelle von Replikationsstress (RS) untersucht. Andererseits stabilisiert RS die R-Loops in aktiv replizierenden und transkribierenden Regionen. Dies deutet auf potenzielle unerforschte Mechanismen hin, durch die Zellen die toxische Akkumulation von RS-bedingten R-Loops abmildern.

Um die Beziehung zwischen Replikationsstress und R-Loops zu untersuchen, haben wir zunächst RS nach Schweregrad in HCT116-Zellen klassifiziert, mit der Überlegung, dass milder RS im Vergleich zu moderatem oder starkem RS eine distinkte Signalsignatur haben könnte. Anschließend setzten wir multiplexe, massenspektrometrie-basierte Phosphoproteomik ein, um die Regulation von R-Loop-auflösenden Enzymen bei Induktion verschiedener RS-Level zu untersuchen. Vergleiche von Aphidicolin- und Onkogen-induziertem RS zeigten Ähnlichkeiten auf Ebene der Phosphoproteomik und einen gemeinsamen Verlust der Phosphorylierung von Faktoren der RNA-Prozessierung. Interessanterweise stellten wir fest, dass RS die Phosphorylierung einer wichtigen R-Loop-Nuklease, RNase H1, an drei Resten (S74, S86 und S80), welche an die Hybrid-Bindungsdomäne angrenzen, reduziert.

Die Kinase-Suche ergab, dass DYRK1A RNase H1 an S74 phosphoryliert. Phosphoproteomik-Analysen zeigten, dass die Reduktion des DYRK1A-Signalwegs eine Triebkraft des RS-vermittelten Verlustes der Phosphorylierung von Faktoren der RNA-Prozessierung ist. Schließlich untersuchten wir die Funktionen der Phosphorylierung von RNase H1. In *in-vitro*-Assays zeigte die dreifach Phosphorylierung-imitierende (S3xD) Mutante von RNase H1 eine verringerte Bindung an RNA:DNA-Hybride im Vergleich zum Wildtyp oder zur nicht-phosphorylierbaren (S3xA) Mutante. Dies deutet darauf hin, dass die Phosphorylierung die Bindung an Hybride direkt beeinflusst. In Zellen zeigte die S3xD-Mutante eine verringerte Trennung zwischen Nukleoli und Nukleoplasma im Vergleich zur S3xA-Mutante. Schließlich deckten Interaktom-Analysen mit den Phosphorylierungs-Mutanten und nach Hemmung von DYRK1A phosphorylierungsbasierte Interaktoren von RNase H1 auf und implizierten GSK3β in der Regulation der Stabilität von RNase H1.

Zusammenfassend präsentiert diese Arbeit bedeutende Fortschritte im Verständnis des DYRK1A-Signalwegs und der Regulation der Aktivität und Stabilität von RNase H1 durch Phosphorylierung. Dies könnte die Verwendung von RNase H1 als Werkzeug in der R-Loop-Biologie beeinflussen und das Verständnis der DYRK1A-Funktion bei Krankheiten vorantreiben.

Contents

List of figures

List of tables

Chapter 1

Introduction

1.1. Post-translational modifications as the backbone of cellular signalling

While genes are information storage of the cells, proteins are the main workers, engaging with each other and other components of the cell and catalyzing biochemical reactions to achieve DNA replication, transcription, RNA processing and modification, translation of RNA into proteins, movement of organelles and cells, organization of cells and organelles, energy production and storage, cellular export and import, recycling of cell components, and many other activities. Whereas the human body contains only about 20,000 proteins, it needs to utilize them for a much larger number of discrete actions. To achieve this, proteins are chemically modified in a process called post-translational modification, vastly increasing their functionality. More than 400 posttranslational modifications (PTMs) have been reported, affecting all amino acid side chains, as well as protein termini[1].

1.1.1. Protein phosphorylation

The best studied PTM is a small, reversible modification of polar amino acids called phosphorylation. It consists of a nucleophilic substitution, where a hydroxyl group of an amino acid side chain (serine, threonine or tyrosine) is replaced with a phosphate group using adenosine triphosphate (ATP) as the donor. The reaction is catalysed by enzymes called kinases and reversed by phosphatases. Serines are by far the most commonly modified amino acid, accounting for over 50% of all phosphorylation events[2]. At least two thirds of human proteins are known to be phosphorylated[3], and the database PhosphoSitePlus reports around 230 000 unique human phosphorylation (phospho-) sites. Phosphorylation is a highly versatile modification, involved in majority of cellular processes, and affecting protein function in multiple ways. It can change protein localization, activity, stability, induce a conformation change or influence interactions of proteins with other molecules, including other proteins[4].

1.1.1.1. Protein kinases

The human proteome consists of around 550 protein kinases divided into over 50 families by evolutionary similarity[5]. Kinases differ by acceptor amino acid and substrate specificity. Broadly, they can be divided into serine-, threonine-, serine/threonine- and tyrosine-directed kinases. They commonly phosphorylate multiple substrates, with serine/threonine kinases having more than 200, and tyrosine kinases more than 1000 target phospho-sites on average[6]. Substrate specificity is commonly determined by a short patch of amino acids around the phosphorylated residue called a substrate motif or a phospho-motif. Phospho-motif averaging of all known substrates of a particular kinase produces a consensus motif, which helps to identify critical residues important for kinase specificity. Based on the substrate specificity, kinases can be divided into groups such as proline-directed (P at +1 position), basophilic (positively charged amino acids around, usually ahead of, the phospho-site), acidophilic (negatively charged amino acids around the phospho-site), glutamine-directed (Q at +1)[7] etc.

1.1.1.2. Hierarchical phosphorylation

Phospho-sites are often spaced closely together, enabling coordinated phosphorylation by one or more kinases. Such events, where multiple closely spaced sites are phosphorylated in a defined order, are called hierarchical phosphorylation[8]. A well-studied example is the CK2 kinase, which can phosphorylate patches of multiple nearby serines interspersed with negatively charged amino acids, with crucial phosphoserines at +1 and +3[8]. Since CK2 is an acidophilic kinase, the negative charge introduced by the first phosphorylation event favors subsequent nearby modifications.

Multiple kinases can also cooperate in hierarchical phosphorylation. The initial event in such cases is called priming, and is performed by a priming kinase. Subsequent phosphorylation events are catalysed by a processive kinase[9]. This stepwise mechanism requiring multiple actors enables more elaborate regulation of the substrate function. For example, processes that inhibit the processive kinase but enable priming might lead to a different functional outcome (mono-phosphorylated substrate) than processes where both kinases are active (multi-phosphorylated substrate). On the other hand, by inhibiting the priming kinase, any phosphorylation of the substrate is turned off, without requiring regulation of the processive kinase.

Common priming kinases include DYRK1A and DYRK2, while GSK3 is a common processive kinase[9]. CK1 can act as both a priming and a processive kinase. For the purposes of this work, I will introduce DYRK1A as an example of a priming kinase, and GSK3B as a processive kinase, and expand on their functions and biology.

1.1.1.3. DYRK1A

DYRK1A is a member of the Dual-specificity tyrosine phosphorylation-regulated kinase (DYRK) family of serine/threonine kinases within the CMGC (containing CDK, MK, GSK and CDK-like members) kinase superfamily. The defining feature of this family, evident in their name, is their dual specificity – DYRK kinases are activated by autophosphorylation on tyrosine residues, but phosphorylate other substrates on serines and threonines[10]. In case of the human DYRK1A, the autophosphorylation occurs on the conserved tyrosine (Y321) within the activation loop[11]. DYRK1A shares a high domain similarity with other members of the family (DYRK1B, 2, 3 and 4), with DYRK1B being the most similar. Aside from domains shared with DYRK1B, DYRK1A contains two additional C-terminal features, a histidine repeat (H) and a serine/threonine repeat (S/T), with the former being required for its recruitment to RNA polymerase II (RNAPII) condensates[12]. Accordingly, DYRK1A phosphorylates RNAPII at its C-terminal domain (CTD), affecting transcriptional elongation[13]. Unlike the pTEF-b complex, which phosphorylates RNAPII genome-wide[14], DYRK1A acts on RNAPII only at a subset of promoters containing a highly conserved palindromic DNA motif required for DYRK1A recruitment. The target genes transcriptionally regulated by DYRK1A are involved in translation, RNA processing and the cell cycle regulation[13].

In line with having two nuclear localization signals (NLSs), DYRK1A localizes to the nucleus, but is also present in the cytoplasm[15], suggesting a wide range of potential substrates. Aside from RNAPII, DYRK1A phosphorylates a variety of other substrates involved in transcription and signaling (DCAF7, NFAT, STAT3, TRAF3, FOXO1), tyrosine kinase regulation (c-met, SPRY2, EGFR), cell cycle regulation (LIN52, Cyclin D1 and D3, p27, CDC23), alternative splicing (SRSF1, SRSF2, SRSF6 and SF3B1) and DNA damage response (RNF169, SIRT1)[10].

The *DYRK1A* gene is encoded on chromosome 21. Trisomy 21 (also known as the Down syndrome, or DS) is the most common autosomal aneuploidy among newborns[16].

Accordingly, the *DYRK1A* gene is amplified in DS, giving rise to the overexpressed protein, including in the brain[17]. DS is characterized by intellectual impairment, early onset Alzheimer's disease, congenital heart disease, hypotonia, and other developmental problems[18]. Multiple studies point to a causative role of DYRK1A disbalance in neurological and neurodegenerative phenotypes of DS individuals[19,20]. Therefore, inhibition of DYRK1A is considered a promising therapeutic target to improve cognitive deficits in DS[18].

Multiple roles have been described for DYRK1A in the cell cycle maintenance. It is necessary for assembly of the G0-maintaining DREAM complex through phosphorylation of its subunit LIN52, and therefore promotes quiescence and senescence[21]. Interestingly, the DREAM complex reduces somatic DNA repair capacity by repressing DNA repair genes, which can be rescued by inhibition of DYRK1A, with positive impacts on *C. elegans* aging. Furthermore, DYRK1A interacts with and phosphorylates the E3 ubiquitin ligase RNF169, enhancing its ability to recruit 53BP1 to sites of DNA double-strand breaks (DSBs). Additionally, DYRK1A is required for maintaining basal 53BP1 expression[22]. Since 53BP1 favors non-homologous end joining (NHEJ) over homologous recombination (HR), depletion of DYRK1A increases HR and resistance to DNA damage[23]. Altogether, this suggests that depletion or inhibition of DYRK1A has positive effects on genome stability maintenance.

1.1.1.4. GSK3β

The glycogen synthase kinase (GSK3) is an important factor in the insulin and Wnt signalling pathways. It has two isoforms, GSK3α and GSK3β. GSK3β phosphorylates S/T residues which were primed at +4[24]. In some cases, negatively charged amino acids can mimic phosphorylation at +4, allowing for processive GSK3β activity[25]. GSK3β activity can be primed by different kinases: CK1, CDK5, PKA, ERK2, CK2 and the DYRK family[26] (see 1.1.1.5). There are multiple instances of sequential phosphorylation by GSK3β, at a series of serines or phospho-serine mimics separated by 4 residues, leading to hierarchical multisite phosphorylation[27].

While functional consequences of GSK3β are varied, I will single out the role of GSK3β in protein degradation. Commonly, GSK3β introduces the so-called phospho-degron into substrates, allowing for ubiquitin-assisted proteasomal degradation[26]. There are at least

29 targets the stability of which is regulated by GSK3β and the proteasome[26]. Typically, the GSK3β-induced phospho-degron recruits the cullin-RING family E3 ubiquitin ligase complex called SCF[28]. Upon K48-linked ubiquitylation of the substrate, it is directed into proteasomal degradation. Among the best studied GSK3β phospho-degron substrates are MYC (T58)[29] and JUN (S251)[30].

1.1.1.5. Cooperation between DYRK1A and GSK3β

A well-characterized example of the DYRK1A-GSK3β cooperation is the microtubule growth protein Tau. DYRK1A primes Tau, and processive phosphorylation is catalysed by GSK3β[31]. Overproduction of DYRK1A in DS results in more priming and increased hyperphosphorylation and aggregation of Tau, which is seen as a causative factor in early onset Alzheimer's disease, a common characteristic of DS[19]. Interestingly, the same study reports that phosphorylation of a translation initiation factor eIF2Bε is primed by DYRK1A and continued by GSK3β[31]. These two kinases also cooperate in phosphorylation of another microtubule growth factor, MAP1B. MAP1B regulates microtubule dynamics at growing axons and growth cones, and DYRK1A depletion compromises neuritogenesis and microtubule dynamics in cortical neurons, possibly by alterations in MAP1B phosphorylation[32].

Interestingly, DS individuals have a reduced incidence of solid tumors[33]. This is likely a consequence of changes in phosphorylation of the transcription factor called nuclear factor of T cells (NFAT). DYRK1A primes NFAT in the nucleus for processive phosphorylation by GSK3β and CK1. Fully phosphorylated NFAT is exported from the nucleus to the cytoplasm, repressing the tumor angiogenesis transcriptional program[34].

1.1.2. Protein ubiquitylation

Aside from small chemical groups (methyl, acetyl, phosphate) serving as PTMs, whole proteins can also be covalently attached to existing proteins. One of the best known examples is the 76 amino acid long, 8.6 kDa protein called ubiquitin, as well as a family of structurally related proteins, termed ubiquitin-like modifiers (including SUMO, NEDD8, UFM1, and others). Ubiquitin typically attaches to the ε-amino group of substrate lysines using its C-terminal glycine. Aside from lysines, other amino acids can also get ubiquitylated, most prominently the N-terminal methionine[35]. A defining feature of ubiquitin

signaling is its modularity: ubiquitin not only attaches itself to a substrate, but can also get ubiquitylated itself, giving rise to polyubiquitin chains. Moreover, ubiquitin has multiple amino acids available for modification, enabling distinct topologies. Sequential ubiquitylation using one of the eight ubiquitin residues (M1, K6, K11, K27, K29, K33, K48 and K63) produces polyubiquitin chains of a particular linkage type (also called homotypic chains). Mixed (heterotypic) chains arise if a single chain contains at least two different linkage types. If a single ubiquitin within a chain gets modified on two distinct residues (eg. K63 and K48), it leads to chain branching. An additional consideration carrying potential functional information is the length of chains and branches. Considering the number of different branch types, as well as the fact that a substrate can get mono-, homo- or heterotypically poly- and branched polyubiquitylated at multiple acceptor amino acids, this enables a complex type of modification determining distinct functional outcomes, a phenomenon called the ubiquitin code.

1.1.2.1. Enzymatic machinery for ubiquitylation

Ubiquitin is attached to target proteins through a series of enzymatic reactions involving ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligases (E3). E3 ligases carry out the final step of this process, covalent transfer of ubiquitin to a substrate, and are crucial for substrate selectivity[36]. The versatile nature of ubiquitin signaling is primarily realized through a large number of E3 enzymes, more than 600 of which have been described in humans. They are classified by domain identity into RING/U-box, HECT and RING-between-RING (RBR) E3 ligases. While HECT and RBR domains have intrinsic catalytic activity, RING and U-box domains function by facilitating the transfer of ubiquitin from the E2 enzyme to the substrate. Therefore, in contrast to HECT and RBR ligases, which themselves determine the linkage type, it is determined by the E2 enzymes in case of RING ligases. Removal of ubiquitin from substrates and disassembly of ubiquitin chains and branches are catalysed by specialized proteases called deubiquitinases (DUBs). Around 100 DUBs have been described in humans, and they are classified by the mode of catalysis into cysteine-based and zinc-binding metalloproteases. Based on selectivity, they can be classified into non-selective, chain type-selective, and chain type- and chain length type-selective DUBs[37]. A defining feature of DUBs is the presence of various ubiquitin-binding interfaces and domains, which help

to recognize and position ubiquitin chains for hydrolysis, enabling particular selectivity characteristics of each enzyme[37].

1.1.2.2. The ubiquitin-proteasome system (UPS)

Similarly to phosphorylation, ubiquitylation can serve a variety of functions, from changing a protein conformation, activity, interactions and localization to, most prominently, regulating protein stability. Canonically, degradation of proteins is induced by modification using the K48-linked ubiquitin chains. Recently, branched ubiquitin chains, typically containing K48 linkages, are being recognized as a widespread and high-efficiency degradation signal[38]. Degradation of ubiquitin-modified proteins is carried out by a large multi-subunit complex called the proteasome. It consists of two subcomplexes: the 20S core particle (CP) and the 19S regulatory particle (RP)[39]. The CP has a cylindrical structure composed of four stacked rings that facilitate entry of substrates into the cylinder and their proteolytic degradation. The inner (β) rings of the CP contain the active sites with trypsin-like, chymotrypsin-like and caspase-like proteolytic activities. The RP is composed of a base and a lid, and caps one or both ends of the CP. It contains receptors recognizing 48-linked polyubiquitin chains, acting as the initial bridge between the substrate and the proteasome. Additionally, it contains AAA+ ATPase and other enzymatic activities necessary for substrate unfolding and translocation into the CP.

1.2. Replication stress

1.2.1. DNA replication

The coordinated steps each somatic cell undertakes before and during cell division is called the cell cycle, and can be divided into four distinct phases: G1, S, G2 and M. A defining feature of the cell cycle is its ability to precisely duplicate genetic information and equally divide it between the daughter cells during division. Genetic information is duplicated during DNA replication, taking place in S phase, with G1 phase serving to prepare cells for DNA replication. DNA origins, or origins of replication, are genomic loci from which DNA replication is initiated. In G1, they are bound by the Origin Recognition Complex (ORC), followed by recruitment of proteins CDC6 and CDT1, which are essential for loading of the double-hexameric helicase complex MCM2-7. The pre-Replication Complex (pre-RC) is formed upon loading of the inactive MCM2-7 complex, followed by phosphorylation-mediated activation driven by S-phase Cyclin-Dependent Kinases (S-

CDKs) and the Dbf4-dependent kinase (DDK, containing the catalytic subunit CDC7)[40]. Upon phosphorylation, additional components are recruited, including CDC45 and the GINS complex. This gives rise to the active replicative helicase termed the CMG (CDC45-MCM-GINS), and enables firing of the replication origins, a process in which the CMG helicase starts bidirectionally unwinding DNA molecules, enabling the replication machinery access to the DNA. Importantly, only a small subset of licensed replication origins are actually fired, the others remaining dormant and available for activation during S phase if required.

As DNA replication cannot be primed *ab novo* with DNA primers, it is initiated by the primase, a DNA-dependent RNA polymerase, which produces short RNA primers. Acting together with the primase, DNA polymerase α then extends the RNA primers with short stretches of DNA. To enable continuous replication, a more processive complex is loaded, containing DNA polymerases ε and δ. Both strands of the DNA are duplicated simultaneously, but considering that replication takes place in the 5'-3' direction, synthesis of one strand (the leading strand) is continuous, whereas synthesis of the complementary strand (the lagging strand) is more complex and involves synthesis of stretches termed the Okazaki fragments (OFs). The OF synthesis requires repeated action of polymerase α-primase and polymerase δ. Active DNA synthesis takes place at the Y-shaped DNA region where the double-stranded parental DNA meets the nascent (newly replicated) DNA, called the replication fork (RF).

The multi-subunit complex containing DNA polymerases and other replication factors enabling replication of both the leading (polymerase ε) and the lagging strand (polymerase δ) is called the replisome. In addition to DNA polymerases, the replisome includes the processivity component called the sliding clamp (PCNA), clamp loader (RFC), the CMG helicase, the single-stranded binding factor Replication Protein A (RPA), topoisomerases relieving torsional stress, DNA ligases involved in sealing DNA nicks and factors involved in RNA primer removal and OF processing (RNase H, FEN1). As replication progresses, replisomes fired in opposite directions from adjacent origins move closer to each other. Once they converge, the CMG helicase is unloaded and disassembled, and DNA polymerases and released. To remove discontinuities in the newly replicated DNA, DNA overhangs (flaps) are removed by FEN1, and RNA primers are degraded by RNase H,

followed by gap filling and nick ligation. The replisome is highly flexible, and different components are recruited through protein-protein interactions (PPIs) depending on requirements of the particular replication step. Multiple mechanisms are employed to ensure fidelity and processivity of replication, as well as coordination between the leading and lagging strand synthesis.

1.2.2. Definition and sources of replication stress

Any stalling or slowing down of the replication fork is defined as the replication stress (RS)[41]. It derives from both endogenous (cellular or organismal) and exogenous (environmental) sources. Structural perturbations in DNA resulting from damaged DNA bases, mis-incorporated ribonucleotides, repetitive DNA sequences, secondary DNA structures such as G-quadruplexes or R-loops (see section 1.3) and transcription-replication conflicts (TRCs)[42] can impede with the progress of the fork and therefore lead to RS. Procedural sources involve depletion of deoxyribonucleotides (dNTPs) and an increase in origin firing rates[43,44]. The most common environmental sources of RS are UV irradiation, products of cellular metabolism such as reactive oxygen species (ROS) or aldehydes, and chemical mutagens and cytostatics. UV irradiation leads to formation of bulky DNA lesions such as thymidine dimers and 6,4-photoproducts, ROS induces oxidation of DNA bases, aldehydes trigger DNA inter-strand crosslinks and DNA-protein crosslinks, while other chemical mutagens cause a variety of chemical modifications of DNA. Hydroxyurea (HU) and aphidicolin (APH) are the most common RS-inducing cytostatics. Mechanistically, HU inhibits the ribonucleotide reductase, thereby depleting the dNTP pool. The tetracyclic diterpenoid APH is a natural product isolated from fungi such as *Cephalosporum aphidicola* and *Nigrospora oryzae* that directly inhibits replicative B-family DNA polymerases, including polymerases α, δ and ε[45]. HU and APH have been vital in studying the cellular response to RS, as they enable precise manipulation of treatment duration and dosage. In addition, owing to its anti-proliferative properties, HU has been used in the clinic for the treatment of clonal myeloproliferative disorders and solid tumors, although the treatment is associated with secondary leukemias as a result of mutagenicity and induction of chromosome instability[46]. Finally, replication stress is often induced by overexpression of oncogenes such as MYC, RAS or Cyclin E. In case of MYC and Cyclin E, RS is induced by premature S-phase entry while the G1 transcriptional

program is still taking place, leading to replication-transcription conflicts in highly transcribed genes[47].

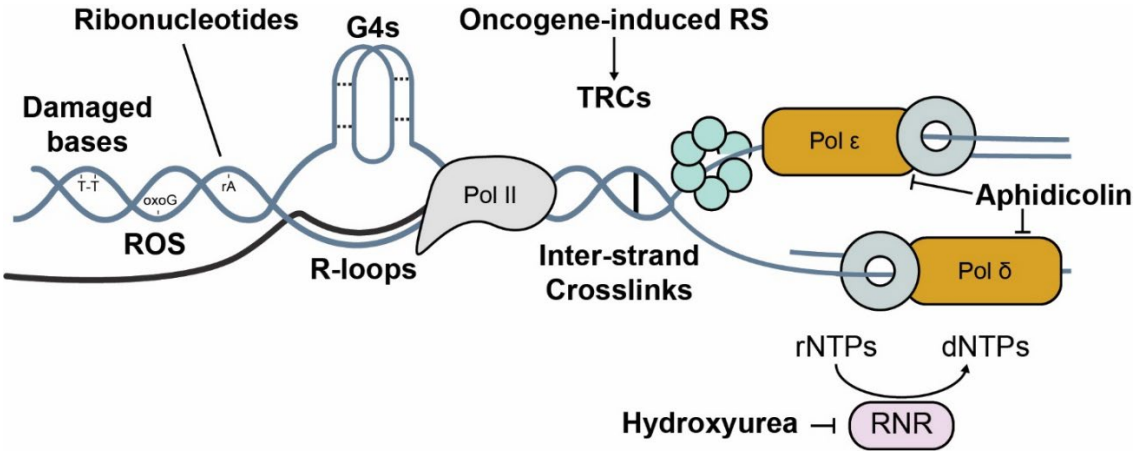


Figure 1: Sources of replication stress.

A number of different sources cause replication stress. This includes: damaged DNA bases; reactive oxygen species (ROS); mis-incorporated ribonucleotides; secondary DNA structures such as R-loops and G4s; inter-strand crosslinks; depletion of dNTPs by hydroxyurea; inhibition of DNA polymerases by aphidicolin; or transcription-replication conflicts, caused by elevated oncogene levels. RNR = ribonucleotide reductase.

1.2.3. Replication stress signaling

A defining attribute of RS is the fork-adjacent exposure of single-stranded DNA (ssDNA) as a result of decoupling of the stalled DNA polymerases with the CMG helicase[48]. Being more sensitive to DNA damage and nuclease attacks[49], such stretches are bound and protected by the replisome component RPA. In fact, RPA is essential for protection of stalled forks, as extensive exposure of single-stranded DNA leads to depletion of the RPA pool and fork breakage[50]. Stretches of RPA-coated ssDNA and ssDNA-dsDNA junctions recruit the Ataxia telangiectasia and Rad3 related (ATR) kinase, the canonical RS response kinase. ATR is activated through interaction with its partner ATRIP, as well as with either of two proteins containing ATR-activating domains[51], TOPBP1[52] and ETAA1[53,54]. ATR belongs to a family of serine/threonine directed Phosphatidylinositol 3-kinase-related kinases (PIKKs), and as such targets proteins at the S/T-Q motif[55]. It targets a multitude of described nuclear substrates, with the most well-known being the checkpoint kinase CHK1. CHK1 helps to spread the ATR-driven phosphorylation signals away from replication forks and ssDNA, which leads to several functional consequences. CHK1 phosphorylates and inhibits the CDC25A and CDC25C phosphatases, which

normally remove the inhibitory phosphorylations on CDKs catalyzed by WEE1 and MYT1 kinases. By preventing CDC25-mediated dephosphorylation and activation of CDKs like CDK1, CHK1 activity stabilizes the inhibitory phosphorylation state of CDKs, thereby inducing the cell cycle arrest. Aside from this, the ATR-CHK1 signaling also suppresses origin firing, stabilizes the replication forks and promotes fork repair and restart[56].

A particularly perplexing relationship is that of RS and origin firing. While the ATR-CHK1 pathway inhibits firing of new origins by inhibiting DDK-CDC7[57] and CDK1[58], RS has also been shown to increase origin firing. Along these lines, the rate of origin firing and replication speed seem to be inversely correlated, meaning that dormant origins are induced after replication stress, and a fork slowdown ensues upon dormant origin firing[44]. One reason for RS-induced dormant origin firing is to allow completion of replication in a timely manner despite reduced fork speed[59–61]. One proposed explanation for origin firing both increasing and decreasing upon RS is that dormant origins in actively replicating DNA regions are fired upon RS to rescue adjacent stalled forks, while dormant origins in non-replicating regions are suppressed[62]. Another aspect might be dosage dependency. For instance, the ATR-CHK1 pathway is activated to a different extent depending on the amount of exposed single-stranded DNA. As such, low doses of replication stress might be more permissive for dormant origin firing. In line with this notion, low levels of replication stress escape the checkpoint, meaning that ATR-CHK1 pathway is not fully activated, and are able to complete replication despite decreased fork speed[63].

1.2.4. Replication fork remodeling and repair pathways elicited by replication stress

The replication forks are highly plastic and amenable to remodeling upon RS[64]. This includes changes in the composition of the replisome and rearrangements of nucleic acid structures at and around the forks. Remodeling serves to stabilize forks and enable their repair, but can also be a source of DNA damage and genome instability if the machinery required for fork protection is mutated or missing. Major pathways of fork remodeling and repair include fork reversal, translesion synthesis, template switching and repriming[65].

During fork reversal, the forks are turned into four-way (Holliday) junctions by the activity of non-redundant DNA translocases HLTF, ZRANB3, SMARCAL1[66] and PICH[67]. Multiple proposed scenarios aim to explain the role of fork reversal. If DNA damage or mismatched

nucleotides are incorporated into DNA, fork reversal can serve to hybridize the damaged DNA with the parental DNA strand, thus enabling repair using template information. On the other hand, Holliday junctions are templates for structure-specific DNA nucleases, which can turn them into double-strand breaks (DSBs), allowing for rescue of stalled forks by break-induced replication (BIR)[68]. Interestingly, fork reversal can also regulate the rate of fork progression. This is evidenced by observations that replication stress not only causes reversal of stalled, but also of ongoing forks[69], suggesting it might act as a genome-wide means to adjust fork speed, possibly to allow extra time for repair or avoid encounters of replication machinery with damaged DNA.

The ends of the reversed forks need to be protected by specialized protein machineries to avoid nuclease-mediated resection. Major fork protectors include the tumor suppressor proteins BRCA1 and BRCA2[70], although plenty more have been described. If these factors are mutated or missing, they allow nascent DNA degradation by nucleases such as MRE11, EXO1[68] and DNA2[71]. Nascent DNA degradation has been proposed as one of the main vulnerabilities of BRCA-mutated tumors, allowing for design of synthetic lethal therapies[72].

Translesion synthesis (TLS) utilizes specialized DNA polymerases with relaxed nucleotide pairing requirements, such as Polη, Polκ, Polι, Polζ and REV1. It thus allows DNA synthesis over damaged DNA, which is however often mutagenic, as the correct template is not available[73]. Template switching (TS) denotes a pathway in which the synthesis of nascent DNA is continued on the sister chromatid upon encountering a fork block, such as a bulky adduct[74]. The process entails homologous recombination (HR), requiring the RAD51 filament-based homology search, DNA synthesis from the sister chromatid template, branch migration and resolution of Holliday junctions.

Finally, repriming refers to initiation of leading-strand replication immediately past the stalled fork using the specialized DNA polymerase PRIMPOL[75]. Repriming results in daughter-strand gaps, regions of ssDNA flanked by dsDNA. Daughter-strand gaps often contain lesions that caused fork stalling in the first place, and that need to be repaired post-replicatively using TLS or TS[76]. While this pathway allows for continued replication past the lesion without requiring prior repair, daughter-strand gaps are also structures

vulnerable to nuclease-mediated expansion and DSB formation, and have been identified as the underlying cause of synthetic lethality between PARP1 inhibitors and BRCA1/2 mutations[77,78]. In addition, extensive use of repriming increases mutagenicity, as it can employ TLS for post-replicative repair. Importantly, different fork remodeling and repair pathways are jointly regulated to ensure the choice of the appropriate pathway based on the identity of the underlying lesion, duration of fork stalling and the availability of the necessary pathway-specific proteins[79].

1.2.5. Replication stress as a source of genome instability

One of the key features of RS is the increase in genome instability (GI), evident by the recurrent changes to the genetic material. A subset of GI is termed chromosome instability or CIN, and can be further separated into numerical/whole-chromosome CIN (W-CIN) and structural CIN (S-CIN). W-CIN refers to gains or losses of whole chromosomes, while S-CIN refers to changes in chromosome sequences as a result of indels, substitutions and translocations[80]. Interestingly, RS can induce both W-CIN and S-CIN, although likely though distinct mechanisms[81]. While W-CIN is thought to occur as a result of chromosome mis-segregation through processes such as increased microtubule instability[82] or defects in centrosome integrity[83], S-CIN is more associated with mutagenic DNA repair processes, such as TLS, repriming[84], mitotic DNA synthesis (MiDAS) and resolution of DNA bridges[85]. S-CIN is particularly common at specific genomic locations called the common fragile sites (CFSs), which were first described cytogenetically as gaps, breaks or constrictions on metaphase chromosomes[86]. CFSs are a result of under-replication at difficult-to-replicate regions, such as secondary DNA structures, tandem repeats or origin-poor regions with long genes transcribed throughout interphase. They are found in all individuals in cells subjected to mild RS. CFSs are causal for neurodegenerative syndromes such as the Fragile X syndrome, and are recognized as a driver of GI in cancer cells[87].

RS and CIN are tightly linked with tumors, and are thought to promote both tumor induction and evolution[88]. While sources of RS in pre-neoplastic lesions and tumors are often difficult to clarify, increased expression of some of the most common oncogenes, such as MYC, RAS and Cyclin E, definitely plays a role[89]. Mutations in tumor suppressors, such as BRCA1 and BRCA2, are also associated with inability of cells to repair RS-associated

lesions[65] and consequent CIN induction. Increasingly, R-loops and G-quadruplexes (G4s) are recognized as cancer-relevant sources of RS, for example in hematological tumors such as acute myeloid leukemia[90]. While we often focus on disease-relevant negative consequences of RS, it is worth asking whether cells “permit” low levels of RS and mutagenicity because they bring a potential advantage. From an evolutionary perspective, RS could be an important driver of genome evolution through events such as gene duplication[91].

1.3. R-loops

1.3.1. Transcription

Aside from serving as a storage of genetic information, DNA fulfills its role by allowing transcription of genes into different classes of RNAs, some of which are translated into proteins. Transcription is carried out by three classes of DNA-dependent RNA polymerases – RNA Pol I, Pol II and Pol III. The majority of protein-coding genes are transcribed into pre-mRNA using Pol II, whereas pre-rRNA genes are transcribed in the nucleoli using Pol I, and tRNAs and other classes of non-coding RNAs are transcribed by Pol III[92]. Pol II-mediated transcription can be divided into initiation, promoter clearance, elongation and termination. During initiation, the pre-initiation complex (PIC) composed of RNA Pol II and various transcription factors is assembled at gene promoters, ~30 bp upstream of the transcription start site (TSS), by binding to defined DNA sequences, such as the TATA box[93]. PIC assembly and interactions with gene enhancers are modulated by a multi-subunit complex called the mediator[94]. As part of the PIC, transcription factor TFIIH, which includes the catalytic subunit CDK7, hyper-phosphorylates the C-terminal domain (CTD) of the major Pol II subunit RPB1 at Ser 5 and Ser 7[95], facilitating transition from initiation to promoter-proximal pausing by interrupting interactions with the mediator[96]. Pausing serves as a regulatory step to allow quick changes in gene expression in response to stimuli[97]. It is mediated by the Negative Elongation Factor (NELF) complex, and eventually Pol II is released into elongation through phosphorylation of NELF and hyper-phosphorylation of the CTD of RPB1 at Ser 2 triggered by the Positive Transcription Elongation Factor (p-TEFb), including the catalytic subunit CDK9[98]. Termination is accompanied by 3' end cleavage and polyadenylation (polyA tailing), which not only enhances termination, but also regulates transcript stability, export to cytoplasm

and translation[99]. During elongation, previously synthesized segments of pre-mRNA are already bound by machineries facilitating splicing, modification (for example, m6A), secondary structure formation and export (TREX-THO complex), with steps of these processes happening either co-transcriptionally or post-transcriptionally.

1.3.2. Definition and sources of R-loops and RNA-DNA hybrids

R-loops are three-stranded nucleic acid structures consisting of an RNA-DNA hybrid and a displaced DNA strand. They are usually formed co-transcriptionally, due to nascent RNA hybridizing with the complementary template DNA[100]. This is favored by negative supercoiling behind the transcribing RNA polymerases and is thought to occur stochastically. Since the nascent RNA hybridizes with the locus from which it has been synthesized, this type of an RNA-DNA hybrid is called *in cis*. The level of transcription at a gene locus strongly positively correlates with the amount of R-loops[101]. Transcribed GC-rich regions are particularly prone to R-loop formation because the displaced DNA strand can form G-quadruplexes (G4s), which stabilizes RNA-DNA hybrid opposite the G4[102]. Additionally, Pol II pausing also favors R-loop formation, as it increases the chance of nascent RNA hybridizing with complementary DNA[103]. Within transcribed genomic regions, R-loops are most common at promoters, transcription start sites and termination sites[104], where they serve important functions. Although we usually consider Pol II to be a major source of R-loops, Pol I and Pol III transcriptional activity also results in R-loop formation[105], in part because of their high transcription rates.

Direct visualization of R-loops using native electron microscopy revealed that genomic R-loops range in size from 30 to 3000 base pairs (bp), with the average size around 180 bp[106]. The same study estimated that a maximum of 0.009% of the genome or around 290 kb is occupied by R-loops, which is much more conservative than the genomic approaches reporting up to 5% of the R-loop-covered genome[101]. This discrepancy in estimates likely stems from the different methodologies used. In any case, R-loops are quite a prevalent secondary DNA structure and, interestingly, do not conform to either the B-form of dsDNA, nor the A-form of dsRNA, instead forming a mixed A/B type structure[107]. Co-transcriptional R-loops have an estimated half-life of 10-20 minutes[101]. This suggests that R-loops are turned over within minutes and their half-life coincides with that of the

promoter-proximal-paused Pol II[108], suggesting that their existence is compatible with dynamics or a regular transcription cycle.

RNA-DNA hybrids *per se* do not necessarily need to form at dsDNA. In other words, ssDNA is also a common template for hybrid formation. For instance, one of the most common genomic sources of hybrids is the replication cycle, in particular Okazaki fragment formation, which includes RNA-primed DNA templates[109]. Moreover, mis-incorporation of ribonucleotide triphosphates (rNTPs), most commonly by polymerase ε, results in “mini-hybrids”[104], and is one of the most common sources of genomic DNA damage[110]. Finally, multiple sources report RNA-DNA hybrids at DSBs[111]. Resection was found to be required for hybrid formation, suggesting the existence of DNA damage-induced long non-coding RNAs (dilncRNAs) paired with resected DNA ends, forming a hybrid rather than an R-loop, and influencing repair of damaged DNA[112].

1.3.3. Physiological roles of R-loops

1.3.3.1. R-loops at repetitive DNA sequences

R-loops are not mere byproducts of transcription, but take on important functions. They are enriched at repetitive genomic regions, such as telomeres, centromeres, retrotransposons and ribosomal DNA (rDNA)[104]. At chromosome ends called telomeres, which are in vertebrates characterized by tandem repeats of the 5’-TTAGGG-3’ sequence, R-loops are formed by a class of long lon-coding RNAs (lncRNAs) called telomere repeat-containing RNAs (TERRA). TERRA hybridizes with the C-rich telomeric DNA strand, which is likely potentiated by G4 formation at the displaced G-rich strand[113]. Similarly to hybrids formed at DSBs, TERRA drives homologous recombination in a telomere maintenance pathway termed alternative lengthening of telomeres (ALT)[114]. Centromeric loci also express lncRNAs that hybridize with centromeric DNA *in cis.* Centromeric R-loops are particularly enriched in mitosis and enable accurate chromosome segregation through an ATR- and Aurora B-driven pathway[115]. Mechanistically, the displaced DNA strand of centromeric R-loops recruits RPA and subsequently activates ATR in an Aurora A-dependent manner. ATR in turn facilitates Aurora B activation, which suppresses lagging chromosomes.

Detection of R-loops by immunofluorescence using multiple strategies reveals that nucleoli, regions where Pol I-mediated transcription of pre-rRNA from rDNA repeats takes

place, are nuclear regions most enriched in R-loops[116]. The rDNA loci are interspersed by intergenic spacers (IGSs), with both regions being prone to R-loop formation. At rDNA promoters, characterized by G-rich CpG islands, R-loop formation prevents DNA methylation and locus silencing, thus supporting production of rDNA[117]. This is a good example of positive roles of promoter-bound R-loops on transcription, which has also been observed across other genomic CpG islands[117]. On the other hand, at IGSs, which are transcribed in the antisense direction by Pol II, persistent R-loops form a shield preventing Pol I-mediated sense transcription. Thus, through restricting Pol I activity to rDNA only, R-loops at IGSs ensure appropriate nucleolar morphology and ribosome biogenesis[118].

1.3.3.2. R-loops at transcription termination sites

Aside from roles at promoters, R-loops are also abundant and functionally relevant at termination sites[101]. Faulty termination can lead to transcription past the TTS, a phenomenon called read-through transcription. Read-through is particularly dangerous in gene-dense regions, as it can affect downstream gene expression by transcriptional interference[119]. A common strategy of termination is RNA polymerase stalling at a G-rich sequence downstream of the polyA sequence[120]. A genome-wide mapping analysis found that around 2000 G-rich termination sites are occupied by R-loops[121]. Although the exact mechanisms by which R-loops aid termination are still discussed, they include triggering Pol II backtracking, Pol II arrest due to torsional stress and recruitment of termination-promoting factors, such as BRCA1, SETX, DDX5, DHX9 and XRN2[104].

1.3.3.3. R-loops in antisense transcription

LncRNAs generated by antisense transcription can form R-loops and often recruit transcription activators or repressors. For instance, lncRNA *ANRASSF1* is produced from the 3'-end of the *RASSF1* locus in the antisense direction and recruits the Polycomb repressive complex 2 (PRC2) to silence the *RASSF1*A promoter through H3K27 trimethylation[122]. An example of an R-loop acting *in trans* and resulting in deposition of activating histone marks is the *GATA3-AS1* lncRNA. It is expressed from the *GATA3* locus and forms an R-loop 3 kb upstream. It upregulates GATA3 expression by recruiting MLL/KMT2A to the locus, resulting in deposition of an activating mark – H3K4 timethylation[123]. Additionally, R-loops can promote DNA demethylation at CpG islands through antisense RNAs. The lncRNA *TARID*, produced in the antisense direction from

the *TCF21* locus, forms a promoter R-loop that recruits the direct R-loop binder GADD45A. Through direct protein-protein interactions, GADD45A recruits the DNA demethylase TET, which removes DNMT3-mediated CpG promoter methylation, thereby enabling expression from the *TFC21* locus[124].

1.3.3.4. R-loops and RNA-DNA hybrids in DNA repair

Multiple labs have detected RNA-DNA hybrids at DSBs and have shown they have direct roles in DSB repair. It is, however, less clear whether three-stranded R-loop structures are also formed at DSBs[104]. Multiple non-mutually exclusive models of hybrid production at breaks were proposed. Increased pausing of Pol II, a decrease in elongation rates and premature transcriptional termination at DSBs could result in co-transcriptional stabilization of hybrids[111]. Along these lines, genome-wide mapping of RNA-DNA hybrids following induction of DSBs revealed their accumulation primarily in regions occupied by Pol II before break induction, arguing for a model in which pre-existing RNAs hybridize with DNA at breaks[125,126]. On the other hand, *de novo* transcription from 3'-ends of resected DSBs is evidenced by recruitment of the PIC, the mediator complex, CDK9 and Pol II to DSBs[127]. This results in production of dilncRNAs[112] that favor repair by homologous recombination.

Regardless of the source of RNA-DNA hybrids at DSBs, they can contribute to repair in multiple ways. RNA-DNA hybrids were shown to both prevent[128] and favor[129] nuclease-mediated resection at DSBs. Downstream of resection, hybrids seem to favor RAD51 nucleofilament formation[130,131], a step essential for strand invasion part of the HR[132]. However, RNA-DNA hybrids need to be removed prior to RAD51 loading[125,129], highlighting their transient nature and tight regulation at DSBs. Aside from classical views of HR, alternative pathways of DSB repair are emerging, such as the RNA-templated repair. Such repair is based on the following premises: (1) mRNA produced from the DSB locus before the DSB induction can span both sides of the DSB, providing a possible template for repair[133]; (2) the translesion synthesis polymerase Polζ is an RNA-templated DNA polymerase[134]; (3) RAD52 and RPA favor annealing of mRNA with complementary DNA at DSBs. Therefore, DNA synthesis from the minimally resected 3'-end of the DSB can be initiated by Polζ in a RAD52-dependent manner, using hybridized RNA as the

template[134]. This type of repair is particularly suitable for G0/G1 phase, when sister chromatids are not available for classical HR-mediated repair[135].

1.3.4. Role in genome instability

So far, I have highlighted functional roles of R-loops. However, a recurring theme in R-loop biology is the transient nature of R-loops and RNA-DNA hybrids, and the need for their regulation and removal. Accordingly, R-loops can impede major DNA- and RNA-related processes that happen on chromatin, such as DNA replication and repair, and transcription[136]. In cases when R-loop levels are elevated and associated with negative functional outcomes, we consider R-loops aberrant, unscheduled, or even pathological[136]. Many of the negative effects attributed to R-loops stem from the fact that they increase replication stress and replication-associated DNA damage[137]. In addition, even in non-dividing cells, where replication-associated processes should not play a major role, R-loops can cause transcription-associated DNA damage[138].

It is widely recognized that ssDNA is more vulnerable than dsDNA to DNA damage[139]. Stabilization of R-loops can thus increase DNA damage by leaving the displaced ssDNA exposed, rendering it more prone to spontaneous chemical modifications, and more accessible to chemical mutagens, oxidative damage, base-modifying enzymes or nucleases[137]. Aside from that, conflicts between the replication and transcription machineries have lately garnered a lot of attention[140]. As both machineries share the same template, they are normally temporally and spatially well separated. Circumstances like oncogene-induced premature S-phase entry and dormant origin firing increase the likelihood of these conflicts[47]. Mechanistically, this is thought to occur through replication taking place on genes that are still actively transcribing. There are two main possible scenarios: the replication and transcription machineries travelling in the opposite direction towards each other (head-on collisions), and both machineries traveling in the same direction (co-directional conflicts). Head-on collisions were shown to be much more dangerous, as they lead to R-loop accumulation, activation of the ATR pathway, and an increase in DSBs[141]. On the other hand, co-directional conflicts activate ATM, and decrease R-loop levels, possibly by the action of the replicative helicase MCM2-7, which can unwind RNA-DNA hybrids[142]. Alternatively, PrimPol could reprime replication

downstream of the R-loop in case of co-directional transcription-replication conflicts (TRCs), as is the case at the GAA repeats[143].

Importantly, R-loops are necessary for induction of DSBs upon head-on TRCs. Activation of ATR by R-loops requires fork reversal and the structure-specific endonuclease MUS81[144], suggesting that forks reverse upon the TRC, and are cleaved by MUS81 to produce a single-ended DSB. Another pathway connecting R-loops to DSB induction is the transcription-coupled nuclear excision repair (TC-NER). R-loops can be processed by the TC-NER endonucleases XPG and XPF, which were shown to create adjacent single-strand breaks, helping to excise the R-loop from the genome, and thereby creating a DSB[145]. Moreover, a recent surprising report found that XPG/XPF-processed nuclear R-loops are a source of cytoplasmic RNA-DNA hybrids, suggesting that they are exported or diffuse from the nucleus to the cytoplasm. In line with cytoplasmic nucleic acid species being immunogenic, these hybrids activate the cGAS-STING pathway, leading to cell death through the innate immune response[146], a mechanism that likely contributes to R-loop-related diseases.

1.3.5. R-loop removing enzymes

Considering the deleterious potential of R-loops and RNA-DNA hybrids, it is not surprising that they are tightly regulated by multiple specialized machineries. Main enzyme classes involved in suppression of RNA-DNA hybrids are helicases, nucleases, topoisomerases and translocases. Many additional factors that do not possess these enzymatic activities can also suppress R-loops, as will be described below.

1.3.5.1. R-loop helicases

Among helicases, a large family called the DEAD-box (DDX) and the DEAH-box (DHX) contains several members known to unwind R-loops. For instance, DHX9 binds RNA-DNA hybrids and unwinds R-loops[147,148]. Along with a couple of other DDX helicases, it associates with the PCNA unloader ATAD5 to suppress R-loops during replication[149]. DDX5[128], DDX19A[150], DDX21[151] and DDX39B[152] were also implicated in the R-loop resolution. Our group has demonstrated that another DEAD-box helicase, DDX41, suppresses aberrant R-loops and consequent DSBs at gene promoters. Depletion of DDX41 leads to increased replication stress, dependency of cancer cells on ATR and inflammatory signaling. Considering that DDX41 mutations are implicated in acute

myeloid leukemia (AML), we proposed that R-loop accumulation might be involved in the etiology of AML[153]. Accordingly, using zebrafish models, Weinreb *et al.* found that Ddx41 regulates the number of hematopoietic stem progenitor cells (HSPCs) and constrains R-loop levels. Moreover, Ddx41 insufficiency led to increased cGAS-STING inflammatory signaling, which is associated with hyperproliferation of HSPCs[154]. These findings substantiate previously discussed roles of R-loops in replication stress, DNA damage induction, inflammatory signaling and disease.

A well-recognized and well-conserved RNA:DNA hybrid helicase is Senataxin (SETX in humans, or Sen1 in yeast[155]). SETX removes R-loops at transcription termination sites, thereby promoting release of nascent transcripts in a manner that depends on the RNA exonuclease XRN2[156]. SETX is also recruited to DSBs that form in transcriptionally active loci, favoring recruitment of RAD51, preventing aberrant joining of DNA ends and thereby translocations[125,157]. Notably, SETX is also present in the nucleoli, where it binds IGS regions, sites of Pol II-mediated R-loop production. SETX regulates levels of R-loop-forming lncRNAs at IGSs, thereby contributing to nucleolar morphology[118]. SETX is particularly important from a disease perspective, as its gain or loss of function mutations contribute to two neurological disorders: Type 2 ataxia with oculomotor apraxia (AOA2) and juvenile amyotrophic lateral sclerosis (ALS4)[158]. Changes in R-loop homeostasis were observed in AOA2 and ALS4 patients and likely contribute to disorder phenotypes.

1.3.5.2. R-loop nucleases

Two *bone fide* nucleases of the Ribonuclease H (RNase H) family have the ability to specifically bind RNA-DNA hybrids and degrade their RNA moiety: RNase H1 and H2. As RNase H1 is discussed in detail in the section 1.4, I will here focus on RNase H2. It is a trimeric nuclear complex containing the catalytic subunit RNase H2A, and two accessory subunits, RNase H2B and H2C. The catalysis takes place via a two-metal ion mechanism using the active site situated on the H2A subunit[159]. The H2B subunit is thought to have a role in protein-protein interactions, evidenced by its PCNA interacting protein-box (PIP-box) region[160]. The H2C is thought to have a largely structural role, as it is connecting H2A and H2B[159]. RNase H2 associates with the replisome through PIP-box-mediated interactions with PCNA. This might be relevant for its contribution to removing RNA primers during Okazaki fragment maturation[161], as well as its role in ribonucleotide

excision repair (RER)[162]. Its essential role in removal of mis-incorporated ribonucleotides is substantiated by the fact that, in the absence of RNase H2 function, topoisomerase Top1 can aberrantly process rNMPs, leading to deletions and DSBs[163–165]. In yeast, RNase H2 expression is cell cycle-regulated and peaks in late S and G2 phase, in line with its post-replicative requirements in RER and R-loop removal[166]. RNase H2 loss results in a stronger genome instability phenotype in yeast than RNase H1 loss[167], likely because RNase H2 is thought to have a broader role in RNA-DNA hybrid removal, and is additionally indispensable for RER. Mutations in RNase H2 cause the Aicardi–Goutières syndrome (AGS), a rare inflammatory neurodevelopmental disorder[168]. Additionally, the autoimmune disease systemic lupus erythematosus is also characterized by recurrent RNase H2 mutations[169], and shares disease features with AGS[170]. Disease pathology of both of these disorders involves inflammatory interferon signaling, which likely occurs as a result of aberrant sensing or accumulation of self-derived immunogenic nucleic acids species[171].

1.3.5.3. R-loop-suppressing topoisomerases and translocases

Topoisomerases are cellular enzymes that regulate the topological state of DNA and RNA[172], and are able to resolve topological stress that results from nucleic acid transactions. Topoisomerase 1 (TOP1) introduces DNA nicks, allowing the DNA duplex to untwist[173], thereby relieving both positive and negative supercoils during replication and transcription[174]. As previously mentioned, transcription leads to negative supercoiling behind the RNA polymerases, which favors R-loop formation. Furthermore, R-loops are especially problematic at head-on TRCs. Accordingly, TOP1 prevents TRCs and replication stress at R-loop-enriched transcription termination sites of highly expressed genes[175]. Moreover, loss of yeast Top1 leads to R-loop stabilization within Pol I-transcribed rDNA repeats, which is exacerbated by the loss of RNase H activity[105].

A recent study implicated another class of enzymes, translocases, in R-loop resolution[176]. Translocases use ATP to anneal strands at DNA junctions, thereby moving the branchpoint between the dsDNA and an alternative DNA structure. Such junctions are found at stalled replication forks or transcription bubbles. Three different translocases, FANCM, ZRANB3 and SMARCAL1, were found to displace R-loops *in vitro* and suppress them *in vivo* in human cells. This suggests that, in addition to triggering fork reversal, fork

remodelers also remove R-loops, perhaps in similar situations when fork reversal is necessary (eg. at TRCs)[176].

1.3.5.4. Other R-loop-suppressing proteins

Productive transcriptional elongation and efficient RNA processing help to keep R-loop levels low[137,177]. Conversely, defects in RNA splicing lead to aberrant R-loop accumulation. This is likely because defects in splicing expose the free 5'-RNA end necessary for R-loop formation[178]. Additionally, excision of introns reduces complementarity of RNA with the DNA template, decreasing the likelihood of R-loop formation. Aquarius (AQR) is an RNA helicase/ATPase and an essential component of the pentameric intron-binding complex[179]. The activity of this complex leads to the activation of the B spliceosome[180]. Depletion of AQR causes strong R-loop accumulation and a consequent increase in DSBs[145]. It is, however, not clear if AQR is a direct RNA-DNA hybrid helicase, whether its splicing role or yet another activity cause the R-loop increase observed after its depletion. Mutations in general splicing factors SRSF2 and U2AF1 are commonly found in the myelodysplastic syndrome (MDS) and also lead to R-loop accumulation, DNA damage and ATR activation. Elevated R-loops have therefore emerged as a unifying mechanism in MDS associated with splicing factor mutations[181].

Even before the discovery of the role of splicing in R-loop homeostasis, it became clear that RNA processing proteins contribute to R-loop regulation. The THO/TREX complex, which directs the export of mRNA from the nucleus to the cytoplasm, prevents R-loop accumulation and enables efficient transcriptional elongation in both yeast and human cells[182,183]. Mechanistically, THO promotes the recruitment of the histone deacetylase Sin3a, leading to transient histone deacetylation, which likely disables R-loop formation behind the RNA polymerase[184].

Finally, two tumor suppressors, core DNA damage and fork stability factors, BRCA1 and BRCA2, have a role in R-loop homeostasis. This role likely contributes to their canonical functions in both DSB repair and fork protection. BRCA1 antagonizes 53BP1 at DSBs and thereby opposes non-homologous end joining (NHEJ) and favors HR[185]. Recent results by multiple groups suggest that BRCA1 is a direct RNA-DNA hybrid binder and is recruited to DSBs by hybrids[131]. BRCA1 is known to favor BRCA2 recruitment to resected DSBs[186], which in turn results in RAD51 loading[187] and homology search. Importantly, BRCA2

seems to recruit RNase H2 to DSBs[131], helping to remove the hybrids, presumably to enable RAD51 loading. Therefore, recent models suggest that BRCA1 is recruited to DSBs by hybrids, and that the hybrids are in turn removed in a BRCA2-dependent manner[188]. Additionally, BRCA1 recruits the helicase SETX to transcription termination sites to resolve R-loops and prevent R-loop-associated single-strand breaks[189]. Whether BRCA1 also recruits SETX to sites of DSBs is less clear. BRCA1 also seems to suppress R-loops at promoter-proximal sites[190]. In neuroblastoma cells, where the oncogenic transcription factor MYCN is expressed, BRCA1 is recruited to sites of promoter-proximal pausing in a MYCN-dependent manner to aid R-loop removal and enable transition into elongation[191]. Altogether, these and many other factors lacking enzymatic activities necessary for R-loop removal maintain R-loop levels through more indirect mechanisms.

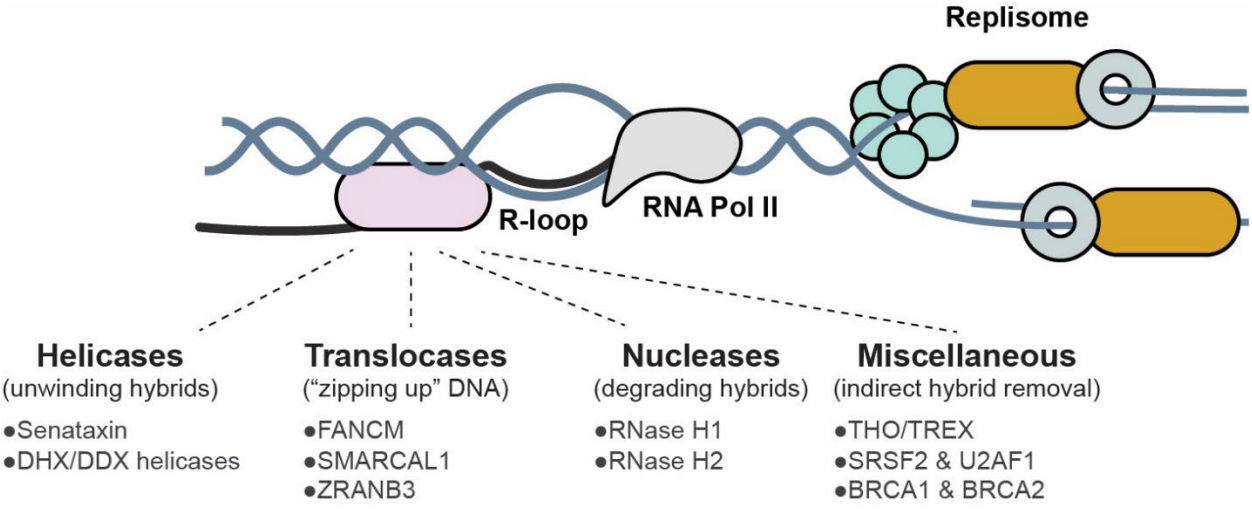


Figure 2: R-loop-removing enzymes.

1.4. RNase H1

1.4.1. Domain composition, localization and conservation

The RNase H enzymes were discovered and isolated from calf thymus samples in the laboratory of Peter Hausen in Tübingen (1969), and given the name indicating the RNA-DNA *hybrid* specificity[192]. RNase H was found in crude isolates of RNA polymerase, decreasing its ability to synthesize RNA from denatured DNA templates five-fold. Along with RNase H2, RNase H1 is the main nuclease able to degrade the RNA strand of the RNA-DNA hybrids in both eukaryotes and prokaryotes. Members of the RNase H family can be found in essentially all organisms, including bacteria and archaea[193]. Human

RNase H1 consists of the N-terminal mitochondrial localization signal (MLS), followed by the hybrid-binding domain (HBD). The catalytic domain (also called the H-domain) is located at the C-terminus, with the unstructured linker connecting the HBD and the catalytic domain[194]. The relatively long linker (around 60 amino acids) allows the catalytic domain to access the hybrid substrates at multiple sites, which may increase enzyme processivity[195]. The full length enzyme is a 286-amino acid protein (32 kDa) expressed ubiquitously in human cells and tissues. Two isoforms of RNase H1 are present in human, both expressed from the same mRNA, the shorter one lacking the MLS. Synthesis of the shorter isoform occurs by alternative translation initiation site at M27, the initial methionine of the HBD[196]. In line with the domain composition, the longer isoform localizes to the mitochondria (mtRNase H1), whereas the shorter localizes to the nucleus (nucRNase H1 or M27-RNase H1). The latter is expressed 7-fold higher than the former, which was explained by the more efficient translational initiation from M27 than from M1. The low expression of mtRNase H1 might be an evolutionary feature essential to avoid toxicity resulting from mitochondrial overexpression[196]. Strong nucleolar enrichment of the nuclear isoform is mediated by the HBD[197]. The truncated version containing only the linker and the catalytic domain localizes to the nucleoplasm, but is excluded from the nucleoli, suggesting that these two domains are sufficient for nuclear localization, although a typical nuclear localization signal (NLS) is missing. Interestingly, RNase H1 is very well conserved. The human version has considerable sequence similarity (33.6% amino acid identity) even with the *E. coli* RNase H1[198]. As expected, the structured domains (HBD and the catalytic domain) are better conserved than the linker.

1.4.2. HBD – the hybrid-binding function of RNase H1

The architecture of RNase H1 is well suited for recognition and removal of RNA-DNA hybrids. The human HBD comprises around 50 amino acids and enhances the processivity of RNase H1[199]. Based on the mutational analysis and biochemical binding assays, three amino acids (W43, K59 and K60) of the HBD have emerged as the critical determinants of the heteroduplex positioning[200]. While K59 and K60 bind the heteroduplex, W43 is responsible for correct positioning of RNase H1 on the substrate for catalysis. Structural models of the HBD with a 12-bp hybrid substrate obtained by Cryo-EM explained HBD's 25-fold preference for hybrids over other nucleic acid substrates by

examining residues critical for binding to either the RNA or the DNA moiety[195]. The RNA-binding loop (D51-A56) mediates binding of the ribose 2'-OH groups to main chains of R52 and A55. The aromatic patch (W43 and F58) is structurally complementary with the DNA strand and critical for hybrid specificity, as substituting deoxyribose with the ribose results in a sterical clash. K60 is one of the few residues directly interacting with the DNA phosphates. These results further substantiate the critical role of the WKK residues in hybrid recognition.

1.4.3. Catalytic domain – the hybrid-degradation function of RNase H1

The catalytic domain provides the endonuclease function and consists of around 150 residues, including four conserved carboxylates (D145, E186, D210, and D274) forming the active site, the former three being essential for the enzymatic activity[201]. Similarly to RNase H2, catalysis occurs through a two-metal ion mechanism. Mg^{2+} is the preferred ion, and Mn^{2+} has an inhibitory effect in the presence of Mg^{2+}[198]. Unlike RNase H2, which can remove single ribonucleotides from dsDNA, RNase H1 requires at least a 6-bp long hybrid substrate, but the optimal cleavage position is 8-12 nucleotides from the 5′-RNA-3′-DNA terminus[202]. While the HBD specializes in hybrid binding, the catalytic domain also contacts both RNA and DNA strands of the hybrid using well-conserved residues in multiple binding pockets that primarily contact the minor groove[203].

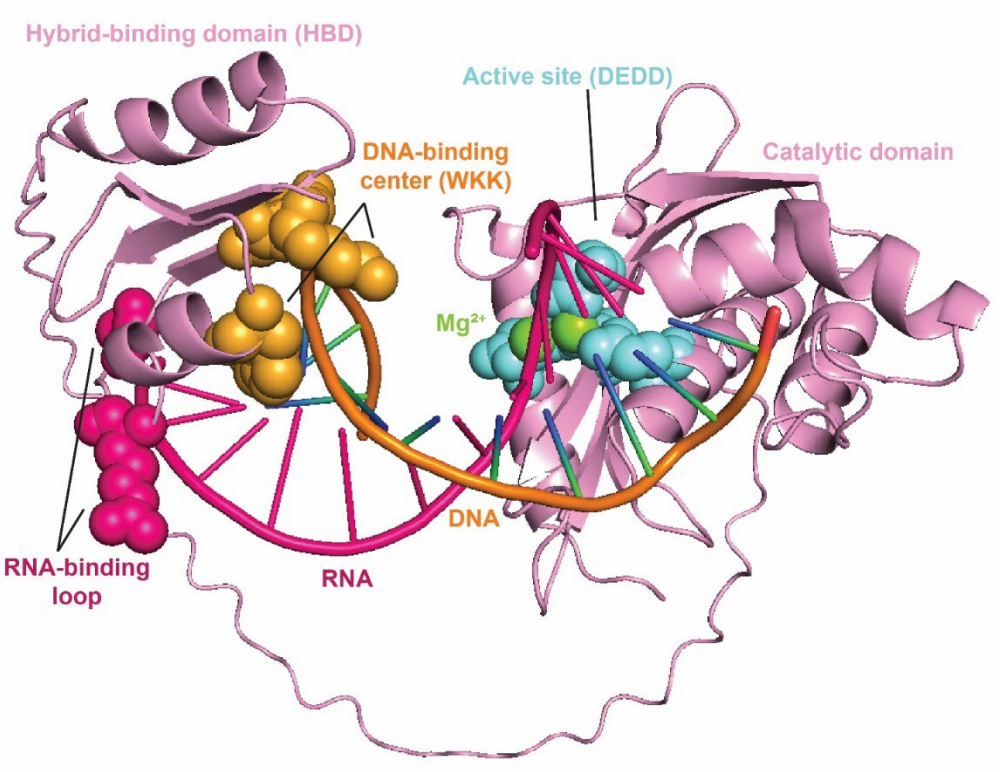


Figure 3: Structure of human RNase H1.

AlphaFold3-modeled structure of human nuclear isoform of RNase H1 in complex with a 12mer RNA:DNA hybrid rendered in PyMol. The main structural features that enable hybrid binding and cleavage are highlighted. Two Mg^{2+} ions positioned in the active site are shown.

1.4.4. Biological roles and regulation

The mammalian RNase H1 has been best studied in the mitochondria. It is essential for processing of RNA primers during the replication of mitochondrial DNA (mtDNA). Likewise, it is involved in the completion of mtDNA replication. In cells lacking RNase H1, linear mtDNAs with deletions are formed instead of proper circular mtDNA[204]. As the complete knockouts of the mouse *Rnaseh1* gene are embryonic lethal[205], conditional knockouts in the mouse heart have been developed to study the RNase H1 function *in vivo*[204]. In the conditional knockout, mitochondrial transcripts are steadily lost during development. While this is likely a consequence of the function of RNase H1 in the replication of mtDNA, RNase H1 also reportedly functions in mitochondrial transcription and translation by regulating mitochondrial mRNAs, rRNAs, and the 7S non-coding RNA[206]. Studying human RNase H1 is disease-relevant, as mutations in the catalytic domain lead to the adult-onset mitochondrial encephalomyopathy[207], also known as the

Autosomal recessive progressive external ophthalmoplegia (PEO) with mitochondrial DNA deletions-2. Recently, the first patients with a mutation in the HBD (Y29C) were identified, displaying symptoms of PEO. Interestingly, this mutation was shown to increase the RNase H1 activity, suggesting that tight coordination of the mtRNase H1 activity is necessary for its mitochondrial function[204].

Several studies in yeast offer valuable insights into functions of the RNase H1, and hint towards potential function of the nuclear isoform in mammalian cells. Yeast RNase H1 functions independently of the cell cycle, but seems to only degrade hybrids upon their accumulation[166]. The yeast *RNH1* gene can be deleted with modest increases in sensitivity to DNA-damaging agents[208]. Although yeast RNase H1 is able to associate with R-loops genome-wide, it removes them only at strong R-loop-forming loci[209]. These include Pol III-transcribed tRNA and 5S rDNA loci[210]. The fact that RNase H1 degrades only a subset of R-loops implies it is somehow restricted from activity, presumably by either protein-protein interactions, PTMs[209], or both.

In humans, nuclear RNase H1 is present at telomeres specifically in ALT cells, where it regulates TERRA. Both depletion and overexpression of RNase H1 have negative consequences on telomere integrity in ALT cells[211]. Furthermore, RNase H1 interacts and co-localizes with RPA[212], with HBD residues R32, R33 and R52 being important for the interaction. However, it is not clear whether this interaction is independent of nucleic acids. Along these lines, the interaction of RNase H1 with RPA depends on the presence of DNA in yeast[210]. *In vitro*, RPA promotes human RNase H1 association with and activity towards hybrids with a DNA overhang or R-loops[212]. As previously mentioned, the nuclear isoform of RNase H1 is present in the nucleoplasm and further enriched in the nucleoli. Upon inhibition of RNA Pol I, RNase H1 and hybrids are transported from the nucleoli to the perinucleolar ring structures. Moreover, after inhibition of TOP1 with camptothecin, RNase H1 is further accumulated in the nucleoli, along with the hybrids. Accordingly, depletion of RNase H1 results in the accumulation of Pol I-transcribed R-loops[197]. This mirrors data from yeast showing that, upon Top1 deletion, RNase H enzymes are critical for removal of Pol I-derived R-loops[105].

Human RNase H1 has been detected at laser-induced microirradiation tracks, representing sites of DSBs. This suggests a potential role of RNase H1 in removal of DSB-associated hybrids. The potential impact of RNase H1 on TRCs is poorly understood, although it has recently been reported that overexpression of RNase H1 increases fork speed[106] and rescues fork degradation upon loss of BRCA2[213], suggesting a role at replication forks.

Antisense nucleotides (ASOs) are DNA oligos that can be used to regulate gene expression and have therapeutic applications, for example in neurological disorders[214]. They are designed to be complementary to endogenous target cytoplasmic or nuclear RNAs, and can thus form RNA:DNA hybrids. Upon hybrid formation, the RNA is cleaved in an RNase H1-dependent manner[215], thereby reducing the pool of target RNA available for translation. Therefore, a better understanding of RNase H1 biology could improve therapeutic applications of ASOs.

1.4.5. A tool in R-loop biology

RNase H1 is emerging as the gold-standard tool to detect and manipulate R-loops in human cells[216]. The tools for hybrid detection based on RNase H1 are increasingly used because the traditional way of detecting R-loops using the S9.6 antibody has been brought into question. Namely, in addition to RNA:DNA hybrids, S9.6 seems to also recognize other RNA species in immunofluorescence approaches, evident from its sensitivity to RNase T1 treatment[217].

Overexpression of RNase H1, either after transient transfection[218] or in inducible cell lines[219], is used to remove R-loops genome-wide. This is often used to argue that the observed phenotypes are R-loop-dependent. Expression of fluorophore-tagged D210N-RNase H1[212] or the HBD-GFP[220] has been used to detect nuclear R-loops by immunofluorescence. A similar approach was developed with purified D210N-RNase H1 for use in fixed cells[221]. Additionally, a recent tool called RHINO consists of a GFP-tagged triple HBD and enables detection of R-loops in live cells[116]. Point mutations of the hybrid-binding WKK residues to alanine offer an additional convenient specificity control, which we have exploited in R-loop proximity proteomics (HBD-APEX2)[153]. A similar approach, based on the D210N-RNaseH1 fused to TurboID[222], was concomitantly developed. Tools

based on RNase H1 are also finding use in genomic approaches to map R-loops, including the Cut&Run-based MapR[223] and the ChIP-seq-based R-ChIP[224].

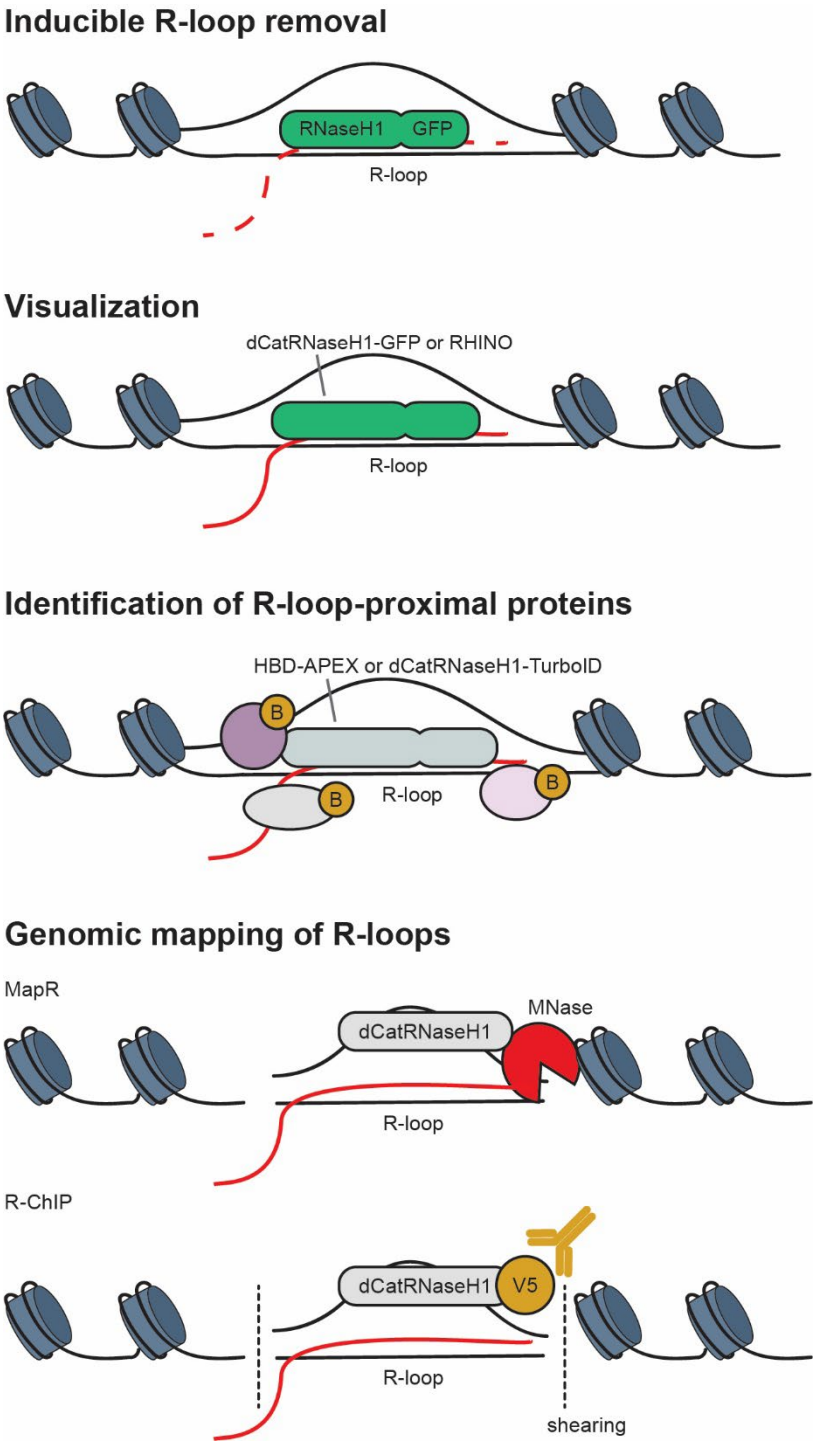


Figure 4: RNase H1 as a tool in R-loop biology.

Genetically engineered RNase H1 can be used to remove or visualize R-loops *in vivo*, to map R-loop-proximal proteins or genomic regions containing R-loops.

1.5. Aims of this work

In this project, we aimed to get a deeper understanding of the relationship between replication stress and the resolution of R-loops. For this purpose, we decided to investigate the phosphorylation response to replication stress, with a focus on R-loop-resolving enzymes. In the course of the project, phosphorylation of the R-loop-degrading nuclease RNase H1 emerged as a promising lead. Considering the ubiquitous use of RNase H1 as a gold-standard research tool in R-loop biology (section 1.4.5), surprisingly little is known about how it is regulated in human cells. For example, it is not clear whether the activity of human RNase H1 is simply determined by its amount, or by additional mechanisms, such as post-translational modifications and protein-protein interactions. Furthermore, it is not clear whether human RNase H1 can recognize all RNA:DNA hybrids or only a subset, and how this specificity is accomplished. Answering these and other questions will improve the design of tools based on RNase H1 and inform the users about the most appropriate tool for their end goal. From the perspective of kinase signaling, we aimed to find novel aspects of DYRK1A signaling, particularly with regards to RNA processing and R-loop biology. DYRK1A is dysregulated in disease, for example in Down syndrome. We realized that its function in RNA processing is understudied in the context of Down syndrome, and might be relevant for some of the pathologies related to this syndrome.

Chapter 2

Materials and methods

2.1. Reagents

2.1.1. Chemicals and kits

Table 1: List of chemicals and kits used in the study.

Name	**Supplier**
Doxycycline	Sigma
Puromycin	Invivogen
Aphidicolin	Santa Cruz
CX-4945	Selleckchem
Harmine	Selleckchem
INDY	MedChem Express
L41	Selleckchem
DMSO	Sigma
CHIR99021	Sigma
Lipofectamine® RNAiMAX	Thermo
QuickStart Bradford 1x Dye Protein Reagent	BioRad
NuPAGE™ LDS Sample Buffer (4X)	Thermo
Dithiothreitol	Thermo
NuPAGE 4-12% BT mini protein gels	Thermo
NuPAGE MOPS SDS Running Buffer (20X)	Thermo
Nitrocellulose blotting membranes	Sigma
Trans-Blot Turbo RTA Transfer Kit, nitrocellulose, mini	BioRad
Trans-Blot Turbo 5x Transfer Buffer	BioRad
Ponceau S	AppliChem
Immobilon® ECL Ultra Western HRP Substrate	Sigma
SuperSignal West Pico	Thermo
Hoechst 33342	Sigma
Urea	Sigma
Thiourea	Sigma
2-Chloroacetamide	Sigma
Trypsin, MS approved	Serva
Sep-Pak C18 Vac Cartridge	Waters

C8 solid phase extraction disc	CDS
TiO2 spheres	GL Sciences
C18 Empore 47 mm extraction disks	CDS
Boric acid	Sigma
Phosphoric acid	Sigma
Ammonium hydroxide	Sigma
Protease Inhibitor cocktail	Sigma
Sodium deoxycholate	AppliChem
Sodium fluoride	Sigma
Sodium orthovanadate	Sigma
β-Glycerophosphate disodium salt hydrate	Sigma
hydroxylamine	Sigma
TMT10plex Isobaric Label Reagent Set plus TMT11-131C Label Reagent	Thermo
TMTPro Isobaric Label Reagent Set	Thermo
Magnesium chloride	New England BioLabs
Acetic Acid	Sigma
Methanol	VWR
TFA	Fisher Chemicals
Acetonitrile	Sigma
BSA	Sigma
Tween-20	Sigma
Fetal Bovine Serum	Thermo
Penicillin-Streptomycin	Thermo
L-glutamine	Thermo
Trypsin-EDTA (0,05%), phenol-red	Thermo
Dulbecco's modified Eagle's medium (DMEM)	Thermo
Roswell Park Memorial Institute media (RPMI)	Thermo
Opti-MEM™	Thermo
Poly-L-Lysine	Sigma
Dialyzed FBS (10,000 molecular weight cut-off)	Thermo
D-MEM for SILAC without lysine and arginine	Thermo
L-arginine and L-lysine	Cambridge Isotope Laboratories
L-arginine [13C6] and L-lysine [2H4]	Cambridge Isotope Laboratories
L-arginine [13C615N4] and L-lysine [13C6-15N2]	Cambridge Isotope Laboratories
Polybrene	Sigma
Linear polyethylenimine transfection (PEI, HCl Max, 40000)	Polysciences

2.1.2. Antibodies

Table 2: *List of primary antibodies used in this study.*

ID	Name	Species	Source	Dilution
-	BrdU (BU1/75)	Rat	Abcam, ab6326	1:500
-	BrdU	Mouse	BD Bioscience, 347580	1:100
101	Phospho-CHK1 (S345)	Rabbit	Cell Signaling, 2341	1:1000
174	CHK1	Mouse	Cell Signaling, 2360	1:1000
102	Phospho-CHK2	Rabbit	Cell Signaling, 2661	1:1000
-	CHK2	Mouse	Cell Signaling, 3440	1:1000
147	γH2AX	Rabbit	Cell Signaling, 2577	1:1000
159	Vinculin	mouse	Sigma, V9264	1:10000
109	HA	Rat	Roche, 11867423001	1:1000
369	RNase H1	Rabbit	Proteintech, 15606-1-AP	1:500
-	Phospho-RNase H1 (S74)	Rabbit	Eurogentec, custom-made	1:50
22	GFP	Mouse	SCBT, sc-9996	1:500
417	DYRK1A	Rabbit	Cell Signaling, 2771	1:1000
197	MYC	Rabbit	Abcam, ab32072	1:1000
-	GSK3α/β	Mouse	Thermo Fisher, 44-610	1:1000

Table 3: List of secondary antibodies used in this study.

ID	Name	Species	Source	Dilution
331	Mouse IgG H&L Cy3.5	Goat	Abcam, ab6946	1:100
332	Rat IgG H&L Cy5	Goat	Abcam, ab6565	1:100
-	Mouse HRP	Goat	Agilent, P044701-2	1:10000
-	Rabbit HRP	Goat	Agilent, P044801-2	1:10000

2.1.3. Solutions

Cell lysis buffer for DNA fiber assay: 200 mM Tris-HCl pH 7.4, 50 mM EDTA, 0.5% SDS

Ponceau S solution: 0.1% (w/v) Ponceau S in 5% (v/v) acetic acid

RIPA buffer: 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 2.5 mM $MgCl_2$

mRIPA buffer: 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 2.5 mM $MgCl_2$

Cell extraction buffer: 10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM NaF, 20 mM Na4P2O7, 2 mM Na3VO4, 1% Triton X-100, 10% glycerol, 0.1% SDS, 0.5% deoxycholate

2x Dilution buffer: 10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA

Phosphatase inhibitors: 100 mM sodium orthovanadate, 500 mM sodium fluoride, 500 mM β-glycerophosphate

Urea:thiorea solution: 6M urea / 2M thiourea in 10 mM HEPES pH 8.0

HE200 buffer: 10 mM HEPES, 200 mM NaCl, 0.05% Tween-20, 50 µM EDTA; final pH 7.5

TE40 buffer: 10 mM Tris-HCl, 40 mM NaCl, 0.05% Tween-20, 50 µM EDTA; final pH 7.5

Duplex buffer: 100 mM potassium acetate, 30 mM HEPES, pH 7.5

2.2. DNA oligonucleotides

Table 4: List of DNA oligonucleotides used in this study.

Name	**Sequence**
heliX oligo 1 (26mer)	GCGCTCGCTCACCCTCCTCTACGGCCTTTATC AGTACTTGTCAACACGAGCAGCCCGTATATTCTCCTACAG CACTA
heliX oligo 2 (26mer)	GGCCGTAGAGGAGGGGAGCGAGCGC
heliX oligo 1 (12mer)	TCCTCTACGGCCTTTATCAGTACTTGTCAACAC GAGCAGCCCGTATATTCTCCTACAGCACTA
RNASEH1-S80A-fwd	GCCCCGGAAGTTGCAGAAGGGCATGAAAATCAACA
RNASEH1-S74A-S76A-rev	AACTTCCGGGGCTGCAGCTTTCCTGACAAAGGC
RNASEH1-S80D-fwd	GACCCGGAAGTTGATGAAGGGCATGAAAATCAACA
RNASEH1-S74D-S76D-rev	AACTTCCGGGTCTGCATCTTTCCTGACAAAGGC

2.3. RNA oligonucleotides

Table 5: List of RNA oligonucleotides used in this study.

For siRNAs, information is provided for the sense (5'-3') strand of the duplex.

Name	Sequence 5'-3'	Source
heliX Oligo 2 (RNA-12mer)	GGCCGUAGAGGA	Sigma
siCTRL (pool)	UGGUUUACAUGUUGUGUGA, UGGUUUACAUGUUUUCUGA, UGGUUUACAUGUUUUCCUA	Sigma
siRNase H1 (pool)	UGGUUUACAUGUCGACUAA, UGGUUUACAUGUUGUGUGA, UGGUUUACAUGUUUUCUGA, UGGUUUACAUGUUUUCCUA	Horizon
siDYRK1A (pool)	GGGAAUGCCUUAUGACCUU, AAAAGUGGAUGGAUCGUUA, UAAGGAUGCUUGAUUAUGA, GAAUUCAACCUUAUUAUGC	Horizon
siAQR	CUGAAUAUGGCGGUGUAGU	Sigma

2.4. Plasmids

Table 6: List of plasmids used in this study.

ID	Name	Construction	Use
684	pLIX402-GFP-M27-RnaseH1	Created previously	Lentiviral transduction
1163	pLIX402-M27-RnaseH1-S3xA-GFP	Site-directed mutagenesis using p684 and oligos RNASEH1-S80A-fwd & RNASEH1-S74A-S76A-rev	Lentiviral transduction
1164	pLIX402-M27-RnaseH1-S3xD-GFP	Site-directed mutagenesis using p684 and oligos RNASEH1-S80D-fwd & RNASEH1-S74D-S76D-rev	Lentiviral transduction
37	psPAX2	Created previously	Lentiviral transduction
38	pMD2.G	Created previously	Lentiviral transduction

2.5. Cell lines

Table 7: List of the mammalian cell lines used in this study.

ID	Cell line	Antibiotics	Source
320	HCT116		Holger Bastians, Uni Göttingen
449	HCT116:M27-RNaseH1-GFP	Puromycin	This study
450	HCT116:M27-RNaseH1-S3xA-GFP	Puromycin	This study
451	HCT116:M27-RNaseH1-S3xD-GFP	Puromycin	This study
102	U2OS:CCNE1-HA	Puromycin	Beli lab
273	U2OS:RNaseH1-3xFLAG-APEX2	G418	Beli lab
-	MEF		Jim Woodgett, LTRI Toronto
-	MEF GSK3A -/-		Jim Woodgett, LTRI Toronto
-	MEF GSK3B -/-		Jim Woodgett, LTRI Toronto

2.6. Methods for molecular cloning

2.6.1. Site-directed mutagenesis

To produce the phospho-dead (S3xA) and phospho-mimicking (S3xD) mutant of RNase H1, site-directed mutagenesis was performed on the p684 vector. Plasmids were amplified by PCR using Q5 polymerase (IMB Core Facilities) and mutant-specific primers (DNA oligo table). 5′-phosphorylation was done with T4 Polynucleotide Kinase (NEB) for 30 minutes at 37°C. Subsequent ligation was carried out by T4 DNA ligase (NEB) for 2h at RT. Cloning results were validated by sequencing (Eurofins Genomics).

2.7. Methods for protein analysis

2.7.1. SDS-PAGE and Western blotting

10-20 µg total protein lysate or immunoprecipitate were resolved on 4-12% gradients SDS-PAGE gels and transferred onto nitrocellulose membranes using the Trans-Blot Turbo system or the XCell II Blot system. Following the transfer, membranes were stained with Ponceau S and blocked for 30 min in a 3 % BSA solution in PBS supplemented with with 0.1% Tween-20 (PBS-T). Membranes were incubated with primary antibodies overnight at 4°C. Following 3× 5 min washes with PBS-T, secondary antibodies coupled

to horseradish peroxidase (Agilent) were applied for 1 h at room temperature. Membranes were then washed 3× 10 min with PBS-T and developed with Immobilon ECL Ultra Western HRP Substrate and SuperSignal West Pico detection reagent. Chemiluminescence signal was recorded on ChemiDoc (Bio-Rad).

2.8. Methods for mammalian cells

2.8.1. Mammalian cell culture

All cells were cultured at 37°C with 5% CO2 and all culture media were supplemented with 10% fetal bovine serum, 100 U/mL penicillin-streptomycin and 5mM L-glutamine. U2OS cells were maintained in Dulbecco's modified Eagle's medium (DMEM) and HCT116 cells were maintained in Roswell Park Memorial Institute media (RPMI). Mouse embryonic fibroblasts (MEFs) were maintained in DMEM (high glucose with GlutaMAX Supplement) with 10% FBS, PenStrep and 1x non-essential amino acids. The three stable HCT116 cell lines expressing RNase H1 WT, S3xA or S3xD GFP-tagged under a doxycycline (DOX)-inducible promoter were cultured in RPMI supplemented with 2 µg/mL puromycin. The stable U2OS cell line expressing Cyclin E1 under a DOX-inducible promoter were cultured in DMEM supplemented with 2 µg/mL puromycin. Cells were induced with 2 µg/mL DOX. All cells were tested on a regular basis for mycoplasma using a PCR-based method.

2.8.2. Lentiviral transduction

All manipulations that included lentiviral particles were conducted in the IMB S2 laboratory. HEK293T packaging cells were transfected using a PEI:DNA ratio of 3:1 (w:w) with a packaging vector (p37), an envelope vector (p38) and either of the three transfer vectors (p684, p1163 or p1164). Media was changed from OptiMEM to complete DMEM after 7 h transfection. After 24 h, supernatant containing the viral particles was collected and filtered through a 0.45 µm PES filter. Polybrene was added to the supernatant and the mix was added to HCT116 recipient cells. After the first cell split, cells were cultured in RPMI supplemented with 2 µg/ml puromycin. During next 2 weeks, the cells were monitored and surviving colonies were downgraded using the lenti-X qPCR kit (QIAGEN). Genotype was confirmed by PCR and Western blotting.

2.8.3. Cell treatments and lysis

Cells were treated with 100-1000 nM APH for 24 h or other duration, as indicated in the figures; 5 μM CX-4945, 10 μM Harmine, 10 μM INDY, 10 μM Leucettine-L41 for 1 h; or 10 μM CHIR99021 for 4 or 24 h. As a negative control an equal volume DMSO was used. Knock downs were performed for 48h with siRNase H1, siDYRK1A or siCTRL using RNAiMAX according to manufacturer's protocol. For cell lysis, cells were washed twice with PBS and lysed in RIPA buffer containing 50 mM Tris HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% NP40, 0.1% Na-deoxycholate, 0.1% SDS supplemented with phosphatase inhibitors, protease inhibitor cocktail, 100 U/mL SmNuclease and 2 mM $MgCl_2$ at 4°C for 30 min on a rotation wheel. Next, samples were supplemented with sodium chloride at 450 mM and incubated for 15 min. Protein concentrations were measured using the Bradford assay and adjusted and the supernatant proteins were mixed with 4x LDS Sample Buffer supplemented with 1 mM DTT and boiled for 10 min at 95°C.

2.8.4. Cellular fractionation

The protocol was performed at 4°C. Cells were lysed using the mRIPA buffer supplemented with protease and phosphatase inhibitors as described above. They were incubated for 20 min rotating, then spun down for 20 min at 10000 xg. The supernatant was saved (soluble fraction). The pellet (chromatin fraction) was resuspended in the cell extraction buffer supplemented with 100 U/mL SmNuclease and 2 mM MgCl2. Protein concentrations were measured using the Bradford assay and adjusted and the supernatant proteins were mixed with 4x LDS Sample Buffer supplemented with 1 mM DTT and boiled for 10 min at 95°C.

2.8.5. GFP-based pulldowns

Cells were lysed as described in the section "Cell treatments and lysis". After taking 50 μg for inputs, samples were diluted using salt-less RIPA (1:1 V:V) supplemented with protease and phosphatase inhibitors. Magnetic GFP binder agarose beads (IMB Protein Production) were equilibrated in the ice-cold 2x Dilution buffer. 25 μl of the bead slurry was combined with an equal mass of each lysate. Samples were rotated overnight at 4°C, then washed 3x with RIPA buffer diluted 1:1 with the 2x Dilution buffer and transferred to a new tube during the last wash. Samples were boiled in 40-80 μl of 2x NuPAGE LDS

Sample Buffer (Life Technologies) supplemented with 2 mM DTT for 15 min at 70°C. Samples were used either for Western blotting, or for interactome analysis.

2.8.6. Cell cycle analysis by flow cytometry

HCT116 cells were collected with trypsin, washed with PBS and fixed overnight with 70% EtOH at 4°C. Pelleting post-fixation was performed at 3500 xg. Cells were washed once in the wash/permeabilization solution (0.05% Triton-X-100 in PBS), and once in PBS with 2% BSA. To remove RNA, cells were treated for 30 min with RNase A (1 mg/ml) in PBS at RT. They were resuspended in PBS containing 1 mg/ml propidium iodide (PI). Acquisition was performed on a BD LSRFortessa flow cytometer and 40 000 cells were recorded. The 355 nm laser and 450/50 band pass filter were used for analyzing PI fluorescence. Further analysis was done using FloJo. After gating for singlets (FSC/SSC, then PI-H/PI-A) cell cycle profile was determined based on PI staining.

2.8.7. DNA fiber assay

HCT116 cells were labeled with 5-Chloro-2′-deoxyuridine (CldU, 30 µM) for 30 min, washed once with warm PBS, then labeled for 30 min with 5-Iodo-2′-deoxyuridine (IdU, 340 µM). Cells were treated with APH for 1.5 h, 30 min before labeling, and 2x30 min during the labeling pulses. After labeling, cells were washed once with warm and 3× with cold PBS, then trypsinized and spun down (300×g, 5 min). They were resuspended in cold PBS, counted and diluted to 5×10^5/ml. Labeled cells were diluted with twice the number of unlabeled cells. 4 µL of the cell suspension were mixed with 7.5 µl of the lysis buffer for fiber assay directly on the SuperFrost Plus microscopy slide (Thermo Scientific) and incubated horizontally for 9 min. The slides were then tilted at 30°−45°, allowing DNA fibers to spread to the bottom of the slide. DNA spreads were air-dried and fixed with 3:1 methanol:acetic acid overnight at 4 °C. The fibers were rehydrated 3 × 3 min in PBS, dipped once in Milli-Q water and denatured in 2.5 M HCl for 1.5 h at RT, then washed 5 × 3 min in PBS. The slides were blocked for 40 min in the blocking solution (2% BSA, 0.1% Tween-20 in PBS) and incubated with primary antibodies (mouse anti-BrdU, 1:100, BD Bioscience and rat anti-BrdU, 1:500, Abcam) at RT for 2.5 h. After 3 × 5 min washes with PBS-T, the slides were incubated with secondary antibodies (goat anti-mouse Cy3.5, Abcam and goat anti-rat Cy5, Abcam) at RT in the dark for 1 h. The spreads were washed

3 × 5 min with PBS-T, dipped in Milli-Q water, and air-dried completely in the dark. The slides were mounted using Prolong Gold AntiFade mountant (Thermo Scientific), imaged with Visiscope 5-Elements Spinning Disc Confocal microscope (Visitron Systems, Germany) (magnification: ×60 Water immersion objective with ×2 extra magnification; laser lines and corresponding emission filters: 640 nm, 692/40 and 561 nm 623/32) and quantified using the Fiji/ImageJ software.

2.8.8. Fluorescent microscopy using Opera Phenix

Knockdowns in HCT116 cells were performed in 6-well plates and cells were re-plated into PhenoPlates (Perkin Elmer) after poly-lysine coating. All washes were performed in PBS. Cells were washed 2× and fixed with ice-cold methanol for 5 min at –20°C, washed 3x, and permeabilized with Triton X-100 (0.3%) for 5 min at room temperature, followed by 2x washes. Cells were then blocked for 30 min with 3% BSA in PBS-T. Incubation with HBD-AF647 (250 nM) diluted in the blocking buffer supplemented with 1 µg/ml Hoechst 33342 was performed overnight at 4 °C and followed by 3× washes with PBS-T. Cells were kept at 4 °C in PBS until imaging. Imaging was performed with an Opera Phenix (PerkinElmer) microscope using a ×40 1.1NA water objective. Image analysis was performed by using Harmony High-Content Imaging and Analysis Software (version 4.4, PerkinElmer). Standard building blocks allowed for nuclei segmentation based on the Hoechst signal and cells on the edges of the field were excluded. Mean intensity measurements were performed for maximum projections and spot detection was calculated by using algorithm B. The live-cell setup was performed at 37°C with 5% CO_2.

2.9. *In vitro* methods

2.9.1. Protein purification

M27-RNaseH1-D210N (phospho-WT, S3xA and S3xD) fused with a His-MBP-3C tag was expressed in BL21 (DE3) *E. coli* cells using 0.5 mM IPTG at 18°C overnight in LB. Lysis was performed using sonication in IMAC buffer with 500 mM NaCl. First round of purification was performed on a HisTrap, followed by 3C protease digest overnight, heparin column purification to remove nucleic acids and concentration of the elution peak

using Amicon spin-concentrator. This was followed by gel filtration (Superdex 75 16/60 pg). Purified proteins were stored snap-frozen and stored at -80°C.

The HBD-AF647 was purified by the IMB Protein Purification. Detailed information on its design, purification and validation will be described in a PhD thesis by redacted (Petra Beli lab).

2.9.2. HeliX assay

Lyophilized single-stranded oligos (Oligo 1: ligand strand + target sequence; Oligo 2: overhang RNA or DNA complement) were resuspended in the Duplex buffer (IDT) at 100 µM. Oligo 1 and 2 were annealed (=conjugated ligand strand) in equimolar amounts (500 nM final) by mixing for 2 min at 94°C and then cooling at RT. Equal volumes of the 500-nM conjugated ligand strand and 400-nM Adapter strand 1-Ra were mixed and incubated for 20 min at RT at 600 rpm in the dark. Next, pre-incubated ligand-adapter mix was further mixed with an equal volume of Adapter strand 2-lfs. Serial dilutions (1/4) of 500 nM dCat M27-RNaseH1 (phospho-WT, S3xA, S3xD) were prepared in HE200 buffer on ice. The heliX device was supplied with reagents as described by the manufacturer (regeneration, passivation, test & standby solution, TE40 buffer, running buffer: HE200, biochip). Samples were run in a randomized order for each replicate, with chip regeneration between each dilution series. K_D values were calculated using continuous amplitude kinetics (real-time referencing + first blank).

2.10. Proteomics methods

2.10.1. Phosphoproteomics

One 10-cm plate was used per treatment. Until acetone precipitation, all steps were performed at 4°C. Cells were collected by washing twice with ice-cold PBS and lysing in RIPA supplemented with protease and phosphatase inhibitors. Lysates were incubated for 15 min, then 5M NaCl was added (1:10 V:V). Cells were sonicated using the Bioruptor Plus (7 cycles, 30s burst, 30s rest, high intensity) to enable chromatin extraction. Lysates were centrifuged for 15 min at 16 000 xg and supernatant concentrations measured using the Bradford assay. Lysates were purified by adding acetone in excess (80% volume)

overnight at -20°C. All subsequent steps were performed at RT. Sampes were peletted at 1000 xg for 5 min and pellets dried on air. They were resolved in the urea:thiourea solution. Samples were reduced with 1 mM DTT for 45 min and alkylated using 5.5 mM chloroacetamide for 30 min. Proteins were digested using 1:50 (w:w) MS-grade trypsin (Serva) overnight at RT. Samples were incubated in TFA (0.5% V/V) for 1 h at 4°C and centrifuged for 10 min at 4,000 *g*. Peptide supernatants were purified using C18 Sep-Pak columns (Waters), eluted in 50% acetonitrile and dried using SpeedVac. Samples were barcoded with 0.1 µg of acetonitrile-dissolved TMTPro or TMT11plex labels (Thermo Scientific) for 1 h and quenched with 5% hydroxylamine for 15 min[225]. An equal aliquot (5% vol) of each sample was mixed and the ratio check was performed as described[226]. After the ratio check, masses were adjusted treatment-wise. Samples were diluted with 0.1% TFA to reach < 3% ACN and purified on SepPak columns, eluted with 50% ACN and acidified with 6% TFA. Phosphopeptides were enriched using titanium dioxide resin as described[227]. Samples were fractionated using either micro-SCX or high-pH reversed phase chromatography (Thermo Fisher) using manufacturer's instructions. Samples were mixed together and desalted using reversed-phase C18 StageTips[228].

2.10.2. GFP-based interactomics

One 15-cm plate was used per treatment. HCT116:M27-RNaseH-GFP (phospho-WT or mutant) cells were treated for 24h with Dox and 1h with L41. DMSO was used as a mock treatment. Pulldowns were performed as described in "GFP-based pulldowns". After boiling in 2x LDS supplemented with 2 mM DTT, samples were alkylated with 10 mM chloroacetamide for 30 min in the dark at RT. To remove detergents, proteins were purified on SP3 beads[229]. Briefly, hydrophilic and hydrophobic Sera-Mag magnetic carboxylate modified particles (Sigma) were combined 1:1 (v:v), washed in water and resuspended in water at 20 mg/ml. 20 µl of combined beads were added to each sample. Absolute ethanol was added to 50% (v/v) and samples were shaken for 5 min at 1000 rpm. After magnet binding, supernatants were removed and the beads were washed three times with 80% ethanol. Proteins were digested on-bead in the digestion buffer (150 mM HEPES, pH 8) with 0.5 µg of MS-approved trypsin (Serva) per sample overnight at 37°C with shaking. After magnet binding, the supernatants were collected, and the beads were washed twice with the digestion buffer. Wash supernatants were combined with

respective peptide supernatants. Digestion was stopped by adding formic acid (FA) to 1%, and precipitates were removed by centrifugation after a 30 min incubation at 4°C. Samples were spun down for 10 min at 4000 xg and purified using StageTips.

2.10.3. FLAG-based interactomics

For SILAC labeling, cells were cultured for at least 5 passages in SILAC DMEM (Invitrogen) supplemented with dialyzed FBS (Invitrogen) and containing either L-arginine and L-lysine (Merck) or L-arginine [13C6] and L-lysine [2H4] (Cambridge Isotope Laboratories). Three 15-cm plates were used per treatment. All steps until bead boiling were performed on ice. U2OS:RNaseH1-FLAG-APEX2 were washed 2x with ice-cold PBS and lysed with mRIPA supplemented with protease inhibitors. To release the chromatin fraction, 2 mM $MgCl_2$ and 1:500 SmNuclease (IMB Protein Production) were added and incubated for 1 h. NaCl was added (450 mM final) and samples were incubated for 15 min. No-salt RIPA was added to reduce NaCl concentration to 150 mM. Samples were centrifuged for 15 min at 16 000 xg and supernatant concentrations measured with Bradford assay. 40 µl of mRIPA-equilibrated FLAG-M2 affinity gel (Merck) were added to equal mass of each sample. Samples were incubated rotating for 1.5 h, then washed 3x with mRIPA. Elution was performed by boiling in 2x LDS supplemented with 1 mM DTT for 15 min at 75°C. Alkylation was performed with 5.5 mM chloroacetamide for 45 min in the dark. Samples were separated by SDS-PAGE on a 4–12% gradient Bis–Tris gel (Invitrogen). Proteins were stained using the Colloidal Blue Staining Kit (Life Technologies) and digested in-gel using 0.6 µg of MS-approved trypsin (Serva) per gel fraction. Peptides were extracted from the gel and desalted using reversed-phase C18 StageTips.

2.10.4. LC-MS/MS

Run parameters for the LC-MS/MS runs are listed in the *Table 8*. Raw data files were analyzed using MaxQuant[230]. Site localization probabilities were determined by MaxQuant using the post-translational modification scoring algorithm. Parent ion and MS2 spectra were searched against a reference proteome database containing human protein sequences obtained from UniProtKB using Andromeda search engine[231]. Spectra were searched with a mass tolerance of 6 p.p.m. in MS mode, 20 p.p.m. in HCD MS2 mode,

strict trypsin specificity, and allowing up to two miscleavages. Cysteine carbamidomethylation was searched as a fixed modification, whereas protein N-terminal acetylation, methionine oxidation, phosphorylation (STY; only for Phosphoproteomics), and N-ethylmaleimide modification of cysteines (mass difference to cysteine carbamidomethylation) were searched as variable modifications. The dataset was filtered based on posterior error probability (PEP) to arrive at a false discovery rate of below 1% estimated using a target-decoy approach[232]. GFP-based interactomics was analyzed using the MSFragger[233].

Table 8: Instrument settings used for MS-based proteomics.

Settings	Phosphoproteomics	FLAG-based interactomics	GFP-based interactomics
LC			
Device name	EASY-nLC 1200	EASY-nLC 1200	Neo Vanquish
Column	55 cm length, 75 mm inner diameter	55 cm length, 75 mm inner diameter	45 cm length, 75 mm inner diameter
Gradient	2.4 to 33.6% ACN, 120 min	2.4 to 33.6% ACN, 90 min	1.6 to 32% ACN, 70 min
MS – full scan			
m/z range	350-1500	300-1650	325-1300
Resolution	60000	60000	120000
Target value (%)	300	300	300
Max injection time (ms)	40	28	25
ddMS2 scan			
Dependent scans	15	15	50
Normalized collision energy (%)	33	30	26
Resolution	15000	15000	-
Target value (%)	100	100	100
Max injection time (ms)	40	40	10
Isolation window (m/z)	0.8	1.4	1.4
TurboTMT	TMTPro	-	-

2.11. Bioinformatics analysis

2.11.1. Statistical analysis of MS data

Statistical analysis and MS data visualization were performed using the RStudio software environment (version 2024.04.2). Potential contaminants, reverse hits, hits only identified by site and hits with no unique peptides were excluded from the analysis. Phospho-sites were filtered based on localization probability (>75%). Statistical significance was calculated using a moderated t-test (limma package)[234].

2.11.2. Kinase predictions

Computational prediction of kinase activity based on phosphoproteomics was performed using one of the three methods. For Kinase-Substrate Enrichment Analysis, it was based on the KSEA algorithm[235] and the R implementation of KSEA App[236]. Kinase-substrate annotations were obtained from PhosphoSitePlus (PSP) and the NetworKIN database[237]. The analysis was performed with a minimum NetworKIN score of 5 for upregulated and downregulated phosphorylation sites with a p-value ≤ 0.05. For Kinase Library[7] method, differential expression-based enrichment analysis was performed with a $\log_2$-fold change threshold of 0.3 and p-value threshold of 0.05. For the KEA2 method[238], significantly enriched phospho-sites were used as input for the literature-based kinase-substrate library (phospho-site level).

2.11.3. Miscellaneous bioinformatics analysis

Gene Ontology (GO) enrichment analysis was done using ViSEAGO package[239]. Phospho-motif analysis was performed on significantly upregulated phospho-sites using iceLogo[240]. Multiple alignment of RNase H1 from different species was performed with Clustal Omega[241] using default parameters. Results were visualized using JalView[242] and colored according to Percentage Identity. Structure modeling was performed with AlphaFold3[243] and visualized using PyMol.

Chapter 3

Results

3.1. Stratification of replication stress by severity

Replication stress (RS) affects dividing cells in complex ways. It leads to changes in replication timing, origin firing, nuclear architecture, chromatin and replication fork remodeling, and activates checkpoint signaling and DNA repair. Importantly, the extent to which these processes are modulated, and the precise functional outcomes[244], depend on RS severity. Traditionally, strong treatments were used to investigate cellular responses to RS, as these treatments elicited easily discernible phenotypes and led to massive changes in signaling. Recently, mild RS has emerged as a new candidate of intense study. This was triggered by the following observations: (1) mild RS allows progress through the cell cycle, enabling long-term accumulation of phenotypical consequences such as mutations[85]; (2) by allowing entry into G2 phase and mitosis, mild RS allows cell cycle-specific pathways, such as post-replicative repair[245] or mitotic DNA synthesis (MiDAS)[246], to take place; (3) mild RS slows down replication forks, while strong RS blocks them, leading to different fork remodeling outcomes[247]; (4) endogenous RS that cells experience as a result of oncogene induction or replication over difficult-to-replicate regions is considered mild in nature[248]; (5) mild RS leads to both numerical and structural changes of chromosomes (chromosome instability or CIN)[81]; (6) mutations and CIN experienced by untransformed somatic cells as a result of mild RS are thought to be tumor-promoting[249]. Therefore, an improved understanding of cellular responses to mild RS is critical for a better understanding of tumorigenesis and improvement of therapeutic outcomes[250]. With this in mind, we set out to study global phosphorylation signaling upon different levels of RS.

As the first step, we stratified RS by severity using multiple phenotypic readouts. To modulate RS levels, we used a naturally occurring organic compound and B-family DNA polymerase inhibitor aphidicolin (APH). As the experimental system of choice, we used

the chromosomally stable colorectal tumor cells HCT116. These are often utilized to study MiDAS and the induction of CIN upon RS[251,252]. We use them throughout the study, unless specified otherwise. Treating cells acutely with increasing doses of APH led to a stepwise decrease in the replication fork progression rate, as measured by the DNA fiber assay (Figure 5A). Likewise, replication and DNA damage checkpoint signaling, measured by the activation of ATR-pCHK1 and ATM-pCHK2 pathways, was triggered in a dose- and time-dependent manner upon APH treatment (Figure 5B). Increased phosphorylation of the histone variant H2AX at S139 (commonly referred to as γH2AX), an early event upon DSB induction and perhaps the most common marker of increased RS and DNA damage, was used as an additional readout. Finally, we investigated the influence of chronic APH treatment (24 h) on the cell cycle profile (Figure 5C). Interestingly, each of the applied RS doses led to a markedly different cell cycle profile. 100 nM APH, the lowest dose tested, caused a mild increase in the early S-phase population, but otherwise did not affect the cell cycle significantly. 400 nM APH, the medium dose, led to an accumulation of cells in G2 phase, likely because it allows replication, albeit at a much slower rate. 1000 nM APH, the strongest dose tested, caused a broad accumulation across S and G2 phase, likely as a result of the inability of early S-phase cells to finish replication, and delayed entry of late S-phase cells into G2. Altogether, these phenotypes allowed us to stratify RS into mild (100 nM APH), moderate (400 nM APH) and strong (1000 nM APH) RS (Figure 5D).

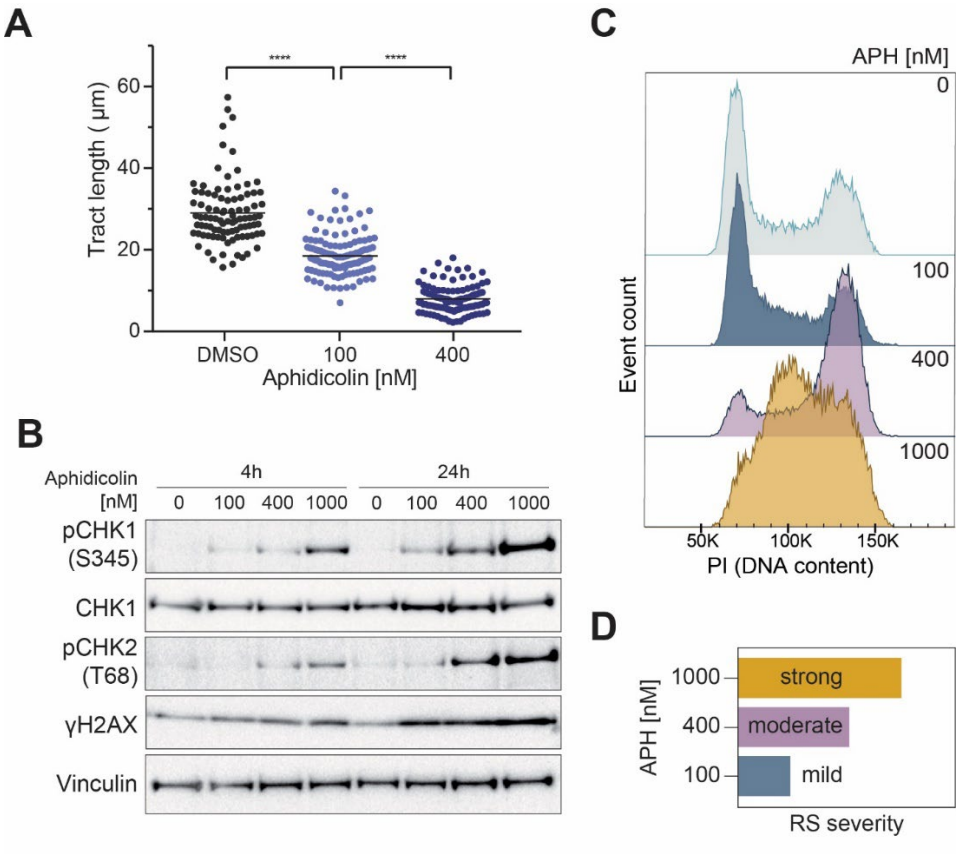


Figure 5: Stratification of replication stress by severity.

A) DNA fiber assay representing combined CldU and IdU track length. Both 30 min pulses and a 30 min pre-treatment were performed with indicated treatments in HCT116. **B)** Western blot analysis of HCT116 lysates treated for 4h or 24h with indicated APH concentrations. Vinculin is used as a loading control. **C)** Flow cytometry histogram after a 24h treatment of HCT116 with indicated APH concentrations. PI = propidium iodide. DNA content is in arbitrary units. **D)** A scheme of replication stress stratified by severity.

3.2. Phosphoproteomics of the dose-dependent response to replication stress

Previous phosphoproteomics studies have focused on the replication stress (RS) levels that strongly activate the checkpoint and do not allow the entry of cells into mitosis[253]. However, it is currently unknown to which extent the various phosphorylation-mediated signaling pathways are induced or inactivated upon lower doses of RS. To study the signaling events associated with different RS levels, we performed multiplexed mass spectrometry-based quantitative phosphoproteomics after a 24-h treatment with the benchmarked RS levels in HCT116 (Figure 6A, left). In this experiment, we identified more than 18 000 high-quality phospho-sites (Figure 6A, right). As expected, the number of consistently regulated phospho-sites rose with the RS dose (Figure 6B). Importantly, mild RS led to an increase of 154 phospho-sites and a decrease of 409 sites (Tier 2 sites, fold change>1.5 and FDR<0.05). This demonstrates that mild RS is sufficient to induce changes in signaling, despite not causing a strong induction of the checkpoint (Figure 5B), nor a large perturbation in the cell cycle profile (Figure 5C). Notably, some of the changes in phosphorylation upon mild RS might be due to a slight accumulation of cells in early S phase.

GO terms analysis of the proteins with a significant gain of phosphorylation upon different RS levels revealed that mild RS was linked with biological processes such as “DNA replication” and “histone demethylation” (Figure 6C). This strengthened the notion that studying global phosphorylation upon mild perturbations offers valuable insights into changes in replication dynamics. Proteins with a phosphorylation loss after mild RS mapped to GO terms associated with RNA processing, such as “RNA splicing”. Consistent with the cell cycle profile (Figure 5C), GO terms analysis of the phospho-sites gained upon moderate RS indicated the G2/M cell cycle arrest (Figure 6D). Phospho-sites gained upon strong RS were enriched for processes associated with DSB repair, in line with strong induction of the DSB markers γH2AX and phospho-CHK2 (Figure 5B). As moderate and strong RS led to considerable perturbation of the cell cycle (Figure 5C), we expect that some phospho-sites regulated after these treatments represent genuine response to RS, whereas other sites simply reflect changes in the cell cycle profile.

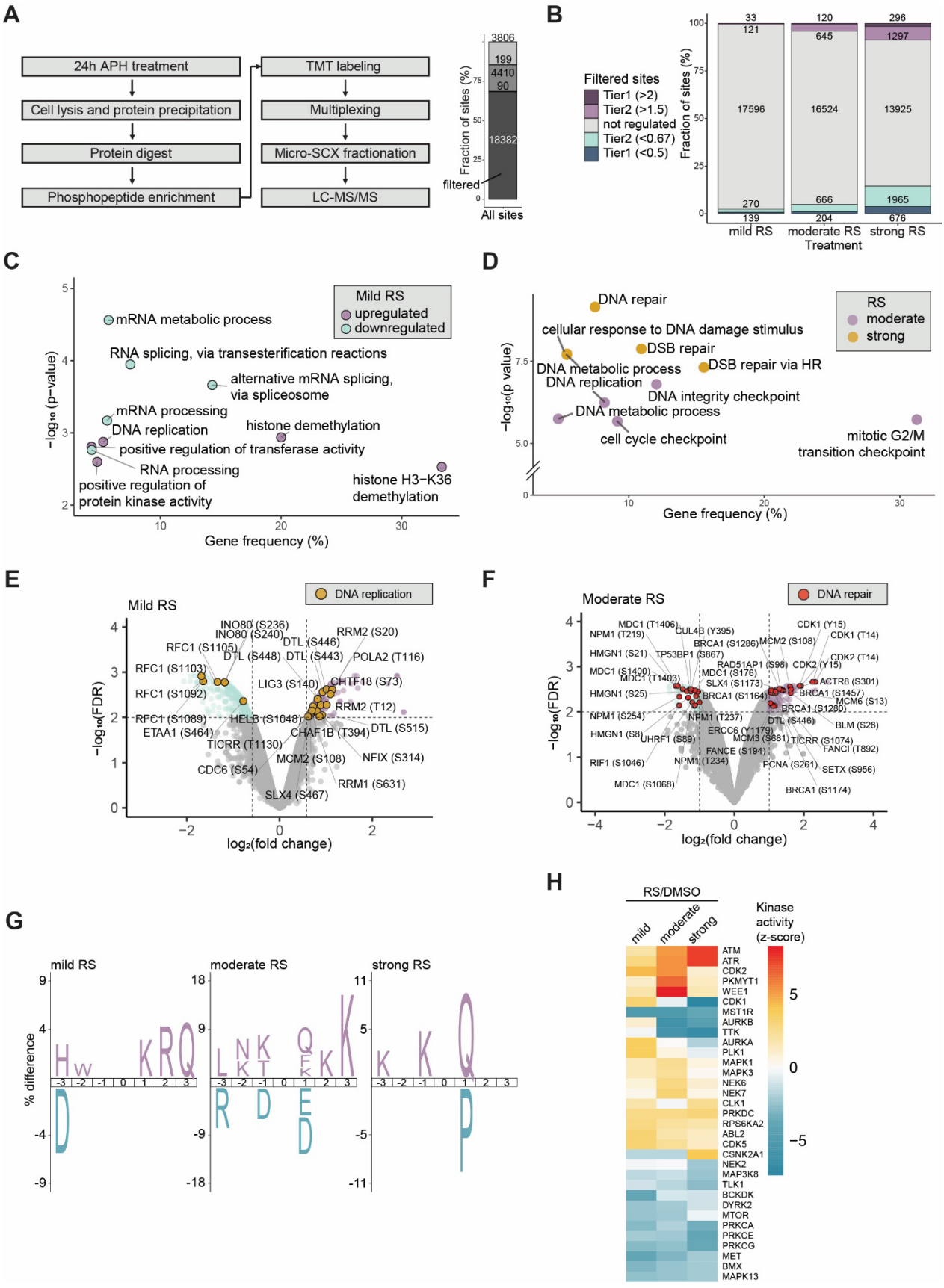

A
24h APH treatment
Cell lysis and protein precipitation
Protein digest
Phosphopeptide enrichment
TMT labeling
Multiplexing
Micro-SCX fractionation
LC-MS/MS
3806
199
4410
90
18382
filtered
Fraction of sites (%)
All sites
B
Filtered sites
Tier1 (>2)
Tier2 (>1.5)
not regulated
Tier2 (<0.67)
Tier1 (<0.5)
33
121
17596
270
139
120
645
16524
666
204
296
1297
13925
1965
676
mild RS
moderate RS
strong RS
Treatment
C
Mild RS
upregulated
downregulated
mRNA metabolic process
RNA splicing, via transesterification reactions
alternative mRNA splicing, via spliceosome
mRNA processing
DNA replication
histone demethylation
positive regulation of transferase activity
RNA processing
positive regulation of protein kinase activity
histone H3−K36 demethylation
−log10 (p−value)
Gene frequency (%)
D
RS
moderate
strong
DNA repair
cellular response to DNA damage stimulus
DSB repair
DNA metabolic process
DSB repair via HR
DNA replication
DNA integrity checkpoint
DNA metabolic process
cell cycle checkpoint
mitotic G2/M transition checkpoint
−log10(p value)
Gene frequency (%)
E
Mild RS
DNA replication
INO80 (S236)
INO80 (S240)
RRM2 (S20)
RFC1 (S1105)
DTL (S446)
DTL (S448)
DTL (S443)
POLA2 (T116)
RFC1 (S1103)
LIG3 (S140)
CHTF18 (S73)
RFC1 (S1092)
RRM2 (T12)
RFC1 (S1089)
HELB (S1048)
ETAA1 (S464)
CHAF1B (T394)
DTL (S515)
TICRR (T1130)
MCM2 (S108)
NFIX (S314)
CDC6 (S54)
RRM1 (S631)
SLX4 (S467)
−log10(FDR)
log2(fold change)
F
Moderate RS
DNA repair
MDC1 (T1406)
NPM1 (T219)
CUL4B (Y395)
BRCA1 (S1286)
MCM2 (S108)
CDK1 (Y15)
CDK1 (T14)
HMGN1 (S21)
TP53BP1 (S867)
CDK2 (T14)
MDC1 (S1400)
MDC1 (S176)
RAD51AP1 (S98)
CDK2 (Y15)
MDC1 (T1403)
SLX4 (S1173)
ACTR8 (S301)
HMGN1 (S25)
BRCA1 (S1164)
BRCA1 (S1457)
MCM6 (S13)
NPM1 (S254)
NPM1 (T237)
BRCA1 (S1280)
DTL (S446)
BLM (S28)
HMGN1 (S8)
UHRF1 (S89)
ERCC6 (Y1179)
MCM3 (S681)
TICRR (S1074)
FANCI (T892)
FANCE (S194)
RIF1 (S1046)
NPM1 (T234)
PCNA (S261)
SETX (S956)
MDC1 (S1068)
BRCA1 (S1174)
−log10(FDR)
log2(fold change)
G
mild RS
moderate RS
strong RS
% difference
H
RS/DMSO
mild
moderate
strong
Kinase activity (z-score)
ATM
ATR
CDK2
PKMYT1
WEE1
CDK1
MST1R
AURKB
TTK
AURKA
PLK1
MAPK1
MAPK3
NEK6
NEK7
CLK1
PRKDC
RPS6KA2
ABL2
CDK5
CSNK2A1
NEK2
MAP3K8
TLK1
BCKDK
DYRK2
MTOR
PRKCA
PRKCE
PRKCG
MET
BMX
MAPK13

Figure 6: Phosphoproteomics of the dose-dependent response to replication stress.

A) (Left) Experimental outline of TMT phosphoproteomics after RS induction. (Right) Number of high-quality phosphorylation sites after filtering. **B)** High-quality sites categorized into up-, down- or non-regulated based on the fold change and false discovery rate (FDR < 0.05). **C)** GO biological processes of sites regulated after mild RS (ViseaGO package). **D)** GO biological processes of upregulated sites after moderate and strong RS (ViseaGO package). **E)** Volcano plot with sites significantly regulated after mild RS (Tier 2 sites). Sites mapping to GO-BP "DNA replication" are highlighted. **F)** Volcano plot with sites significantly regulated after moderate RS (Tier 1 sites). Sites mapping to GO-BP "DNA repair" are highlighted. **G)** Motif analysis of upregulated sites using iceLogo. **H)** Computational prediction of kinase activity using KSEA.

Analysis of individual phospho-sites affected by mild RS revealed increased phosphorylation of multiple DNA replication proteins (Figure 6E). This included the replicative helicase component MCM2, DNA polymerase α subunit POLA2, single-strand break repair ligase LIG3, and origin firing factors CDC6 and Treslin (TICRR). Additionally, ribonucleotide reductase subunits RRM1 and RRM2 were affected, possibly reflecting changes in dNTP production rates upon RS. Among most strongly phosphorylated substrates after moderate RS were the cyclin-dependent kinases CDK1 and CDK2 at T14/Y15 (Figure 6F). These phosphorylation events inhibit CDK activity, and are strongly indicative of DNA damage-induced cell cycle arrest. Moreover, under moderate RS, phosphorylation of multiple DNA repair and fork stability proteins was apparent, such as of the Fanconi anemia pathway components FANCI and FANCE, DSB factors BRCA1 and MDC1, structure specific endonuclease SLX4, and DSB pathway choice regulators T53BP1 and RIF1 (Figure 6F). This opens up opportunities for further mechanistic study of these phosphorylation events.

Phosphorylation motif analysis showed enrichment of the S/T-Q motif after moderate and strong RS, indicating activation of the PIKK family of kinases, including the DSB kinase ATM and the RS response kinase ATR[254] (Figure 6G). Strong RS reduced phosphorylation at phospho-residues followed by a proline (P at +1 position), which we interpret as further indication of the CDK inhibition[255]. To investigate how the global kinase response is affected by different RS levels, we performed computational prediction of kinase activity using Kinase-Substrate Enrichment Analysis (KSEA) (Figure 6H). This

confirmed the stepwise activation of ATM and ATR, as well as a stepwise reduction of cell cycle-promoting CDKs and mitosis-promoting Aurora A/B and PLK1[256]. Interestingly, CDK1 was not predicted as downregulated after mild RS, potentially explaining the lack of cell cycle arrest in this condition.

3.3. Similarities in signaling between Cyclin E- and aphidicolin-induced mild replication stress

Although aphidicolin and hydroxyurea are commonly used as tools to study RS, they are not the most physiological source of RS. Cancer-relevant RS occurs as a consequence of increased expression of oncogenes such as MYC, KRAS or Cyclin E[89]. An improved overview of the signaling induced by oncogenes would benefit our understanding of the pathways relevant for tumorigenesis. We therefore sought to compare APH- and Cyclin E1-induced phosphoproteomes. For this, we developed and made use of the U2OS cell line expressing the CCNE1 gene in a doxycycline-inducible manner (Figure 7A). TMT-based phosphoproteomics after a 24-h treatment with doxycycline (Figure 7B) revealed over 100 sites each that were either upregulated of downregulated upon CCNE1 overexpression (p-value <0.01). Phosphorylation of CCNE1 was strongly increased, acting as a positive control (Figure 7C). Y19 on CDK1/2/3 was more phosphorylated after CCNE1 induction, suggesting changes in CDK activity. Interestingly, APH and doxycycline treatment correlated very well (Pearson correlation coefficient = 0.61), revealing similarities in signaling (Figure 7D). The overlap was most noticeable among downregulated sites, which were enriched for RNA processing factors.

Kinase predictions after CCNE1 overexpression and APH-induced mild RS revealed further similarities at the signaling level (Figure 7E). While it was expected that Cyclin E increases CDK activity, in particular CDK2[257], we observed a similar response after APH treatment. This suggests that after mild RS, CDK activity is in general not diminished, in line with the observation that cells do not get arrested in S or G2 (Figure 5C). Kinase prediction also retrieved plethora of potential kinases with less activity after RS, allowing further exploration.

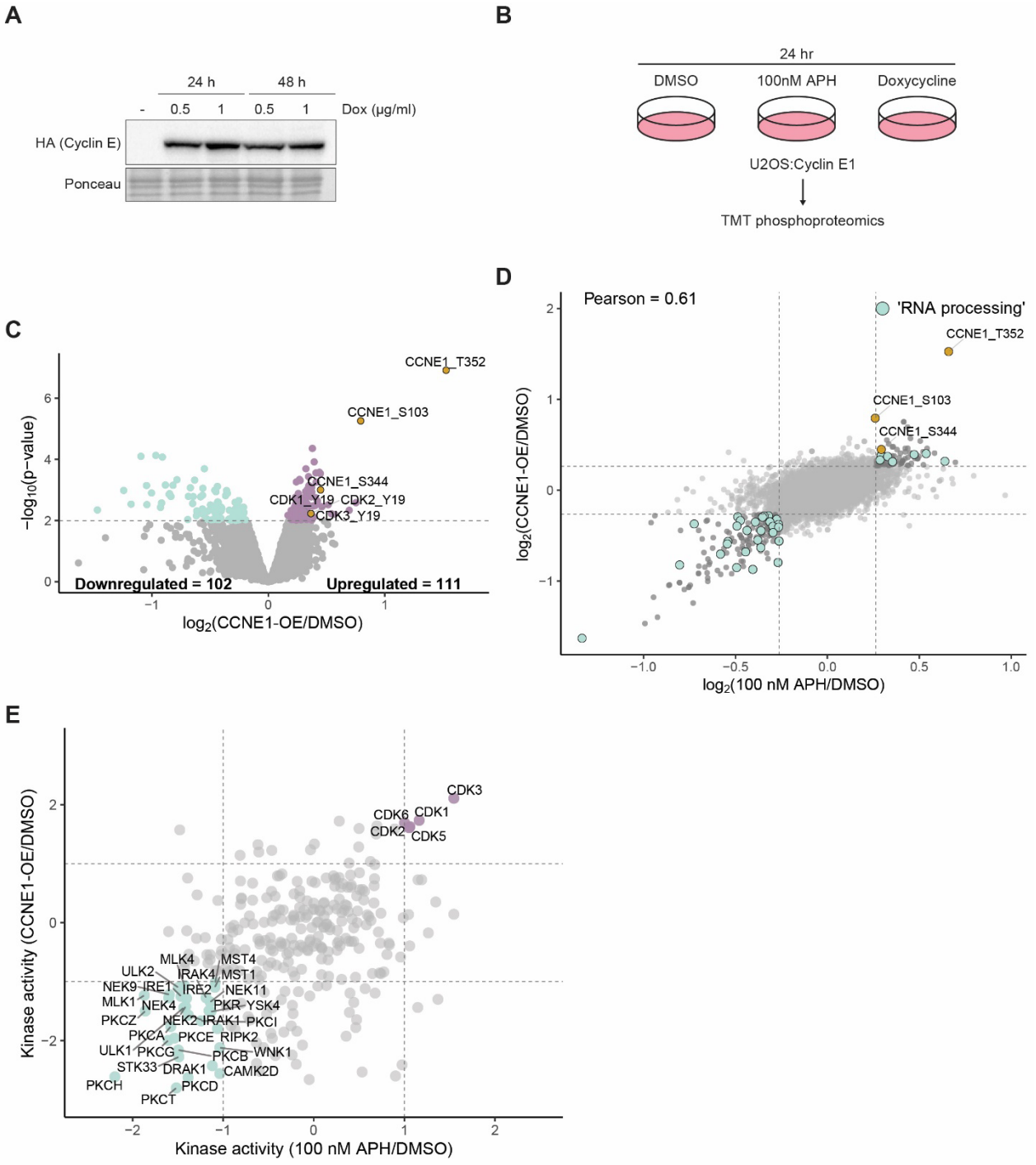


Figure 7: Similarities in signaling between Cyclin E- and aphidicolin-induced mild replication stress.

A) Western blot analysis of the U2OS:Cyclin E1 cell line treated for 24h or 48h with indicated doxycycline (dox) concentrations. Ponceau is used as a loading control. **B)** Treatment scheme of TMT phosphoproteomics upon APH-induced mild RS or doxycycline-mediated Cycline E1 overexpression. **C)** Volcano plot with sites significantly regulated (p-value < 0.01) after Cyclin E1 overexpression. The number of significantly up-

and downregulated sites is indicated at the bottom. Selected sites are highlighted. **D)** Scatter plot comparing APH treatment and Cyclin E1 overexpression. RNA processing factors are highlighted. Pearson correlation coefficient is indicated. **E)** Scatter plot comparing kinase activity predictions using the Kinase Library after indicated treatments. Kinases predicted to be commonly regulated are highlighted.

3.4. Analysis of replication stress-induced changes in phosphorylation of R-loop regulators reveals regulation of RNase H1 and Senataxin by phosphorylation

RS has a two-way relationship with R-loops. Namely, aberrant accumulation of R-loops increases RS, and vice versa. However, evidence for the latter is mostly based on the frequently criticized immunofluorescence using the S9.6 antibody[217,258,259]. To strengthen the notion that APH increases R-loop levels, we used a tool recently developed in our lab called the HBD-AF647 (Figure 8A). The tool utilizes purified hybrid binding domain of the human RNase H1 covalently linked to a fluorophore (Alexa Fluor 647), enabling fluorescent imaging of RNA:DNA hybrids and R-loops in fixed cells without the need for a secondary antibody (developed by redacted). Chronic RS (24 h), including mild RS, significantly increased the nuclear HBD-AF647 staining (Figure 8B). As expected, depletion of the RNA:DNA hybrid nuclease RNase H1, as well as the helicase AQR[145], also considerably increased nuclear hybrid levels.

While a large number of R-loop-modulating helicases, nucleases, translocases and other R-loop binders have been characterized to date, their regulation by PTMs upon stresses that cause R-loop accumulation is just beginning to emerge[260]. To find out if R-loop factors are regulated by phosphorylation upon RS, we overlapped our phosphoproteomics data with a dataset of 186 high-confidence R-loop regulators reported in the R-loopBase[261] (Figure 8C). This revealed that phosphorylation of 36 R-loop regulators at 91 sites is significantly affected by at least one RS treatment (Figure 8D). Among them, we found multiple well-studied DDR proteins with R-loop-related and -unrelated functions, such as BRCA1 (13 sites), MDC1 (10 sites), FANCI (4 sites), FANCD2 (1 site), BLM (2 sites) and RPA1 (2 sites). We noticed that proteins with a loss of phosphorylation, such as HNRNPD, NPM1, NCL, DDX5, ADAR, SAFB2, PAF1 and NONO, were enriched for transcription-

and RNA processing-related functions (Figure 8D). Critically, we found increased phosphorylation of a bone fide R-loop helicase Senataxin/SETX[125,262] at 5 serine residues (S947, S950, T952, S956 and S1345), four of which are SQ sites, and decreased phosphorylation of one of the major R-loop nucleases RNase H1 at three serine residues (S74, S76 and S80). Interestingly, the reported phospho-serines on RNase H1 are located in an intrinsically disordered region immediately downstream of its hybrid-binding domain (HBD) (Figure 8E), prompting us to investigate their function.

Sequence comparisons from yeast to humans showed that HBD is universally followed by an array of phosphorylatable residues, most often serines, which we termed the phospho-cluster region (Figure 8F). Out of the three serines present in humans, S76 and S80 were the best conserved. Aiming to find the kinase phosphorylating RNase H1, we explored its interactome. To this end, we created a U2OS cell line with endogenously tagged RNase H1 fused with APEX and FLAG tags using CRISPR-Cas9. FLAG-based affinity purification coupled to LC-MS/MS revealed RNase H1 as the top hit (Figure 8G). In line with the role of RNase H1 in RNA processing, subunits of RNA Pol II[263], the transcription elongation PAF1 complex, multiple histones and different R-loop-related DDX/DHX helicases were significantly enriched. Among enriched kinases, we found transcription- and splicing-related CDK11B and CDK12, as well as two subunits of the CK2 holoenzyme (CSNK2A1 and CSNK2B). S80 of RNase H1 is surrounded by three negatively charged residues (E78, E81 and E84) (Figure 8F), which resembles the motif necessary for hierarchical phosphorylation by CK2[8], prompting us to investigate it as a candidate kinase. Phosphoproteomics upon chemical inhibition of CK2 using CX-4945[264] revealed that phosphorylation of RNase H1 at S74/S76 is strongly affected, even more so than after APH (Figure 8H). However, a closer inspection of the RS- and CX-4945-dependent substrates showed that two phospho-sites of another kinase, DYRK1A, are strongly affected by both treatments (Figure 8I). Additionally, a recent report highlighted DYRK1A as an off-target of CX-4945[265].

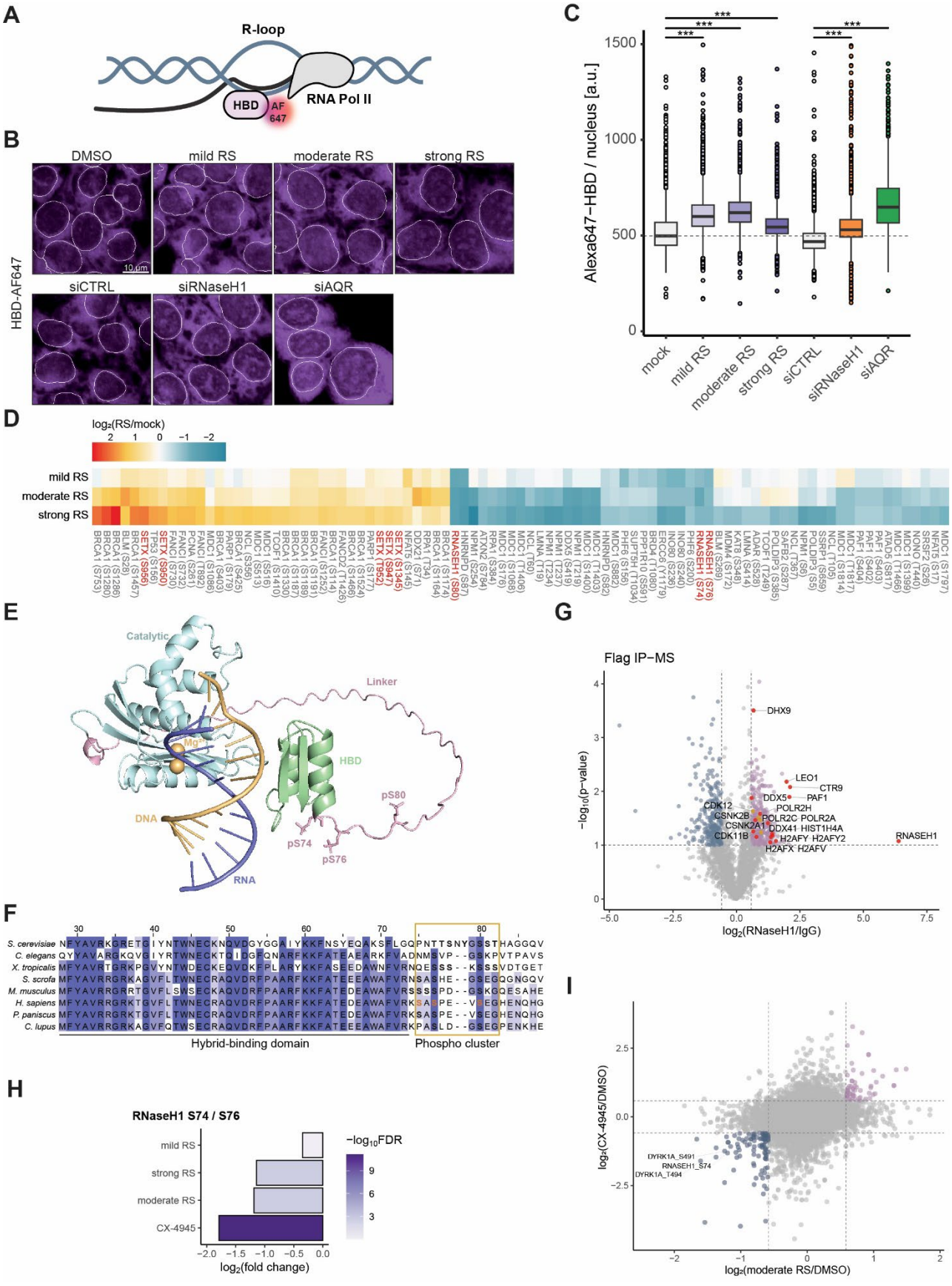

A
R-loop
HBD
AF 647
RNA Pol II
B
DMSO
mild RS
moderate RS
strong RS
HBD-AF647
10 µm
siCTRL
siRNaseH1
siAQR
C
Alexa647−HBD / nucleus [a.u.]
1500
1000
500
0
***
mock
mild RS
moderate RS
strong RS
siCTRL
siRNaseH1
siAQR
D
log2(RS/mock)
2
1
0
−1
−2
mild RS
moderate RS
strong RS
BRCA1 (S753)
BRCA1 (S1280)
BRCA1 (S1286)
BLM (S28)
BRCA1 (S1457)
SETX (S956)
SETX (S960)
FANCI (S730)
FANCI (S732)
PCNA (S261)
FANCI (T882)
MDC1 (S1086)
BRCA1 (S403)
PARP1 (S179)
BRCA1 (S425)
MDC1 (S515)
NCL (S356)
MDC1 (S513)
TCOF1 (S1410)
BRCA1 (S1330)
BRCA1 (S1187)
BRCA1 (S1189)
BRCA1 (S1191)
FANCI (S1252)
BRCA1 (S114)
FANCD2 (T1426)
BRCA1 (S1466)
BRCA1 (S1524)
PARP1 (S762)
BRCA1 (S1174)
NFAT5 (S145)
DDX21 (S71)
RPA1 (T34)
BRCA1 (S1164)
RNASEH1 (S80)
HNRNPD (S87)
RPA1 (S384)
ATXN2 (S734)
RPA1 (S268)
MDC1 (S176)
MDC1 (T1406)
MDC1 (S1068)
MDC1 (T1403)
MDC1 (S1400)
DDX21 (S219)
NPM1 (T237)
NPM1 (T234)
LMNA (T19)
NCL (T60)
MDC1 (S882)
HNRNPD (S82)
MDC1 (S1403)
SUPT5H (T1034)
PHF6 (S156)
PPP1R10 (S591)
BRD4 (T1080)
ERCC6 (Y1179)
INO80 (S240)
PHF6 (S200)
RNASEH1 (S76)
RNASEH1 (S74)
BLM (S289)
MDM4 (S172)
KAT8 (S348)
LMNA (S414)
ADAR (S528)
TCOF1 (S1385)
SAFB2 (S287)
NCL (T367)
NPM1 (S6)
SSRP1 (S659)
POLDIP3 (S5)
NCL (T105)
MDC1 (S1814)
MDC1 (T1817)
PAF1 (S402)
PAF1 (S404)
ATAD5 (S817)
MDC1 (T1466)
MDC1 (S1399)
NONO (T440)
SRSF2 (S26)
NFAT5 (S17)
MDC1 (S1797)
E
Catalytic
Linker
Mg2+
HBD
pS80
DNA
pS74
pS76
RNA
F
S. cerevisiae
C. elegans
X. tropicalis
S. scrofa
M. musculus
H. sapiens
P. paniscus
C. lupus
Hybrid-binding domain
Phospho cluster
G
Flag IP−MS
−log10(p−value)
DHX9
LEO1
CTR9
DDX5
PAF1
CDK12
POLR2H
CSNK2B
POLR2C POLR2A
CSNK2A1
DDX41 HIST1H4A
CDK11B
H2AFY H2AFY2
H2AFX H2AFV
RNASEH1
log2(RNaseH1/IgG)
H
RNaseH1 S74 / S76
mild RS
strong RS
moderate RS
CX-4945
−log10FDR
log2(fold change)
I
log2(CX-4945/DMSO)
DYRK1A_S491
RNASEH1_S74
DYRK1A_T494
log2(moderate RS/DMSO)

Figure 8: Analysis of replication stress-induced changes in phosphorylation of R-loop regulators reveals modification of RNase H1 and Senataxin by phosphorylation.

A) Scheme of the fluorescence-based RNA:DNA hybrid sensor HBD-AF647**. B)** Representative images of HBD-AF647 staining in fixed HCT116 after the indicated treatments. Nuclei are outlined in white. **C)** Quantification of immunofluorescence from the experiment shown in B. Unpaired, two-sided t-test with Bonferroni correction**,** ***: $p \leq 0.001$**. D)** Heatmap representing the overlap of the sites from the TMT phosphoproteomics (**Figure 6**) with the R-loopBase. Hits with at least one significantly changing treatment are shown. The scale bar represents log_2-transformed fold changes. **E)** AlphaFold3-modeled interaction of the M27-RNase H1 with a 12-mer hybrid substrate rendered in PyMol. Domains are highlighted and labeled in different colors. Phospho-sites of interest are indicated. Two Mg^{2+} ions positioned in the active site are shown. **F)** Multiple sequence alignment containing the HBD and the phospho-cluster region of RNase H1 in selected species**.** Phosphorylatable residues within the cluster are shown in bold. Sites phosphorylated in humans are shown in orange. **G)** Flag-based immunoprecipitation in the U2OS:RNaseH1-FLAG-APEX2 cell line. Selected hits are highlighted in red, enriched kinases in yellow. **H)** Barplot showing regulation of the RNase H1 phospho-site containing S74 and S76 in HCT116 after treatment with RS or CX-4945. **I)** Scatter plot comparing moderate RS to the CX-4945 treatment. Selected sites are indicated.

3.5. Replication stress-mediated reduction in DYRK1A signaling leads to a loss of RNase H1 phosphorylation

We therefore hypothesized that DYRK1A, rather than CK2, phosphorylates RNase H1. To facilitate the kinase search and further mechanistic exploration of RNase H1 phosphorylation, we developed two new tools. Using lentiviral transduction, we stably integrated a doxycycline (dox)-inducible construct bearing the GFP-tagged nuclear isoform of RNase H1, including the triple phospho-dead (S74A, S76A and S80A; abbreviated as S3xA) and the triple phospho-mimicking (S74D, S76D, S80D; abbreviated as S3xD) version into the HCT116 cell line (Figure 9A). Furthermore, we developed a custom phosphorylation-specific RNase H1 antibody raised against a peptide mono-phosphorylated at S74. Although the antibody displayed unspecific, knockdown-resistant bands at 30 kDa in inputs and pulldowns of endogenous RNase H1 (not shown), the phosphorylation was readily detectable in GFP-based pulldowns using the dox-inducible RNase H1 cell line (Figure 9B). Critically, the phospho-S74 signal was dox-dependent, and not visible upon expression of the S3xA nor the S3xD mutant of RNase H1, indicating that the antibody indeed recognizes the phospho-group on one of the three serines of

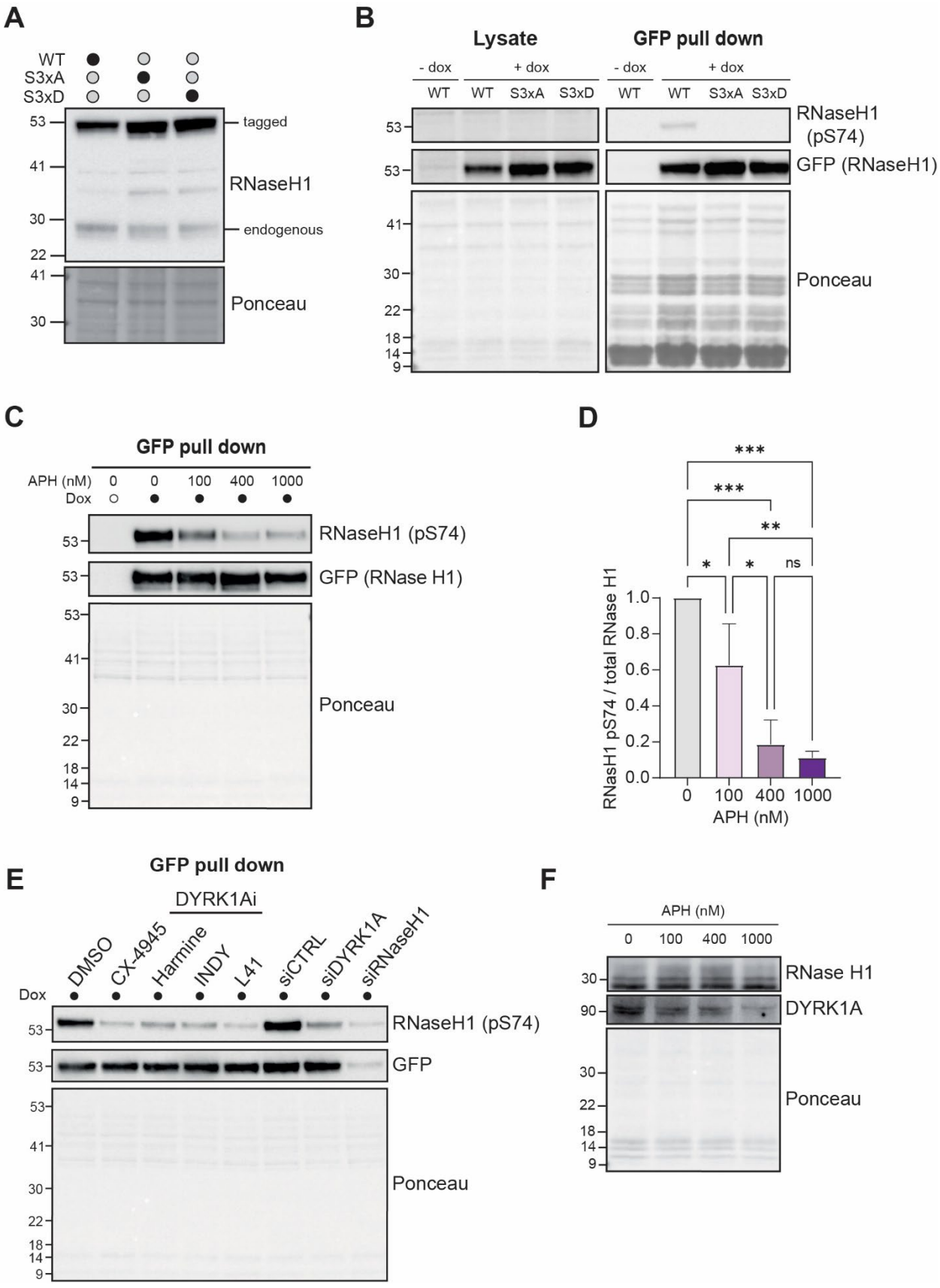

A
WT
S3xA
S3xD
53
41
30
22
tagged
RNaseH1
endogenous
Ponceau
B
Lysate
GFP pull down
- dox
+ dox
WT
S3xA
S3xD
RNaseH1 (pS74)
GFP (RNaseH1)
Ponceau
C
GFP pull down
APH (nM)
Dox
RNaseH1 (pS74)
GFP (RNase H1)
Ponceau
D
RNasH1 pS74 / total RNase H1
APH (nM)
ns
E
GFP pull down
DYRK1Ai
DMSO
CX-4945
Harmine
INDY
L41
siCTRL
siDYRK1A
siRNaseH1
Dox
RNaseH1 (pS74)
GFP
Ponceau
F
APH (nM)
RNase H1
DYRK1A
Ponceau

Figure 9: Replication stress-mediated reduction in DYRK1A signaling leads to a loss of RNase H1 phosphorylation.

A) Western blot showing expression of the GFP-tagged RNase H1 phospho-mutants in HCT116 after a 24h treatment with 2 µg/ml doxycycline. Ponceau: loading control. **B)** Western blot validating the antibody specific for pS74 of RNase H1 in GFP-based pulldowns. Ponceau: loading control. **C)** Western blot showing dose-dependent decrease of RNase H1 pS74 after a 24h treatment with APH. Ponceau: loading control. **D)** Quantification of results from C in 3 biological replicates (*: $p \leq 0.05$, **: $p \leq 0.01$; ***, $p \leq 0.001$). **E)** Western blot of GFP-based pulldowns showing effects of DYRK1A inhibition of knockdown on the levels of pS74. Ponceau: loading control. **F)** Western blot showing the effects of different doses of APH after a 24h-treatment on total levels of DYRK1A. Ponceau: loading control.

interest. The antibody-based approach confirmed that RS decreases phosphorylation of RNase H1 in a dose-dependent manner, while total levels of GFP-tagged RNase H1 remain unchanged (Figure 9C,D).

Strikingly, inhibition of DYRK1A using three different chemical inhibitors (Leucettine L41[266], INDY[267] and harmine[268]), as well as CX-4945, almost completely reversed phosphorylation at S74, while depletion of DYRK1A using siRNA reduced it (Figure 9E). Additionally, total levels of DYRK1A decreased with the RS dose, offering an explanation for reduced phosphorylation of RNase H1 upon RS (Figure 9F).

3.6. Reduction in DYRK1A signaling is a driver of the replication stress-induced phosphorylation loss on RNA processing factors

Assuming RNase H1 is not the only RS-regulated DYRK1A target, we next asked whether its other substrates are decreased in phosphorylation upon RS. To this end, we performed phosphoproteomics after strong RS, chemical inhibition of DYRK1A using L41 or INDY, or depletion of DYRK1A (Figure 10A). Principal component analysis (PCA) showed a clustering of the DYRK1A inhibitors, and their similarity with depletion of DYRK1A (Figure 10B). As expected, APH treatment displayed a distinct response. Pairwise comparison of the strong RS with either INDY or L41 revealed a subset of shared substrates (Figure 10C,D). Critically, both RS and inhibition of DYRK1A led to a reduction of its auto-phosphorylated activity-promoting residue S529[269], strengthening the observation that RS

reduces DYRK1A signaling. Overlapping the downregulated phospho-sites dependent on RS, INDY and L41 revealed a subset of 27 substrates (Figure 10E). Among them were proteins like the DHX/DDX family RNA helicases DHX38 and DDX21, spliceosome B factor SF3B1 and alternative splicing factors RBMX and SAFB (Figure 10F). Kinase prediction using KEA2 confirmed that DYRK1A activity is reduced both after RS and DYRK1A inhibition (Figure 10G). In addition, it revealed pathways common to DYRK1A inhibitors, such as increased CDK2 and GSK3B activity, which merits further investigation. GO terms analysis revealed that targets with a loss of phosphorylation after DYRK1A inhibition or depletion are predominantly involved in RNA processing and splicing (Figure 10H). Finally, by overlapping sites regulated after DYRK1A inhibition or depletion, we generated a high-confidence set of DYRK1A targets (Figure 10I). Altogether, our results suggest that the reduction in DYRK1A signaling is one of the drivers of the previously observed loss of phosphorylation of RNA processing factors upon RS.

3.7. Functional implications of RNase H1 phosphorylation

We hypothesized that DYRK1A-mediated phosphorylation of RNase H1 could impact its biological role in the following ways: it could change the propensity of the HBD to bind RNA:DNA hybrids, change RNase H1 localization, chromatin binding, activity, localization, or protein-protein interactions. To test the first hypothesis, we turned to *in vitro* nucleic acid-protein binding assays (Figure 11A). To this end, we purified phospho-WT or phospho-substituent catalytically inactive (D210N) RNase H1 lacking the mitochondrial localization signal (Figure 11B). Using surface plasmon resonance and a 12-mer RNA:DNA hybrid substrate, we showed that the S3xD mutant has a dissociation constant (K_D) 1.8-fold higher than either WT or S3xA mutant, suggesting decreased hybrid binding (Figure 11C). Closer inspection of the binding dynamics revealed that S3xD associates with hybrids less efficiently, but is more prone to being retained on the hybrid upon washout (Figure 11D). As a specificity control, we compared binding of RNase H1 to dsDNA, confirming that RNase H1 binds to hybrids with much higher efficiency than to

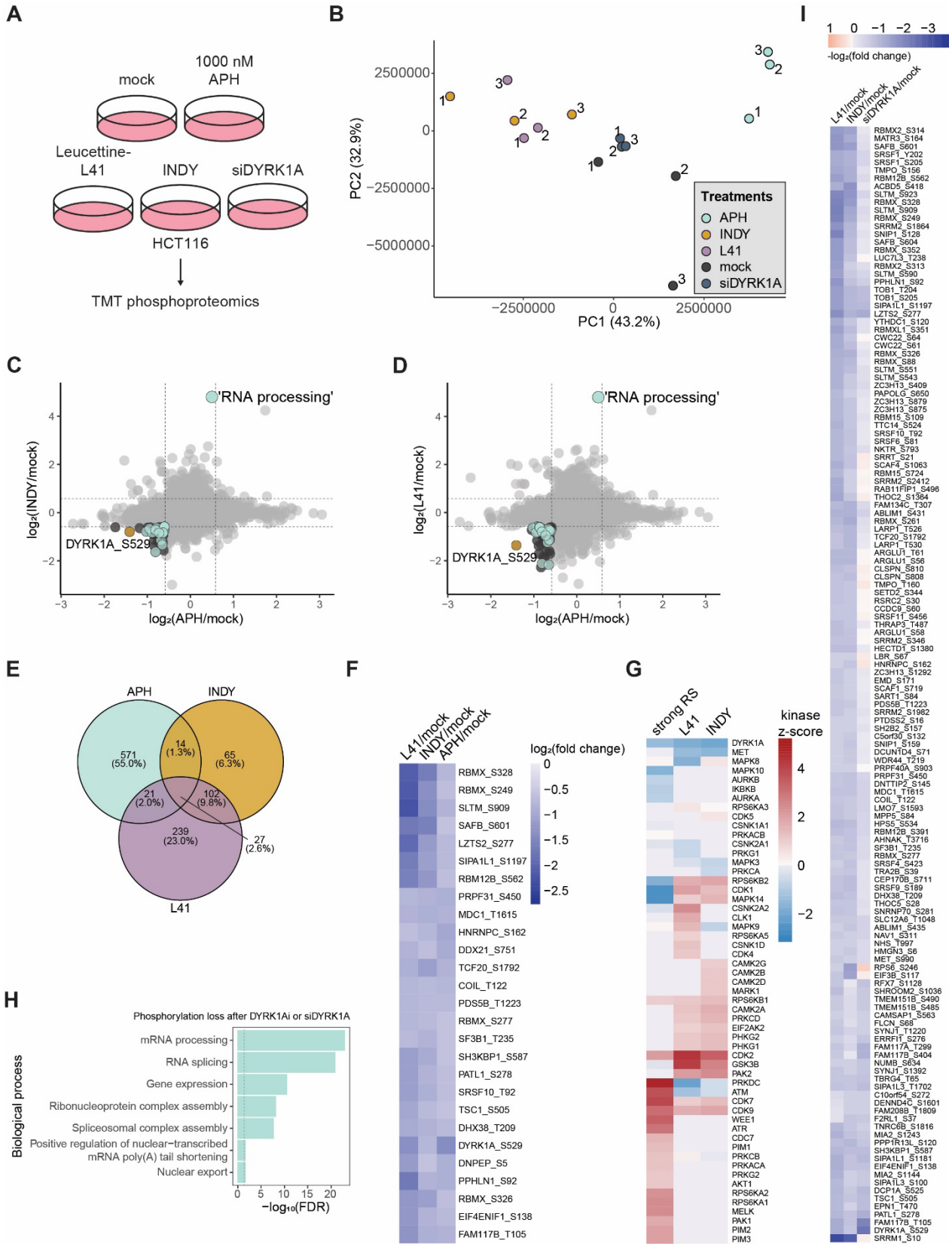

A
mock
1000 nM APH
Leucettine-L41
INDY
siDYRK1A
HCT116
TMT phosphoproteomics
B
PC2 (32.9%)
PC1 (43.2%)
Treatments
APH
INDY
L41
mock
siDYRK1A
C
'RNA processing'
log2(INDY/mock)
log2(APH/mock)
DYRK1A_S529
D
'RNA processing'
log2(L41/mock)
log2(APH/mock)
DYRK1A_S529
E
APH
INDY
L41
571 (55.0%)
14 (1.3%)
65 (6.3%)
21 (2.0%)
102 (9.8%)
239 (23.0%)
27 (2.6%)
F
L41/mock
INDY/mock
APH/mock
log2(fold change)
RBMX_S328
RBMX_S249
SLTM_S909
SAFB_S601
LZTS2_S277
SIPA1L1_S1197
RBM12B_S562
PRPF31_S450
MDC1_T1615
HNRNPC_S162
DDX21_S751
TCF20_S1792
COIL_T122
PDS5B_T1223
RBMX_S277
SF3B1_T235
SH3KBP1_S587
PATL1_S278
SRSF10_T92
TSC1_S505
DHX38_T209
DYRK1A_S529
DNPEP_S5
PPHLN1_S92
RBMX_S326
EIF4ENIF1_S138
FAM117B_T105
G
strong RS
L41
INDY
kinase z-score
DYRK1A
MET
MAPK8
MAPK10
AURKB
IKBKB
AURKA
RPS6KA3
CDK5
CSNK1A1
PRKACB
CSNK2A1
PRKG1
MAPK3
PRKCA
RPS6KB2
CDK1
MAPK14
CSNK2A2
CLK1
MAPK9
RPS6KA5
CSNK1D
CDK4
CAMK2G
CAMK2B
CAMK2D
MARK1
RPS6KB1
CAMK2A
PRKCD
EIF2AK2
PHKG2
PHKG1
CDK2
GSK3B
PAK2
PRKDC
ATM
CDK7
CDK9
WEE1
ATR
CDC7
PIM1
PRKCB
PRKACA
PRKG2
AKT1
RPS6KA2
RPS6KA1
MELK
PKN1
PIM2
PIM3
H
Phosphorylation loss after DYRK1Ai or siDYRK1A
Biological process
mRNA processing
RNA splicing
Gene expression
Ribonucleoprotein complex assembly
Spliceosomal complex assembly
Positive regulation of nuclear−transcribed mRNA poly(A) tail shortening
Nuclear export
−log10(FDR)
I
-log2(fold change)
L41/mock
INDY/mock
siDYRK1A/mock
RBMX2_S314
MATR3_S164
SAFB_S601
SRSF1_Y202
SRSF1_S205
TMPO_S156
RBM12B_S562
ACBD5_S418
SLTM_S923
RBMX_S328
SLTM_S909
RBMX_S249
SRRM2_S1864
SNIP1_S128
SAFB_S604
RBMX_S352
LUC7L3_T238
RBMX2_S313
SLTM_S590
PPHLN1_S92
TOB1_T204
TOB1_S205
SIPA1L1_S1197
LZTS2_S277
YTHDC1_S120
RBMXL1_S351
CWC22_S64
CWC22_S61
RBMX_S326
RBMX_S88
SLTM_S551
SLTM_S543
ZC3H13_S409
PAPOLG_S650
ZC3H13_S879
ZC3H13_S875
RBM15_S109
TTC14_S524
SRSF10_T92
SRSF6_S81
NKTR_S793
SRRT_S21
SCAF4_S1063
RBM15_S724
SRRM2_S2412
RAB11FIP1_S496
THOC2_S1364
FAM134C_T307
ABLIM1_S431
RBMX_S261
LARP1_T526
TCF20_S1792
LARP1_T530
ARGLU1_T61
ARGLU1_S56
CLSPN_S810
CLSPN_S808
TMPO_T160
SETD2_S344
RSRC2_S30
CCDC9_S60
SRSF11_S456
THRAP3_T487
ARGLU1_S58
SRRM2_S346
HECTD1_S1380
LBR_S67
HNRNPC_S162
ZC3H13_S1292
EMD_S171
SCAF1_S719
SART1_S84
PDS5B_T1223
SRRM2_S1982
PTDSS2_S16
SH2B2_S157
C5orf30_S132
SNIP1_S159
DCUN1D4_S71
WDR44_T219
PRPF40A_S903
PRPF31_S450
DNTTIP2_S145
MDC1_T1615
COIL_T122
LMO7_S1593
MPP5_S84
HPS5_S534
RBM12B_S391
AHNAK_T3716
SF3B1_T235
RBMX_S277
SRSF4_S423
TRA2B_S39
CEP170B_S711
SRSF9_S189
DHX38_T209
THOC5_S28
SNRNP70_S281
SLC12A6_T1048
ABLIM1_S435
NAV1_S311
NHS_T997
HMGN3_S6
MET_S990
RPS6_S246
EIF3B_S117
RFX7_S1128
SHROOM2_S1036
TMEM151B_S490
TMEM151B_S485
CAMSAP1_S563
FLCN_S68
SYNJ1_T1220
ERRFI1_S276
FAM117A_T299
FAM117B_S404
NUMB_S634
SYNJ1_S1392
TBRG4_T65
SIPA1L3_T1702
C10orf54_S272
DENND1C_S1601
FAM208B_T1809
F2RL1_S37
TNRC6B_S1816
MIA2_S143
PPP1R13L_S120
SH3KBP1_S587
SIPA1L1_S1181
EIF4ENIF1_S138
MIA2_S1144
SIPA1L3_S100
DCP1A_S525
TSC1_S505
EPN1_T470
PATL1_S278
FAM117B_T105
DYRK1A_S529
SRRM1_S10

Figure 10: Reduction in DYRK1A signaling is a driver of the replication stress-induced phosphorylation loss on RNA processing factors.

A) Scheme of the TMT phosphoproteomics after indicated treatments in HCT116. **B)** Principal component analysis (PCA) of the TMT intensities corresponding to replicates of different treatments. Replicate number is indicated. **C)** Scatter plot comparing phosphoproteomes after strong RS (APH) and INDY. RNA processing factors are highlighted. **D)** Scatter plot comparing phosphoproteomes after strong RS (APH) and L41. RNA processing factors are highlighted. **E**) Venn diagram showing overlap of the sites downregulated (FC < 0.67) after indicated treatments. **F)** Heatmap showing regulation of 27 substrates with shared downregulation from E. **G)** Heatmap showing kinase activity predictions (KEA2) after indicated treatments. **H)** GO terms analysis of phospho-substrates commonly downregulated after L41, INDY and siDYRK1A. Substrates downregulated in at least 2 conditions are shown. **I**) Heatmap showing phospho-substrates analyzed in H.

dsDNA. Interestingly, S3xD also showed reduced binding to dsDNA, likely because RNase H1 binds dsDNA and hybrids in a similar way. Next, we used a live-cell fluorescent microscopy approach to investigate cellular localization of the phospho-mutants. While phospho-WT displayed an overall similar nuclear distribution to S3xA and S3xD, we noticed that the S3xA mutant displays a clear separation between the nucleolar and nucleoplasmic compartments, whereas this separation is impaired in the S3xD mutant (Figure 11E). This was further substantiated by the quantification of the coefficient of variation (CV) of nuclear intensity (Figure 11F), a measure which was previously used to study the impact of phosphorylation on protein distribution in the nuclear sub-compartments[270]. The S3xD mutant had a significantly reduced CV% in comparison to the S3xA, with WT being in between, likely because it is a mixture of the non-phosphorylated and phosphorylated species.

To test the impact of phosphorylation on chromatin levels, we monitored the amount of RNase H1 on chromatin and in the soluble fraction throughout a timecourse of a strong RS treatment (Figure 11G). While the chromatin levels of RNase H1 did not change throughout the timecouse, the levels of soluble RNase H1 slightly decreased at later timepoints, particularly at 24 h. This was preceded by the appearance of a faint higher-molecular weight band, possibly corresponding to ubiquitylated RNase H1. Therefore, we concluded that phosphorylation of RNase H1 does not influence the total levels of RNase H1 on chromatin, but it might be involved in the regulation of its stability.

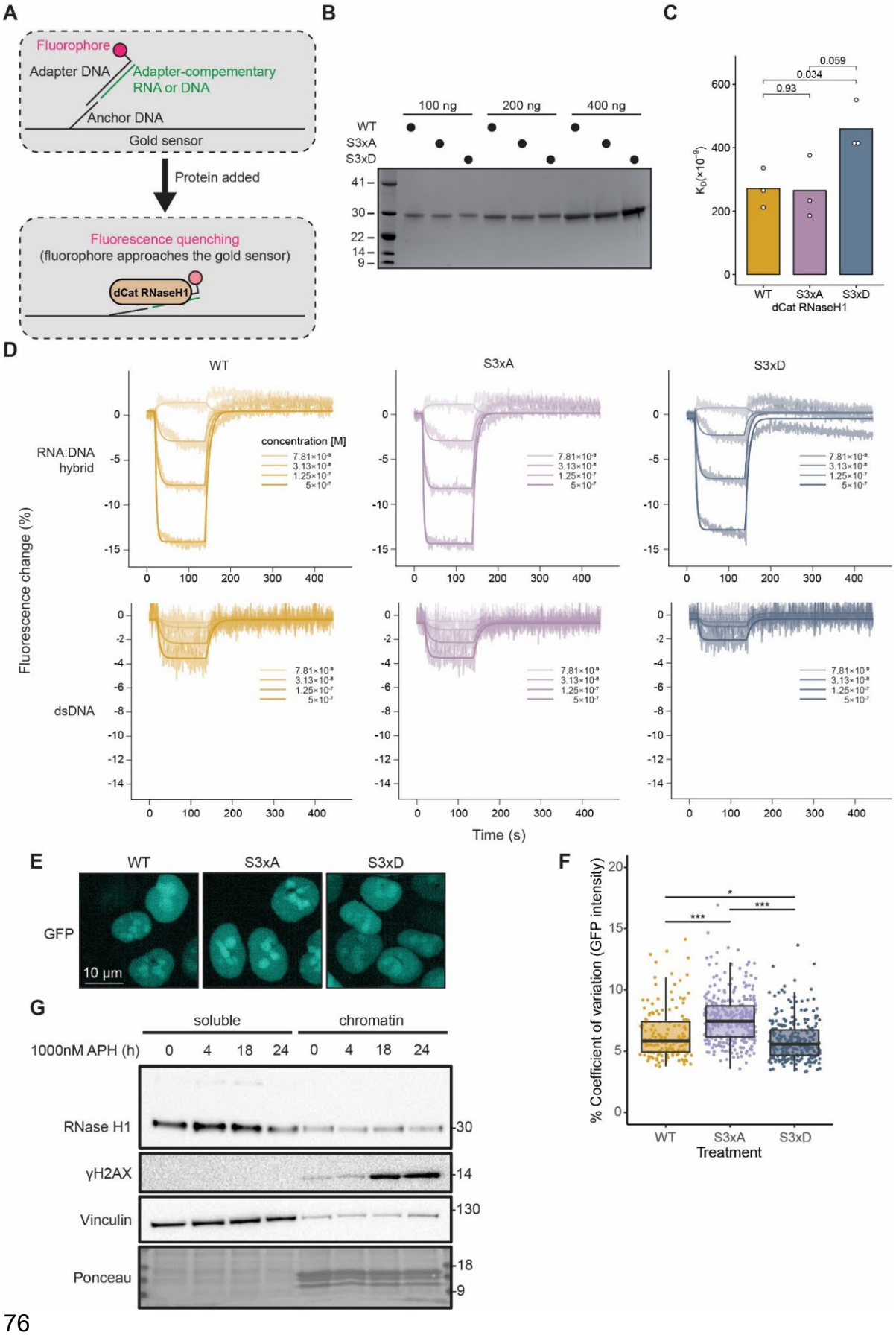

A
Fluorophore
Adapter DNA
Adapter-compementary RNA or DNA
Anchor DNA
Gold sensor
Protein added
Fluorescence quenching
(fluorophore approaches the gold sensor)
dCat RNaseH1
B
100 ng
200 ng
400 ng
WT
S3xA
S3xD
41
30
22
14
9
C
0.059
0.034
0.93
$K_D(\times 10^{-9})$
WT
S3xA
S3xD
dCat RNaseH1
D
WT
S3xA
S3xD
RNA:DNA hybrid
dsDNA
concentration [M]
7.81×10⁻⁹
3.13×10⁻⁸
1.25×10⁻⁷
5×10⁻⁷
Fluorescence change (%)
Time (s)
E
WT
S3xA
S3xD
GFP
10 µm
F
% Coefficient of variation (GFP intensity)
*
***
***
WT
S3xA
S3xD
Treatment
G
soluble
chromatin
1000nM APH (h)
0
4
18
24
0
4
18
24
RNase H1
γH2AX
Vinculin
Ponceau
30
14
130
18
9

Figure 11: Functional implications of RNase H1 phosphorylation.

A) Outline of the HeliX+ assay, a surface plasmon resonance approach to measure nucleic acid-protein binding. **B)** Purified catalytically-inactive (D210N) M27-RNase H1 phospho-mutants loaded at indicated amounts and stained with Colloidal Blue. S3xD shows a size shift as a result of phospho-mimics. **C)** Quantification of the HeliX+ assay measuring K_D of interaction of indicated M27-RNase H1 mutants with a 12-mer RNA:DNA hybrid substrate in triplicates. Unpaired, two-sided Student's t-test with indicated p-values. **D)** Trace plots showing % change in fluorescence upon addition and washout of the dilution series of the M27-RNase H1 mutants. 12-mer hybrid and 26-mer dsDNA substrates were used. **E)** Representative images of live-cell fluorescent microscopy after a 24h doxycycline induction in HCT116:M27-RNase H1-GFP. **F)** Dot plots showing % coefficient of variation of GFP intensity from experiment in E. Boxplots show the median (central line), first and third quartiles (box limits), and 1.5 times the interquartile range (whiskers). Two-sided Mann-Whitney U test (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). **G)** Western blot showing levels of endogenous RNase H1 after RS treatment for specified duration and cellular fractionation.

3.8. Analysis of the phosphorylation-based changes in the RNase H1 interactome reveals a role of GSK3B in RNase H1 stability

To investigate the possibility that phosphorylation of RNase H1 changes its interaction profile, we performed the LC/MS-MS-based interactome analysis after GFP-based pulldowns in HCT116 cells expressing WT or mutant RNase H1 (Figure 12A). Additionally, we treated WT-expressing cells with L41 to trigger the loss of phosphorylation. As expected, RNase H1 was by far the most enriched hit after doxycycline induction in WT cells (Figure 12B). Pairwise comparisons of S3xA-expressing cells and L41-treated cells revealed factors with a preference for binding either phosphorylated or non-phosphorylated RNase H1 (Figure 12C). Multiple interactors were RNA-binding proteins known to reside in the nucleoli, mirroring nucleolar localization of RNase H1. Among them, nucleolin (NCL), FUS and SFPQ preferentially interacted with non-phosphorylated RNase H1. These proteins are involved in condensate formation (phase separation), potentially explaining our observation that S3xA mutant has a better nucleolar-nucleoplasmic separation than S3xD. On the other hand, phosphorylated RNase H1 interacted with a subunit of the rDNA-transcribing RNA Pol I subunit POLR1F and, in

addition, with two factors involved nucleolar transcription at rDNA and its intergenic spacers (IGSs), TOP1[271] and RAD51C[272]. Additionally, among the hits with a preference for phosphorylated RNase H1 was the PRMT5 arginine methyltransferase complex, including the catalytic subunit PRMT5, the non-catalytic component WDR77 and the substrate adaptor RIOK1. This suggests RNase H1 might be methylated in a phosphorylation-dependent manner.

Treatment with L41 led to a strong loss of interaction of RNase H1 with the GSK3β kinase (Figure 12D). This caught our interest, as GSK3β is known to phosphorylate substrates primed by DYRK1A[31]. GSK3β-mediated phosphorylation often introduces a phospho-degron, enabling substrate ubiquitylation by the SCF ubiquitin ligase and subsequent degradation by the proteasome[28]. In line with this, the S3xA mutant showed decreased binding to multiple proteasome subunits (Figure 12E), and had higher overall expression levels than WT (Figure 9A), indicating a potential defect in its ubiquitin-mediated degradation. S3xD mutant interacted with the phospho-degron reader FBXL6 (Figure 12F), which acts together with the SCF ligase[273], offering a possible candidate linking the RNase H1 phospho-cluster with the SCF.

To test if GSK3β regulates total levels of RNase H1, we treated HCT116 cells with a selective GSK3β inhibitor CHIR99021[274]. Prolonged treatment with CHIR99021 led to a strong increase in total levels of RNase H1, and a concomitant increase in pS74 (Figure 12G). This confirmed that GSK3β regulates levels of RNase H1 and that DYRK1A, rather than GSK3β, phosphorylates pS74. To exclude cell type-specific or inhibitor off-target effects, we used mouse embryonic fibroblasts (MEFs) with a biallelic knockout (KO) of GSK3α[275] or GSK3β[276]. While GSK3α KO did not have an effect on RNase H1 stability, GSK3β KO led to a 1.5-fold increase in total levels of RNase H1 (Figure 12H,I). In summary, our results suggest that phosphorylation of RNase H1 affects its interaction landscape and might thus regulate its localization or methylation. Additionally, we find strong evidence that DYRK1A-mediated phosphorylation of RNase H1 serves to recruit GSK3β, which could introduce a phospho-degron and therefore induce proteasomal degradation of RNase H1.

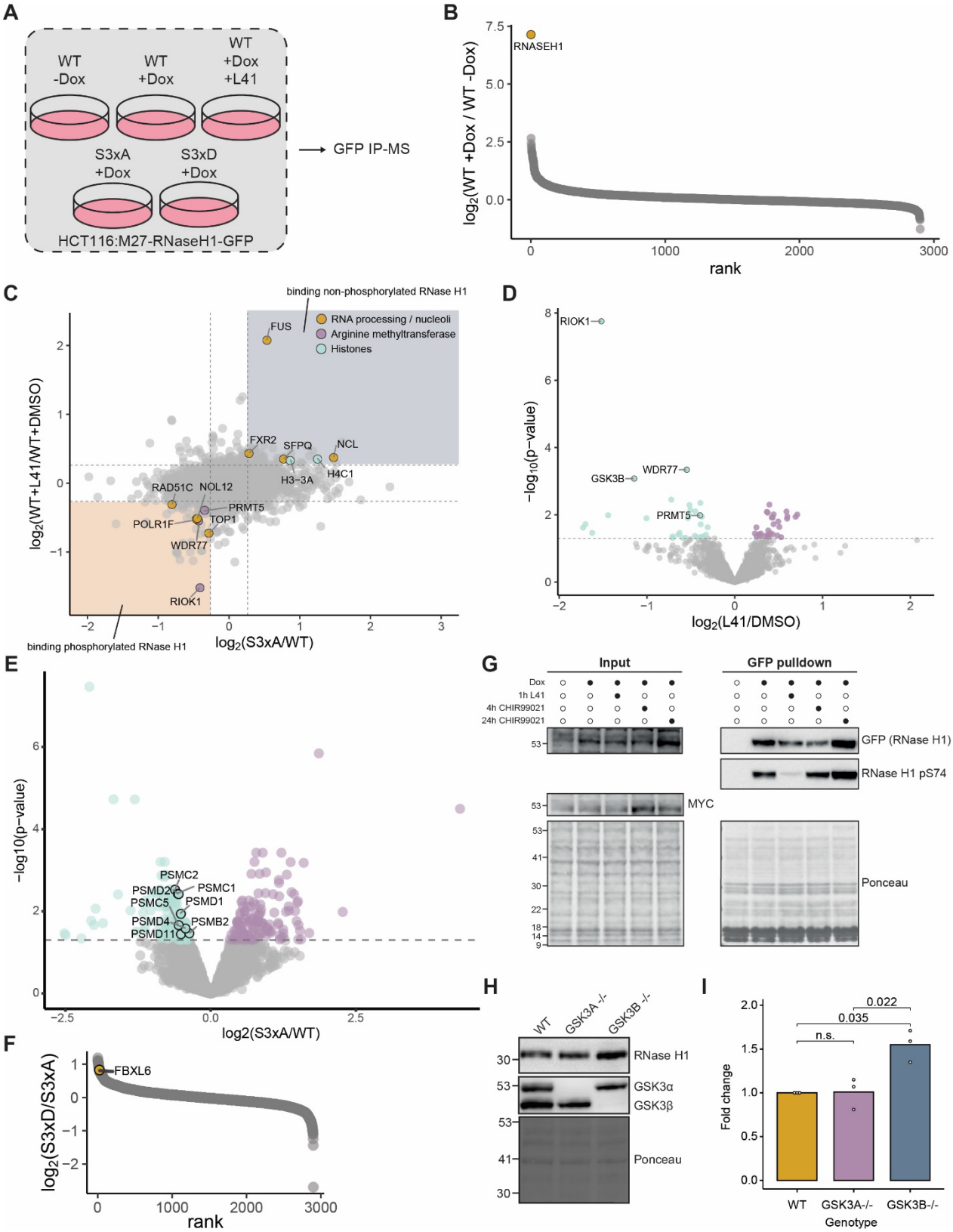

A
WT -Dox
WT +Dox
WT +Dox +L41
S3xA +Dox
S3xD +Dox
GFP IP-MS
HCT116:M27-RNaseH1-GFP
B
RNASEH1
log2(WT +Dox / WT -Dox)
rank
C
binding non-phosphorylated RNase H1
RNA processing / nucleoli
Arginine methyltransferase
Histones
FUS
FXR2
SFPQ
NCL
H4C1
H3−3A
RAD51C
NOL12
PRMT5
POLR1F
TOP1
WDR77
RIOK1
binding phosphorylated RNase H1
log2(WT+L41/WT+DMSO)
log2(S3xA/WT)
D
RIOK1
GSK3B
WDR77
PRMT5
−log10(p−value)
log2(L41/DMSO)
E
PSMC2
PSMD2
PSMC1
PSMC5
PSMD1
PSMD4
PSMB2
PSMD11
−log10(p−value)
log2(S3xA/WT)
F
FBXL6
log2(S3xD/S3xA)
rank
G
Input
GFP pulldown
Dox
1h L41
4h CHIR99021
24h CHIR99021
GFP (RNase H1)
RNase H1 pS74
MYC
Ponceau
H
WT
GSK3A -/-
GSK3B -/-
RNase H1
GSK3α
GSK3β
Ponceau
I
0.022
0.035
n.s.
Fold change
WT
GSK3A-/-
GSK3B-/-
Genotype

Figure 12: Analysis of the phosphorylation-based changes in the RNase H1 interactome reveals a role of GSK3B in RNase H1 stability.

A) Outline of the GFP-based pulldowns followed by LC-MS/MS after indicated treatments in HCT116:M27-RNaseH1-GFP. **B)** Rank plot of proteins from GFP-based pulldowns after 24h doxycycline or DMSO treatment based on *n* = 3 biologically independent experiments. RNase H1 is highlighted in yellow. **C)** Scatter plot comparing L41 treatment to DMSO in WT-expressing cells, and S3xA to WT expression. Factors belonging to selected terms are selected (see legend). **D)** Volcano plot showing differential binding to phospho-WT RNase H1 in L41- or DMSO-treated cells. Selected factors are highlighted. **E)** Volcano plot showing differential binding to WT and S3xA mutant. Proteasomal subunits are highlighted. **F)** Rank plot of proteins differentially binding to S3xA and S3xD mutant. FBXL6 is highlighted in yellow. **G)** Western blot showing GFP-based pulldowns in HCT116:M27-RNaseH1-GFP after a treatment with specified chemical inhibitors. Ponceau: loading control. **H)** Western blot showing total levels of RNase H1 in MEF cells with a GSK3α or GSK3β knockout. Ponceau: loading control. **I)** Quantification of the fold changes based on results from H, n = 3. Values are normalized on the protein content (Ponceau).

Chapter 4

Discussion

4.1. Relationship between replication stress and the cell cycle

Our efforts to stratify RS by severity revealed major cell cycle differences between the RS treatments in HCT116 (Figure 5C). For example, while mild RS did not significantly affect the cell cycle, moderate RS caused a shift towards the G2 phase, and strong RS towards the S phase. It would be interesting to explore whether longer treatments with mild RS eventually stall the cell cycle, considering that DNA repair processes are carried over into the subsequent stages of the cell cycle (G2 in case of post-replicative gap filling[277], mitosis in case of MiDAS[246], or G1 in case of the post-mitotic DNA synthesis[278]). In particular, inheriting DNA lesions into the subsequent cell cycle increases duration of the G1 phase[63].

Interestingly, live-cell imaging using the PCNA chromobody in U2OS cells revealed that treatment with moderate RS (400 nM APH) increases the duration of S phase from roughly 6 h to more than 24 h[279]. Therefore, the shift towards G2 after moderate RS can be explained, at least in part, by the severely prolonged duration of the S phase. In addition, multiple lines of evidence (Figure 5B,C; Figure 6D,F,H) suggest DNA damage checkpoint activation after moderate RS, likely as a result of unrepaired DNA lesions. Considering that strong RS almost completely blocks the progression of the replication forks[280], it is expected that cells are unable to finish the S phase. The cell cycle effects we observed are likely cell-type dependent, as each cell line has a unique cell cycle duration and sensitivity to RS can vary depending on the mutational landscape and the expression profile of the cell line.

The interpretation of the aphidicolin-induced changes in phosphorylation is complicated by the cell cycle effects. Therefore, to understand which events are a direct result of the RS, and which are purely cell-cycle driven, one would need to categorize significantly

changing sites after 24-h RS using the database of phosphorylation abundance changes across the cell cycle stages[281].

4.2. Comparison of Cyclin E- and APH-induced signaling

The cell cycle phase transitions are governed by the complexes of cyclins and CDKs[282]. In complex with CDK2, Cyclin E promotes the G1/S transition, which is premature if Cyclin E is amplified, leading to replication stress induction[47]. Comparison of different sources of RS revealed that APH unexpectedly mimics Cyclin E induction at the signaling level. Namely, phosphorylation of Cyclin E itself was increased after both Cyclin E induction and APH. Two of the commonly increased phospho-sites (T352, corresponding to T395 on the canonical CCNE1 isoform; and S344, corresponding to S387 on the canonical CCNE1 isoform) are CDK2 targets[283]. This is significant for two reasons. Firstly, it validates the expected increase in CDK2 activity after Cyclin E induction, which is further evidenced by the kinase predictions (Figure 7E). Secondly, it implies that CDK2 activity is also elevated after APH treatment. Of note, CDKs have similar phosphorylation motifs and overlapping sets of targets[284], making it difficult to tell them apart in kinase prediction approaches. Moreover, CCNE1 was also shown to bind to and activate CDK1[285]. Thus, our results provide further evidence that CDK activity is elevated, rather than inhibited, after mild RS. This is significant as CDK1 has been reported as the driver of increased origin firing after mild RS[252,286].

It is important to note that the comparison between CCNE1- and APH-induced RS was performed in the U2OS background, while other experiments were performed in HCT116. We noticed that U2OS cells were less overall less affected by RS than HCT116 cells, which might stem from their different sensitivities to RS.

4.3. RNase H1 interactome reveals transcription- and RNA processing-related roles

Despite the ubiquitous use of RNase H1 as a tool in R-loop biology, relatively little is known about how its function in human cells is determined by its interactors. Our

interaction profiling of RNase H1 revealed that it is largely associated with the transcription machinery and RNA processing factors (Figure 8G). For example, we have detected three RNA Pol II subunits (POLR2A, POLR2C and POLR2H) interacting with RNase H1. This suggests RNase H1 might have a co-transcriptional hybrid removal activity aided by interactions with core transcription components. This is supported by recent reports of RNase H1 phase-separating into Pol II transcriptional condensates[263]. Therefore, we speculate that the separation in RNase H1 and H2 functions might be determined by the fact that RNase H1 is associated with transcription machinery, while RNase H2 associates with the replication machinery[161]. RNase H1 also interacted with three subunits of the transcription elongation PAF1 complex, PAF1, CTR9 and LEO1 (Figure 8G). Interestingly, depletion of the components of the PAF1 complex leads to accumulation of R-loops[177]. We suggest testing whether the ability of the PAF1 complex to suppress R-loops, at least in part, depends on RNase H1.

4.4. Unresolved questions about the phosphorylation of RNase H1

Our phosphoproteomics results suggest that RNase H1 is less phosphorylated on all three serine residues (S74, S76 and S80) after RS (Figure 8D). However, it was so far difficult to discern whether semi-phosphorylated intermediates exist, or whether RNase H1 is either non-phosphorylated or triple-phosphorylated. Considering that, unlike S74 and S76, phosphorylation at S80 has not been reported on PhosphoSitePlus, this suggests it might be a rarer event that happens sub-stoichiometrically in comparison to phosphorylation at S74 and S76. One possibility is that phosphorylation at S80 introduces a phospho-degron, which will be discussed below. This would explain why it is more difficult to detect. In that case, treatment with proteasome inhibitors, such as MG132, should stabilize such phosphorylation and enable its easier detection, which we plan to test. Experiments with CX-4945 showed that DYRK1A might phosphorylate both S74 and S76 (Figure 8H). We consider this likely for the following reason. Although we have convincingly shown S74 is the substrate of DYRK1A, S76 better fits its consensus phosphorylation motif. Namely, DYRK1A has a preference for positively charged residues ahead of the serine/threonine,

particularly for arginine (R) at -3, and for proline (P) at +1[287]. The amino acid environment of S76 thus slightly better supports its phosphorylation (V**RK**S$_{74}$A**S**$_{76}$**P**EVS$_{80}$). An additional aspect is the phosphorylation occupancy, which can be tested in stringent pulldowns of GFP-tagged RNase H1[288]. Including inhibitor treatments in this setup will hopefully enable discerning the contribution of DYRK1A and GSK3β to phosphorylation of the three residues of interest. This can be further confirmed by *in vitro* kinase assays.

While our initial phosphoproteomics results do not distinguish the two isoforms (Figure 8D), the results using the M27-RNase H1-GFP suggest that DYRK1A phosphorylates the nuclear isoform (Figure 9C). Since DYRK1A was also found to phosphorylate mitochondrial proteins[289], it is possible that it acts on mitochondrial RNase H1 in a similar way. Alternatively, if RNase H1 is phosphorylated in the mitochondria or the cytoplasm, this could be mediated by another kinase.

4.5. Functions of RNase H1 phosphorylation

Our functional analyses suggest that phosphorylation of RNase H1 might regulate multiple biological roles, and that there might indeed be functional differences between semi- or fully-phosphorylated RNase H1. While we have started studying the effects of phosphorylation on hybrid binding, this should be further expanded. For example, we have so far shown that the S3xD mutant has reduced binding to hybrids in comparison to the S3xA mutant (Figure 11C,D). However, it is not clear whether S3xD functionally mimics phosphorylation of RNase H1. Therefore, another approach would be to test binding to hybrids upon *in vitro* phosphorylation using purified DYRK1A and phospho-WT or S3xA dCat RNase H1. Furthermore, the impact of phosphorylation on R-loop/hybrid degradation should be tested *in vitro*[212] and *in vivo*. To test this *in vivo*, we are planning to perform immunofluorescence experiments using the RHINO probe[116], as well as genomic mapping of R-loops using MapR[223].

Previous research had demonstrated that the HBD is critical for the nucleolar localization of RNase H1[197]. Our results demonstrate that the phospho-cluster regulates the degree of nucleolar-nucleoplasmic separation, likely either through interactions of RNase H1 with nucleic acid species, or with other proteins. This should be further tested, using treatments

expected to perturb RNase H1 phosphorylation, such as APH or DYRK1Ai. Considering that APH treatment reduces phosphorylation at all 3 serine residues (Figure 8D), we expect it would behave similarly like the S3xA mutant, although the effects would likely not be as striking. Treatment with DYRK1Ai would be less straightforward to interpret, as we are not sure it would perturb phosphorylation of all three serines. On the other hand, we are not sure if phosphorylation on all 3 sites is necessary for the change in localization. This could be tested by expressing GFP-tagged semi-phosphorylated phospho-mutants of RNase H1.

4.6. DYRK1A – a new R-loop regulator?

The phosphoproteome we performed after inhibition or depletion of DYRK1A offers valuable insights into DYRK1A signaling. For example, kinase prediction analysis revealed that DYRK1A inhibition increases activity of GSK3β and CDK2. Critically, DYRK1A inhibits GSK3β by phosphorylation at T356[290], providing an explanation for why GSK3β activity would be elevated upon DYRK1A inhibition. That CDK2 activity is increased is in agreement with previous phosphoprotemics experiments[287]. Mechanistically, DYRK1A inhibition might boost CDK activity by promoting the accumulation of cyclins[291]. Therefore, our results suggest that an increase in CDK activity observed after APH-induced mild RS (Figure 7E) might be due to lower DYRK1A activity (Figure 9F). This could be tested by overexpressing DYRK1A under conditions of mild RS and measuring CDK activity.

Although we did not manage to quantify sites on RNase H1 in the DYRK1Ai phosphoproteomics (Figure 10), likely due to the nature of data-dependent acquisition[292], we found a loss of phosphorylation at a double-phosphorylated RNase H1 peptide (S74/S76) upon chemical inhibition of DYRK1A in a previously published phosphoproteome[287].

Considering the intimate relationship between transcription, RNA processing and R-loop levels, as well as the large number of RNA processing factors affected by inhibition of DYRK1A, we speculate that it might have a role in R-loop homeostasis. For example, we found that inhibition of DYRK1A lowers phosphorylation of the DDX/DHX-box proteins

DDX21 and DHX38 (Figure 10F), which have been associated with R-loop resolution[151,293]. This role might be disease-relevant, considering that DYRK1A imbalance is implicated in neurological disorders and cancer[10]. Interestingly, Down syndrome is characterized by early-onset hematological cancers[294], which often occur due to R-loop mis-regulation[90,153,154,181]. Moreover, dose imbalance of DYRK1A in Down syndrome is associated with premature aging and unrepaired DNA damage[295], phenotypes that R-loops could also contribute to. We therefore plan to test R-loop levels after DYRK1A depletion or inhibition.

4.7. GSK3β – stabilization of RNase H1 through a ubiquitin-dependent mechanism?

The interactome of RNase H1 revealed that it interacts with GSK3β in a DYRK1A-dependent manner (Figure 12D). Based on this, we hypothesized that DYRK1A primes RNase H1 for GSK3β-mediated phosphorylation, as has previously been demonstrated for multiple substrates[32]. Although we did not experimentally test this, we find it likely that GSK3β would phosphorylate S80 of pre-phosphorylated RNase H1. Using AlphaFold3, we modeled GSK3β interaction with an RNase H1 peptide pre-phosphorylated at S74 and S76 (Figure 13A). The model with the highest iPTM score (0.8) showed an interaction of the GSK3β positively charged pocket, comprised of the residues R96, R180 and K205, with pS74 of RNase H1 (Figure 13B). This pocket is known to bind phosphorylated residues on primed GSK3β substrates[296]. Interestingly, 4 out of 5 models predicted interaction of the positively charged pocket with pS74, while one showed the interaction with pS76. Only the interaction with pS74 would allow the ATP to be placed proximally to S80 of RNase H1, and thereby enable its phosphorylation (Figure 13C). It is important to note that GSK3β usually phosphorylates substrates primed downstream of the phospho-site, at the +4 position[296]. In the case of RNase H1, we predict GSK3β would bind the priming residue at the -6 position, which, if experimentally validated, would reveal a new class of GSK3β priming preferences.

We plan to further investigate the possibility that GSK3β introduces a phospho-degron into RNase H1. This will consist of testing whether: (1) S80 of RNase H1 is responsive to

GSK3β inhibition; (2) DYRK1A is necessary for GSK3β-mediated phosphorylation; (3) the S3xA mutant loses sensitivity to GSK3β inhibition for its stabilization; (4) RNase H1 is ubiquitylated in a manner dependent on GSK3β, a phospho-degron binder (eg. FBXL6 - Figure 12F), the SCF E3 ligase[26] and NEDD8[297]; (5) the S3xA mutant loses GSK3β-dependent ubiquitylation.

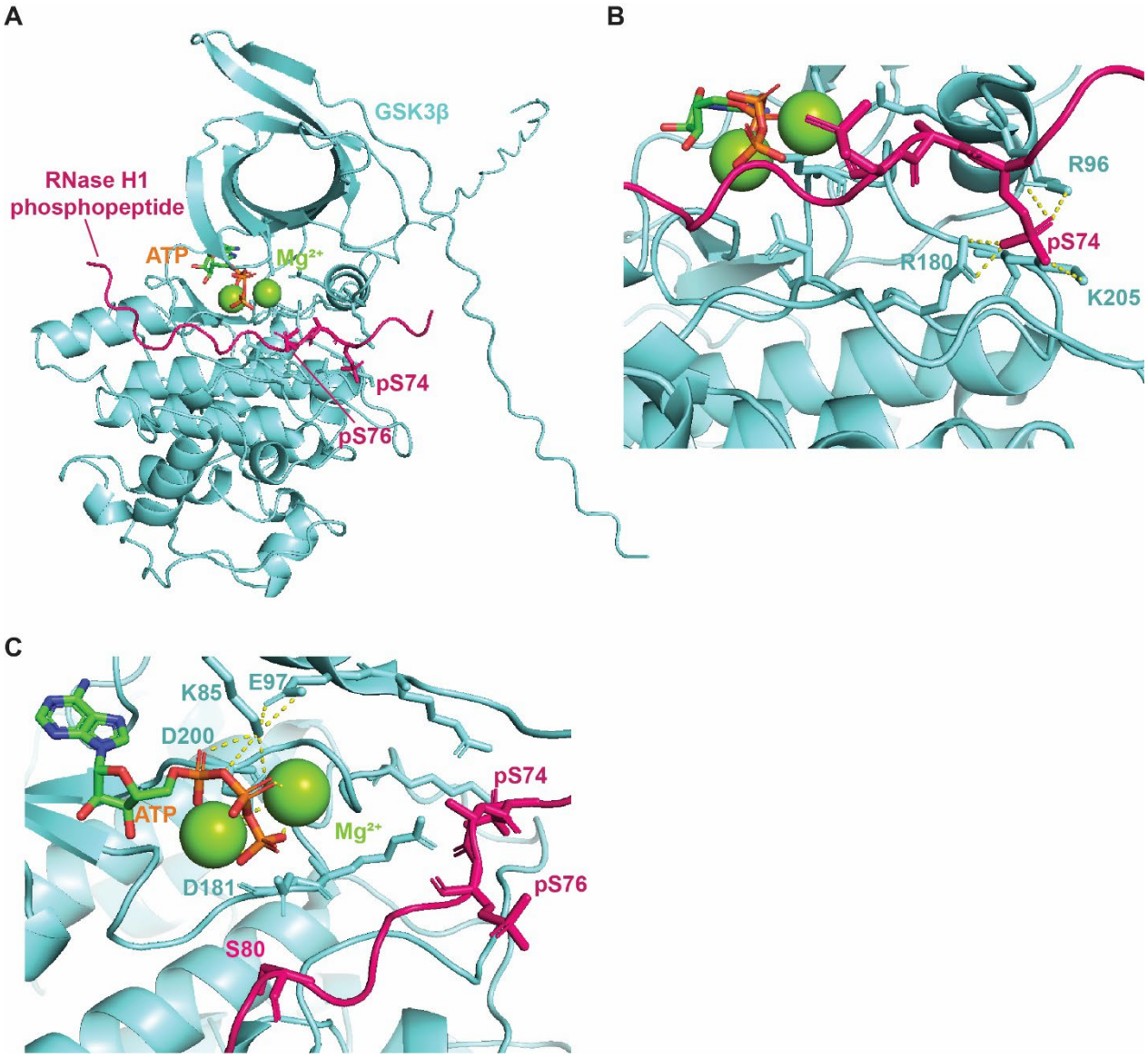


Figure 13: Modeling the interaction of GSK3β with the RNase H1 phospho-cluster region.

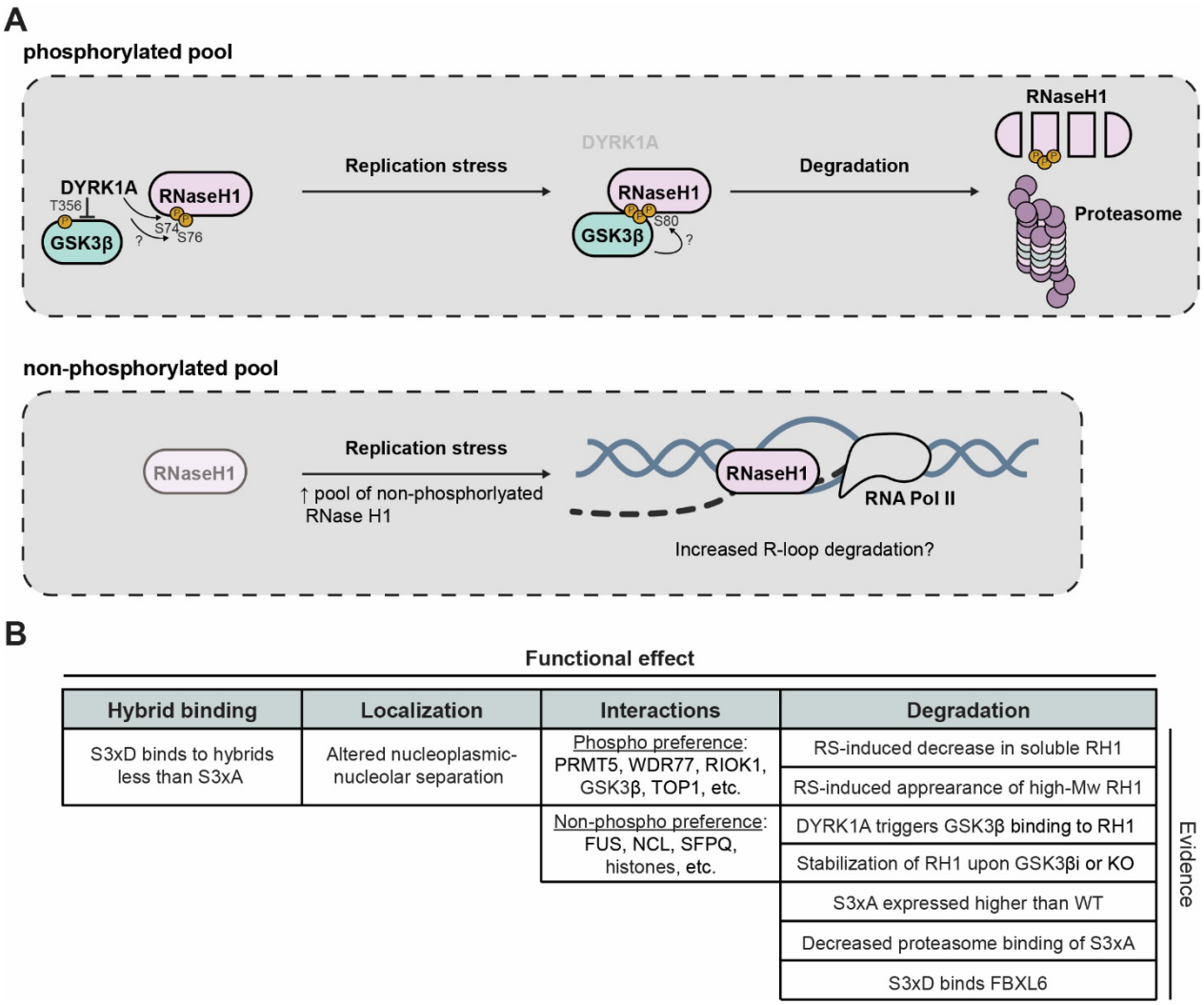
A) AlphaFold3-modeled interaction of the GSK3β with the double-phosphorylated (pS74-pS76) RNase H1 phosphopeptide rendered in PyMol. ATP and Mg^{2+} ions are also shown. **B)** Zoomed-in view of RNase H1 pS74 interacting with the positively charged pocket of

GSK3β. **C)** Zoomed-in view of the active site of GSK3β occupied by ATP and Mg^{2+} ions, positioned proximally to S80 of RNase H1.

4.8. Conclusions

Based on the presented results, we propose a model of regulation of RNase H1 phosphorylation and its function under replication stress (Figure 14A). In unchallenged conditions, RNase H1 is phosphorylated by DYRK1A at S74 (Figure 9E) and (likely) S76. DYRK1A also inhibits GSK3β activity by phosphorylating it at T356[290]. Under prolonged replication stress, DYRK1A levels are decreased (Figure 9F) and GSK3β is therefore activated. GSK3β then phosphorylates S80 of RNase H1 (Figure 12D,G-I; Figure 13), introducing a phospho-degron. The phosphorylated pool of RNase H1 is therefore degraded in a proteasome-dependent manner (Figure 12E). Because the non-phosphorylated pool is continuously synthesized, this increases its relative amount after replication stress (Figure 9C,D). Since non-phosphorylated version has a higher affinity towards RNA:DNA hybrids (Figure 11C,D), this should increase the degradation of R-loops. Importantly, this model does not require the activity of a phosphatase on RNase H1, as phosphorylation is lost through degradation. Alternatively, if effects of GSK3β on the levels of RNase H1 are more indirect, the loss of DYRK1A activity is sufficient to decrease the pool of phosphorylated RNase H1 upon replication stress.

We also propose that the semi-phosphorylated RNase H1 might affect the distribution of RNase H1 in the nucleoli and the nucleoplasm, its protein-protein interactions and the RNA:DNA hybrid binding (Figure 14B). Fully phosphorylated RNase H1 should, on the other hand, drive the degradation of RNase H1. The reader should keep in mind that parts of this model are preliminary and require further study.



B

Functional effect			
Hybrid binding	**Localization**	**Interactions**	**Degradation**
S3xD binds to hybrids less than S3xA	Altered nucleoplasmic-nucleolar separation	Phospho preference: PRMT5, WDR77, RIOK1, GSK3β, TOP1, etc.	RS-induced decrease in soluble RH1
			RS-induced apprearance of high-Mw RH1
		Non-phospho preference: FUS, NCL, SFPQ, histones, etc.	DYRK1A triggers GSK3β binding to RH1
			Stabilization of RH1 upon GSK3βi or KO
			S3xA expressed higher than WT
			Decreased proteasome binding of S3xA
			S3xD binds FBXL6

Evidence

Figure 14: Model – regulation of RNase H1 by phosphorylation.

A) DYRK1A-, RS- and GSK3β- mediated regulation of RNase H1 phosphorylation and stability. See section 4.8 for explanation. **B)** Proposed functions of RNase H1 phosphorylation and its loss. Evidence for each functional effect collected in this study is presented in the table.

Chapter 5

Appendix

5.1 Abbreviations

AOA2	Type 2 ataxia with oculomotor apraxia
AGS	Aicardi–Goutières syndrome
AML	Acute myeloid leukemia
APH	Aphidicolin
AQR	Aquarius
ALS4	juvenile amyotrophic lateral sclerosis
ALT	Alternative telomere lengthening
ATM	Ataxia-telangiectasia mutated
ATP	Adenosine-5’-diphosphate
ATR	Ataxia telangiectasia and Rad3 related
BER	Base excision repair
BIR	Break-induced replication
bp	Base pair
BRCA	Breast cancer
BrdU	5-Bromo-2'-deoxyuridine
BSA	Bovine serum albumin
CDK	Cyclin-dependent kinase
CFS	Common fragile site

ChIP	Chromatin immunoprecipitation
CIN	Chromosome instability
CldU	5-Chloro-2′-deoxyuridine
CP	Core particle
CTD	C-terminal domain
DDK	Dbf4-dependent kinase
DDX	DEAD-box
DHX	DEAH-box
dilncRNA	DNA damage-induced long non-coding RNA
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
Dox	Doxycycline
DS	Down syndrome
DSB	Double-strand break
dsDNA	Double-stranded deoxyribonucleic acid
DTT	dithiothreitol
DUB	Deubiquitylating enzyme
DYRK	Dual-specificity tyrosine phosphorylation-regulated kinase
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
EDTA	Ethylenediaminetetraacetic acid

G4	G-quadruplex
GI	Genome instability
GSK3	Glycogen synthase kinase
HBD	Hybrid-binding domain
HSPC	Hematopoietic stem progenitor cell
HR	Homologous recombination
HU	Hydroxyurea
IdU	5-Iodo-2′-deoxyuridine
IF	Immunofluorescence
IGS	Intergenic spacer
lncRNA	Long non-coding RNA
MCM	Minichromosome maintenance protein complex
MDS	Myelodysplastic syndrome
MiDAS	Mitotic DNA synthesis
MLS	Mitochondrial localization signal
NELF	Negative elongation factor
NER	Nucleotide excision repair
NFAT	Nuclear factor of T cells
NLS	Nuclear localization signal
NHEJ	Non-homologous end joining
nt	Nucleotide
OF	Okazaki fragment
ORC	Origin recognition complex

p-TEFb	Positive transcription elongation factor
PARP	Poly (ADP-Ribose) Polymerase
PBS	Phosphate-buffered saline
PCNA	Proliferating cell nuclear antigen
PEI	Polyethyleneimine
PEO	Progressive external ophthalmoplegia
Phospho	Phosphorylation
PIC	Pre-initiation complex
PIKK	Phosphatidylinositol 3-kinase-related kinase
PIP-box	PCNA interacting protein-box
PRC2	Polycomb repressive complex 2
pre-RC	pre-Replication Complex
PRIMPOL	Primase and DNA directed polymerase
PTM	Post-translational modification
rDNA	Ribosomal DNA
RBR	RING-between-RING
RER	Ribonucleotide excision repair
RF	Replication fork
RNA	Ribonucleic acid
RNR	Ribonucleotide reductase
RNAPII / Pol II	RNA polymerase II
rNTP	Ribonucleotide triphosphate
ROS	Reactive oxygen species

RP	Regulatory particle
RPA	Replication protein A
RS	Replication stress
S-CIN	Structural chromosome instability
ssDNA	Single-stranded DNA
TC-NER	Transcription-coupled nucleotide excision repair
TERRA	Telomere repeat-containing RNA
TLS	Translesion synthesis
TOP1	Topoisomerase 1
TRC	Transcription-replication conflict
TS	Template switching
TSS	Transcription start site
UPS	Ubiquitin-proteasome system
W-CIN	Whole-chromosome instability

5.2 Publications

1. Papadopoulos*, D.,Ha*, S. A.,Fleischhauer*, D.,Uhl, L.,Russell, T. J.,**Mikicic, I.**,Schneider, K.,Brem, A.,Valanju, O. R.,Cossa, G.,Gallant, P.,Schuelein-Voelk, C.,Maric, H. M.,Beli, P.,Büchel, G.,Vos, S. M. & Eilers, M. The MYCN oncoprotein is an RNA-binding accessory factor of the nuclear exosome targeting complex. *Molecular Cell* **84**, 2070-2086.e20 (2024).
2. Bhatia, P.,Bickle, M.,Agrawal, A. A.,Truss, B.,Nikolaidi, A.,Brockmann, K.,Reinhardt, L.,Vogel, S.,Szegoe, E. M.,Pal, A.,Hermann, A.,**Mikicic, I.**,Yun, M.,Falkenburger, B. & Sterneckert, J. Axonal Lysosomal Assays for Characterizing the Effects of LRRK2 G2019S. *Biology* **13**, 58 (2024).
3. Suryo Rahmanto*, A.,Blum*, C. J.,Scalera, C.,Heidelberger, J. B.,Mesitov, M.,Horn-Ghetko, D.,Gräf, J. F.,**Mikicic, I.**,Hobrecht, R.,Orekhova, A.,Ostermaier, M.,Ebersberger, S.,Möckel, M. M.,Krapoth, N.,Da Silva Fernandes, N.,Mizi, A.,Zhu, Y.,Chen, J. X.,Choudhary, C. *et al.* K6-linked ubiquitylation marks formaldehyde-induced RNA-protein crosslinks for resolution. *Molecular Cell* **83**, 4272-4289.e10 (2023).
4. Stier*, A.,Gilberto*, S.,Mohamed*, W. I.,Royall, L. N.,Helenius, J.,**Mikicic, I.**,Sajic, T.,Beli, P.,Müller, D. J.,Jessberger, S. & Peter, M. The CUL4B -based E3 ubiquitin ligase regulates mitosis and brain development by recruiting phospho-specific DCAFs . *The EMBO Journal* **42**, (2023).
5. Shi, J.,Hauschulte, K.,**Mikicic, I.**,Maharjan, S.,Arz, V.,Strauch, T.,Heidelberger, J. B.,Schaefer, J. V.,Dreier, B.,Plückthun, A.,Beli, P.,Ulrich, H. D. & Wollscheid, H. P. Nuclear myosin VI maintains replication fork stability. *Nature Communications* **14**, 1–13 (2023).
6. Solvie*, D.,Baluapuri*, A.,Uhl*, L.,Fleischhauer, D.,Endres, T.,Papadopoulos, D.,Aziba, A.,Gaballa, A.,**Mikicic, I.**,Isaakova, E.,Giansanti, C.,Jansen, J.,Jungblut, M.,Klein, T.,Schülein-Völk, C.,Maric, H.,Doose, S.,Sauer, M.,Beli, P. *et al.* MYC multimers shield stalled replication forks from RNA polymerase. *Nature* **612**, 148–155 (2022).
7. Mosler*, T.,Baymaz*, H. I.,Gräf, J. F.,**Mikicic, I.**,Blattner, G.,Bartlett, E.,Ostermaier, M.,Piccinno, R.,Yang, J.,Voigt, A.,Gatti, M.,Pellegrino, S.,Altmeyer, M.,Luck, K.,Ahel, I.,Roukos, V. & Beli, P. PARP1 proximity proteomics reveals interaction partners at stressed replication forks. *Nucleic Acids Research* **50**, 11600–11618 (2022).
8. Gräf*, J. F.,**Mikicic*, I.**,Ping, X.,Scalera, C.,Mayr, K.,Stelzl, L. S.,Beli, P. & Wagner, S. A. Substrate spectrum of PPM1D in the cellular response to DNA double-strand breaks. *iScience* **25**, (2022).
9. Wegmann*, S.,Meister*, C.,Renz, C.,Yakoub, G.,Wollscheid, H. P.,Takahashi, D. T.,**Mikicic, I.**,Beli, P. & Ulrich, H. D. Linkage reprogramming by tailor-made E3s reveals polyubiquitin chain requirements in DNA-damage bypass. *Molecular Cell* **82**, 1589-1602.e5 (2022).
10. Mosler, T.,Conte, F.,Longo, G. M. C.,**Mikicic, I.**,Kreim, N.,Möckel, M. M.,Petrosino, G.,Flach, J.,Barau, J.,Luke, B.,Roukos, V. & Beli, P. R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nature Communications* **12**, 1–17 (2021).

* indicates joint authorship

5.3 Curriculum Vitae

Redacted for data protection reasons.

Acknowledgements

Redacted for data protection reasons.

Chapter 6

Literature

1. Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Reports 2011 11* **1**, 1–5 (2011).
2. Wilson, L. J. *et al.* New Perspectives, opportunities, and challenges in exploring the human protein kinome. *Cancer Res.* **78**, 15–29 (2018).
3. Kitata, R. B. *et al.* A data-independent acquisition-based global phosphoproteomics system enables deep profiling. *Nat. Commun.* **12**, 1–14 (2021).
4. Humphrey, S. J., James, D. E. & Mann, M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol. Metab.* **26**, 676–687 (2015).
5. Kostich, M. *et al.* Human members of the eukaryotic protein kinase family. *Genome Biol.* **3**, research0043.1 (2002).
6. Sugiyama, N., Imamura, H. & Ishihama, Y. Large-scale Discovery of Substrates of the Human Kinome. *Sci. Reports 2019 91* **9**, 1–12 (2019).
7. Johnson, J. L. *et al.* An atlas of substrate specificities for the human serine/threonine kinome. *Nat. 2023 6137945* **613**, 759–766 (2023).
8. St-Denis, N. *et al.* Systematic investigation of hierarchical phosphorylation by protein kinase CK2. *J. Proteomics* **118**, 49–62 (2015).
9. Aoki, K., Yoshida, K., Aoki, K. & Yoshida, K. Biological Consequences of Priming Phosphorylation in Cancer Development. *Protein Phosphorylation* (2017) doi:10.5772/INTECHOPEN.70039.
10. Rammohan, M. *et al.* The chromosome 21 kinase DYRK1A: emerging roles in cancer biology and potential as a therapeutic target. *Oncogene 2022 4114* **41**, 2003–2011 (2022).

11. Kentrup, H. *et al.* Dyrk, a Dual Specificity Protein Kinase with Unique Structural Features Whose Activity Is Dependent on Tyrosine Residues between Subdomains VII and VIII. *J. Biol. Chem.* **271**, 3488–3495 (1996).

12. Lu, H. *et al.* Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nat. 2018 5587709* **558**, 318–323 (2018).

13. Vona, C. D. *et al.* Chromatin-wide Profiling of DYRK1A Reveals a Role as a Gene-Specific RNA Polymerase II CTD Kinase. *Mol. Cell* **57**, 506–520 (2015).

14. Price, D. H. P-TEFb, a Cyclin-Dependent Kinase Controlling Elongation by RNA Polymerase II. *Mol. Cell. Biol.* **20**, 2629–2634 (2000).

15. Kaczmarski, W. *et al.* Intracellular distribution of differentially phosphorylated dual-specificity tyrosine phosphorylation-regulated kinase 1A (DYRK1A). *J. Neurosci. Res.* **92**, 162–173 (2014).

16. Sperling, K., Scherb, H. & Neitzel, H. Population monitoring of trisomy 21: problems and approaches. *Mol. Cytogenet.* **16**, 1–12 (2023).

17. Dowjat, W. K. *et al.* Trisomy-driven overexpression of DYRK1A kinase in the brain of subjects with Down syndrome. *Neurosci. Lett.* **413**, 77–81 (2007).

18. Feki, A. & Hibaoui, Y. DYRK1A Protein, A Promising Therapeutic Target to Improve Cognitive Deficits in Down Syndrome. *Brain Sci. 2018, Vol. 8, Page 187* **8**, 187 (2018).

19. Liu, F. *et al.* Overexpression of Dyrk1A contributes to neurofibrillary degeneration in Down syndrome. *FASEB J.* **22**, 3224–3233 (2008).

20. Neumann, F. *et al.* DYRK1A inhibition and cognitive rescue in a Down syndrome mouse model are induced by new fluoro-DANDY derivatives. *Sci. Reports 2018 81* **8**, 1–12 (2018).

21. Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P. & Decaprio, J. A. DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. *Genes Dev.* **25**, 801–813 (2011).

22. Guard, S. E. *et al.* The nuclear interactome of DYRK1A reveals a functional role in

DNA damage repair. *Sci. Reports 2019 91* **9**, 1–12 (2019).

23. R. Menon, V. *et al.* DYRK1A regulates the recruitment of 53BP1 to the sites of DNA damage in part through interaction with RNF169. *Cell Cycle* **18**, 531–551 (2019).

24. Fiol, C. J., Wang, A., Roeske, R. W., Roach, P. J. & Hunter, T. Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates. *J. Biol. Chem.* **265**, 6061–6065 (1990).

25. Lee, C. J. *et al.* FBXW7-mediated stability regulation of signal transducer and activator of transcription 2 in melanoma formation. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 584–594 (2020).

26. Robertson, H., Hayes, J. D. & Sutherland, C. A partnership with the proteasome; the destructive nature of GSK3. *Biochem. Pharmacol.* **147**, 77–92 (2018).

27. Dobrowolski, R. & De Robertis, E. M. Endocytic control of growth factor signalling: multivesicular bodies as signalling organelles. *Nat. Rev. Mol. Cell Biol. 2011 131* **13**, 53–60 (2011).

28. Willems, A. R., Schwab, M. & Tyers, M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 133–170 (2004).

29. Yada, M. *et al.* Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *EMBO J.* **23**, 2116–2125 (2004).

30. Pérez-Benavente, B. *et al.* GSK3-SCF(FBXW7) targets JunB for degradation in G2 to preserve chromatid cohesion before anaphase. *Oncogene* **32**, 2189–2199 (2013).

31. Woods, Y. L. *et al.* The kinase DYRK phosphorylates protein-synthesis initiation factor eIF2Bε at Ser539 and the microtubule-associated protein tau at Thr212: potential role for DYRK as a glycogen synthase kinase 3-priming kinase. *Biochem. J.* **355**, 609–615 (2001).

32. Scales, T. M. E., Lin, S., Kraus, M., Goold, R. G. & Gordon-Weeks, P. R. Nonprimed and DYRK1A-primed GSK3β-phosphorylation sites on MAP1B regulate microtubule dynamics in growing axons. *J. Cell Sci.* **122**, 2424–2435 (2009).

33. Hasle, H., Haunstrup Clemmensen, I. & Mikkelsen, M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet* **355**, 165–169 (2000).

34. Mancini, M. & Toker, A. NFAT proteins: emerging roles in cancer progression. *Nat. Rev. Cancer 2009 911* **9**, 810–820 (2009).

35. McClellan, A. J., Laugesen, S. H. & Ellgaard, L. Cellular functions and molecular mechanisms of non-lysine ubiquitination. *Open Biol.* **9**, (2019).

36. Zheng, N. & Shabek, N. Ubiquitin ligases: Structure, function, and regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017).

37. Lange, S. M., Armstrong, L. A. & Kulathu, Y. Deubiquitinases: From mechanisms to their inhibition by small molecules. *Mol. Cell* **82**, 15–29 (2022).

38. Lee, J. S. *et al.* A Multiubiquitin Chain Is Confined to Specific Lysine in a Targeted Short-Lived Protein. *Science (80-.).* **243**, 1576–1583 (1989).

39. Murata, S., Yashiroda, H. & Tanaka, K. Molecular mechanisms of proteasome assembly. *Nat. Rev. Mol. Cell Biol. 2009 102* **10**, 104–115 (2009).

40. Parker, M. W., Botchan, M. R. & Berger, J. M. Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **52**, 107–144 (2017).

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

42. Goehring, L., Huang, T. T. & Smith, D. J. Transcription-Replication Conflicts as a Source of Genome Instability. *Annu. Rev. Genet.* **57**, 157–179 (2023).

43. Pai, C. C. & Kearsey, S. E. A Critical Balance: dNTPs and the Maintenance of Genome Stability. *Genes (Basel).* **8**, 57 (2017).

44. Zhong, Y. *et al.* The level of origin firing inversely affects the rate of replication fork progression. *J. Cell Biol.* **201**, 373 (2013).

45. Baranovskiy, A. G. *et al.* Structural basis for inhibition of DNA replication by aphidicolin. *Nucleic Acids Res.* **42**, 14013–14021 (2014).

46. Sakano, K., Oikawa, S., Hasegawa, K. & Kawanishi, S. Hydroxyurea Induces Site-

specific DNA Damage via Formation of Hydrogen Peroxide and Nitric Oxide. *Japanese J. Cancer Res.* **92**, 1166–1174 (2001).

47. Macheret, M. & Halazonetis, T. D. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* **555**, 112–116 (2018).

48. Byun, T. S., Pacek, M., Yee, M. C., Walter, J. C. & Cimprich, K. A. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev.* **19**, 1040–1052 (2005).

49. Anindya, R. Single-stranded DNA damage: Protecting the single-stranded DNA from chemical attack. *DNA Repair (Amst).* **87**, 102804 (2020).

50. Toledo, L. I. *et al.* ATR Prohibits Replication Catastrophe by Preventing Global Exhaustion of RPA. *Cell* **155**, 1088–1103 (2013).

51. Thada, V. & Cortez, D. Common motifs in ETAA1 and TOPBP1 required for ATR kinase activation. *J. Biol. Chem.* **294**, 8395 (2019).

52. Mordes, D. A., Glick, G. G., Zhao, R. & Cortez, D. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes Dev.* **22**, 1478 (2008).

53. Haahr, P. *et al.* Activation of the ATR kinase by the RPA-binding protein ETAA1. *Nat. Cell Biol. 2016 1811* **18**, 1196–1207 (2016).

54. Lee, Y. C., Zhou, Q., Chen, J. & Yuan, J. RPA-Binding Protein ETAA1 Is an ATR Activator Involved in DNA Replication Stress Response. *Curr. Biol.* **26**, 3257–3268 (2016).

55. Traven, A. & Heierhorst, J. SQ/TQ cluster domains: concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. *BioEssays* **27**, 397–407 (2005).

56. Saxena, S. & Zou, L. Hallmarks of DNA replication stress. *Mol. Cell* **82**, 2298–2314 (2022).

57. Costanzo, V. *et al.* An ATR- and Cdc7-Dependent DNA Damage Checkpoint that Inhibits Initiation of DNA Replication. *Mol. Cell* **11**, 203–213 (2003).

58. Moiseeva, T. N. *et al.* An ATR and CHK1 kinase signaling mechanism that limits

origin firing during unperturbed DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 13374–13383 (2019).

59. Woodward, A. M. *et al.* Excess Mcm2–7 license dormant origins of replication that can be used under conditions of replicative stress. *J. Cell Biol.* **173**, 673–683 (2006).

60. Ge, X. Q., Jackson, D. A. & Blow, J. J. Dormant origins licensed by excess Mcm2–7 are required for human cells to survive replicative stress. *Genes Dev.* **21**, 3331–3341 (2007).

61. Ibarra, A., Schwob, E. & Méndez, J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8956–8961 (2008).

62. Yekezare, M., Gó mez-González, B. & Diffley, J. F. X. Controlling DNA replication origins in response to DNA damage – inhibit globally, activate locally. *J. Cell Sci.* **126**, 1297–1306 (2013).

63. Lezaja, A. & Altmeyer, M. Inherited DNA lesions determine G1 duration in the next cell cycle. *Cell Cycle* **17**, 24–32 (2018).

64. Berti, M., Cortez, D. & Lopes, M. The plasticity of DNA replication forks in response to clinically relevant genotoxic stress. *Nat. Rev. Mol. Cell Biol.* **21**, 633–651 (2020).

65. Cybulla, E. & Vindigni, A. Leveraging the replication stress response to optimize cancer therapy. *Nat. Rev. Cancer 2022 231* **23**, 6–24 (2022).

66. Quinet, A., Lemaçon, D. & Vindigni, A. Replication Fork Reversal: Players and Guardians. *Mol. Cell* **68**, 830–833 (2017).

67. Tian, T. *et al.* The ZATT-TOP2A-PICH Axis Drives Extensive Replication Fork Reversal to Promote Genome Stability. *Mol. Cell* **81**, 198-211.e6 (2021).

68. Lemaçon, D. *et al.* MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat. Commun. 2017 81* **8**, 1–12 (2017).

69. Mutreja, K. *et al.* ATR-Mediated Global Fork Slowing and Reversal Assist Fork Traverse and Prevent Chromosomal Breakage at DNA Interstrand Cross-Links. *Cell*

Rep. **24**, 2629-2642.e5 (2018).

70. Tye, S., Ronson, G. E. & Morris, J. R. A fork in the road: Where homologous recombination and stalled replication fork protection part ways. *Semin. Cell Dev. Biol.* **113**, 14–26 (2021).

71. Porebski, B. *et al.* WRNIP1 Protects Reversed DNA Replication Forks from SLX4-Dependent Nucleolytic Cleavage. *iScience* **21**, 31 (2019).

72. Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science (80-.).* **355**, 1152–1158 (2017).

73. Sale, J. E. Translesion DNA Synthesis and Mutagenesis in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).

74. Adar, S., Izhar, L., Hendel, A., Geacintov, N. & Livneh, Z. Repair of gaps opposite lesions by homologous recombination in mammalian cells. *Nucleic Acids Res.* **37**, 5737–5748 (2009).

75. Mourón, S. *et al.* Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nat. Struct. Mol. Biol. 2013 2012* **20**, 1383–1389 (2013).

76. Wong, R. P., Petriukov, K. & Ulrich, H. D. Daughter-strand gaps in DNA replication – substrates of lesion processing and initiators of distress signalling. *DNA Repair (Amst).* **105**, 103163 (2021).

77. Panzarino, N. J. *et al.* Replication Gaps Underlie BRCA-deficiency and Therapy Response. *Cancer Res.* **4421**, canres.1602.2020 (2020).

78. Cong, K. *et al.* Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency. *Mol. Cell* **81**, 3128-3144.e7 (2021).

79. Quinet, A., Tirman, S., Cybulla, E., Meroni, A. & Vindigni, A. To skip or not to skip: choosing repriming to tolerate DNA damage. *Mol. Cell* **81**, 649–658 (2021).

80. Thompson, S. L., Bakhoum, S. F. & Compton, D. A. Mechanisms of Chromosomal Instability. *Curr. Biol.* **20**, R285–R295 (2010).

81. Burrell, R. A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).

82. Böhly, N., Kistner, M. & Bastians, H. Mild replication stress causes aneuploidy by deregulating microtubule dynamics in mitosis. *Cell Cycle* **18**, 2770–2783 (2019).

83. Wilhelm, T. *et al.* Mild replication stress causes chromosome mis-segregation via premature centriole disengagement. *Nat. Commun.* **10**, 1–14 (2019).

84. Gyüre, Z. *et al.* Spontaneous mutagenesis in human cells is controlled by REV1-Polymerase ζ and PRIMPOL. *Cell Rep.* **42**, 112887 (2023).

85. Mocanu, C. & Chan, K. L. Mind the replication gap. *R. Soc. Open Sci.* **8**, (2021).

86. Glover, T. W., Wilson, T. E. & Arlt, M. F. Fragile sites in cancer: more than meets the eye. *Nat. Rev. Cancer 2017 178* **17**, 489–501 (2017).

87. Bjerregaard, V. A., Özer, Ö., Hickson, I. D. & Liu, Y. The detection and analysis of chromosome fragile sites. *Methods Mol. Biol.* **1672**, 471–482 (2018).

88. Bakhoum, S. F. & Landau, D. A. Chromosomal Instability as a Driver of Tumor Heterogeneity and Evolution. *Cold Spring Harb Perspect Med* **7**, a029611 (2017).

89. Kotsantis, P., Petermann, E. & Boulton, S. J. Mechanisms of oncogene-induced replication stress: Jigsaw falling into place. *Cancer Discov.* **8**, 537–555 (2018).

90. Hirayama, M., Shinriki, S. & Matsui, H. R-loops in normal and malignant hematopoiesis. *Front. Hematol.* **2**, 1297657 (2023).

91. Chuong, J. N. *et al.* DNA replication errors are a major source of adaptive gene amplification. *bioRxiv* 2024.05.03.589936 (2024) doi:10.1101/2024.05.03.589936.

92. Cramer, P. Eukaryotic Transcription Turns 50. *Cell* **179**, 808–812 (2019).

93. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol. 2018 1910* **19**, 621–637 (2018).

94. Poss, Z. C., Ebmeier, C. C. & Taatjes, D. J. The Mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **48**, 575–608 (2013).

95. Coin, F. & Egly, J. M. Revisiting the Function of CDK7 in Transcription by Virtue of a Recently Described TFIIH Kinase Inhibitor. *Mol. Cell* **59**, 513–514 (2015).

96. Wong, K. H., Jin, Y. & Struhl, K. TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Mol.*

Cell **54**, 601–612 (2014).

97. Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* **33**, 960–982 (2019).

98. Egloff, S. CDK9 keeps RNA polymerase II on track. *Cell. Mol. Life Sci. 2021 7814* **78**, 5543–5567 (2021).

99. Neve, J., Patel, R., Wang, Z., Louey, A. & Furger, A. M. Cleavage and polyadenylation: Ending the message expands gene regulation. *RNA Biol.* **14**, 865–890 (2017).

100. Belotserkovskii, B. P., Tornaletti, S., D'Souza, A. D. & Hanawalt, P. C. R-loop generation during transcription: Formation, processing and cellular outcomes. *DNA Repair (Amst).* **71**, 69–81 (2018).

101. 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).

102. 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).

103. Castillo-Guzman, D. & Chédin, F. Defining R-loop classes and their contributions to genome instability. *DNA Repair (Amst).* **106**, 103182 (2021).

104. Niehrs, C. & Luke, B. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol. 2020 213* **21**, 167–178 (2020).

105. El Hage, A., French, S. L., Beyer, A. L. & Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* **24**, 1546–1558 (2010).

106. Stoy, H. *et al.* Direct visualization of transcription-replication conflicts reveals post-replicative DNA:RNA hybrids. *Nat. Struct. Mol. Biol. 2023 303* **30**, 348–359 (2023).

107. 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).

108. Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **2014**, (2014).

109. Burgers, P. M. Solution to the 50-year-old Okazaki-fragment problem. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 3358–3360 (2019).

110. Williams, J. S. & Kunkel, T. A. Ribonucleotides in DNA: Origins, repair and consequences. *DNA Repair (Amst).* **19**, 27–37 (2014).

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

112. Michelini, F. *et al.* Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat. Cell Biol. 2017 1912* **19**, 1400–1411 (2017).

113. Rippe, K. & Luke, B. TERRA and the state of the telomere. *Nat. Struct. Mol. Biol. 2015 2211* **22**, 853–858 (2015).

114. Arora, R. *et al.* RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* **5**, 1–11 (2014).

115. Kabeche, L., Nguyen, H. D., Buisson, R. & Zou, L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science (80-.).* **359**, 108–114 (2018).

116. Martin, R. M., De Almeida, M. R., Gameiro, E. & De Almeida, S. F. Live-cell imaging unveils distinct R-loop populations with heterogeneous dynamics. *Nucleic Acids Res.* **51**, 11010–11023 (2023).

117. 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).

118. Abraham, K. J. *et al.* Nucleolar RNA polymerase II drives ribosome biogenesis. *Nat. 2020 5857824* **585**, 298–302 (2020).

119. Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science (80-.).* **352**, (2016).

120. Gromak, N., West, S. & Proudfoot, N. J. Pause Sites Promote Transcriptional Termination of Mammalian RNA Polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996 (2006).

121. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chédin, F. GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res.* **23**, 1590–1600 (2013).

122. 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**, e1003705 (2013).

123. Gibbons, H. R. *et al.* Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. *Front. Immunol.* **9**, 365395 (2018).

124. Arab, K. *et al.* Long Noncoding RNA TARID Directs Demethylation and Activation of the Tumor Suppressor TCF21 via GADD45A. *Mol. Cell* **55**, 604–614 (2014).

125. Cohen, S. *et al.* Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nat. Commun.* **9**, 1–14 (2018).

126. Bader, A. S. & Bushell, M. DNA:RNA hybrids form at DNA double-strand breaks in transcriptionally active loci. *Cell Death Dis. 2020 114* **11**, 1–7 (2020).

127. Pessina, F. *et al.* Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nat. Cell Biol. 2019 2110* **21**, 1286–1299 (2019).

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

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

130. 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).

131. D'Alessandro, G. *et al.* BRCA2 controls DNA:RNA hybrid level at DSBs by

mediating RNase H2 recruitment. *Nat. Commun. 2018 91* **9**, 1–17 (2018).

132. McIlwraith, M. J. *et al.* Reconstitution of the strand invasion step of double-strand break repair using human Rad51 Rad52 and RPA proteins. *J. Mol. Biol.* **304**, 151–164 (2000).

133. Keskin, H. *et al.* Transcript-RNA-templated DNA recombination and repair. *Nat. 2014 5157527* **515**, 436–439 (2014).

134. Meers, C. *et al.* Genetic Characterization of Three Distinct Mechanisms Supporting RNA-Driven DNA Repair and Modification Reveals Major Role of DNA Polymerase ζ. *Mol. Cell* **79**, 1037-1050.e5 (2020).

135. Wei, L. *et al.* DNA damage during the G0/G1 phase triggers RNA-templated, Cockayne syndrome B-dependent homologous recombination. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3495–E3504 (2015).

136. Rinaldi, C., Pizzul, P., Longhese, M. P. & Bonetti, D. Sensing R-Loop-Associated DNA Damage to Safeguard Genome Stability. *Front. Cell Dev. Biol.* **8**, 1–13 (2021).

137. Petermann, E., Lan, L. & Zou, L. Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids. *Nat. Rev. Mol. Cell Biol. 2022 238* **23**, 521–540 (2022).

138. Woolley, P. R. *et al.* Regulation of transcription patterns, poly(ADP-ribose), and RNA-DNA hybrids by the ATM protein kinase. *Cell Rep.* **43**, 113896 (2024).

139. Lindahl, T. Instability and decay of the primary structure of DNA. *Nat. 1993 3626422* **362**, 709–715 (1993).

140. García-Muse, T. & Aguilera, A. Transcription–replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol. 2016 179* **17**, 553–563 (2016).

141. 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 (2017).

142. Shin, J. H. & Kelman, Z. The Replicative Helicases of Bacteria, Archaea, and Eukarya Can Unwind RNA-DNA Hybrid Substrates. *J. Biol. Chem.* **281**, 26914–

26921 (2006).

143. Šviković, S. *et al.* R-loop formation during S phase is restricted by PrimPol-mediated repriming. *EMBO J.* **38**, (2019).

144. Matos, D. A. *et al.* ATR Protects the Genome against R Loops through a MUS81-Triggered Feedback Loop. *Mol. Cell* **77**, 514-527.e4 (2020).

145. Sollier, J. *et al.* Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. **56**, 777–785 (2015).

146. Crossley, M. P. *et al.* R-loop-derived cytoplasmic RNA–DNA hybrids activate an immune response. *Nat. 2022 6137942* **613**, 187–194 (2022).

147. Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst).* **10**, 654–665 (2011).

148. 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).

149. Kim, S. *et al.* ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork. *Nucleic Acids Res.* **48**, 7218–7238 (2020).

150. Pinter, S. *et al.* A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation. *Nucleic Acids Res.* **49**, 4350–4370 (2021).

151. 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).

152. Pérez-Calero, C. *et al.* UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev.* **34**, 898–912 (2020).

153. Mosler, T. *et al.* R-loop proximity proteomics identifies a role of DDX41 in

transcription-associated genomic instability. *Nat. Commun. 2021 121* **12**, 1–17 (2021).

154. Weinreb, J. T. *et al.* Excessive R-loops trigger an inflammatory cascade leading to increased HSPC production. *Dev. Cell* **56**, 627-640.e5 (2021).

155. 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).

156. 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).

157. Arnould, C. *et al.* Chromatin compartmentalization regulates the response to DNA damage. *Nat. 2023 6237985* **623**, 183–192 (2023).

158. Gatti, V., De Domenico, S., Melino, G. & Peschiaroli, A. Senataxin and R-loops homeostasis: multifaced implications in carcinogenesis. *Cell Death Discov. 2023 91* **9**, 1–10 (2023).

159. Shaban, N. M., Harvey, S., Perrino, F. W. & Hollis, T. The Structure of the Mammalian RNase H2 Complex Provides Insight into RNA·DNA Hybrid Processing to Prevent Immune Dysfunction. *J. Biol. Chem.* **285**, 3617–3624 (2010).

160. 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).

161. Bubeck, D. *et al.* PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Res.* **39**, 3652–3666 (2011).

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

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

164. Kim, N. *et al.* Mutagenic processing of ribonucleotides in DNA by yeast

topoisomerase I. *Science (80-.).* **332**, 1561–1564 (2011).

165. 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).

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

167. O'Connell, K., Jinks-Robertson, S. & Petes, T. D. Elevated Genome-wide instability in yeast mutants lacking RNase H activity. *Genetics* **201**, 963–975 (2015).

168. Crow, Y. J. *et al.* Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. *Nat. Genet. 2006 388* **38**, 910–916 (2006).

169. Günther, C. *et al.* Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. *J. Clin. Invest.* **125**, 413–424 (2015).

170. Lee-Kirsch, M. A., Wolf, C. & Günther, C. Aicardi–Goutières syndrome: a model disease for systemic autoimmunity. *Clin. Exp. Immunol.* **175**, 17–24 (2013).

171. Okude, H., Ori, D. & Kawai, T. Signaling Through Nucleic Acid Sensors and Their Roles in Inflammatory Diseases. *Front. Immunol.* **11**, 625833 (2021).

172. Deweese, J. E., Osheroff, M. A. & Osheroff, N. DNA Topology and Topoisomerases: Teaching a "Knotty" Subject. *Biochem. Mol. Biol. Educ.* **37**, 2 (2008).

173. Li, M. & Liu, Y. Topoisomerase I in Human Disease Pathogenesis and Treatments. *Genomics. Proteomics Bioinformatics* **14**, 166–171 (2016).

174. Pommier, Y., Sun, Y., Huang, S. Y. N. & Nitiss, J. L. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol. 2016 1711* **17**, 703–721 (2016).

175. Promonet, A. *et al.* Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nat. Commun.* **11**, 1–12 (2020).

176. Hodson, C. *et al.* Branchpoint translocation by fork remodelers as a general mechanism of R-loop removal. *Cell Rep.* **41**, 111749 (2022).

177. Shivji, M. K. K., Renaudin, X., Williams, Ç. H. & Venkitaraman, A. R. BRCA2 Regulates Transcription Elongation by RNA Polymerase II to Prevent R-Loop Accumulation. *Cell Rep.* **22**, 1031–1039 (2018).

178. Kim, A. & Wang, G. G. R-loop and its functions at the regulatory interfaces between transcription and (epi)genome. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1864**, 194750 (2021).

179. De, I. *et al.* The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nat. Struct. Mol. Biol. 2014 222* **22**, 138–144 (2015).

180. Schmitzová, J., Cretu, C., Dienemann, C., Urlaub, H. & Pena, V. Structural basis of catalytic activation in human splicing. *Nat. 2023 6177962* **617**, 842–850 (2023).

181. Chen, L. *et al.* The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol. Cell* **69**, 412-425.e6 (2018).

182. 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**, e1002386 (2011).

183. Gómez-González, B. *et al.* Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J.* **30**, 3106–3119 (2011).

184. 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).

185. Daley, J. M. & Sung, P. 53BP1, BRCA1, and the Choice between Recombination and End Joining at DNA Double-Strand Breaks. *Mol. Cell. Biol.* **34**, 1380–1388 (2014).

186. Sy, S. M. H., Huen, M. S. Y. & Chen, J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7155–7160 (2009).

187. Yuan, S. S. *et al.* BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res.* **59**, 3547–51 (1999).

188. Alonso, M. S. M. & Noordermeer, S. M. Untangling the crosstalk between BRCA1 and R-loops during DNA repair. *Nucleic Acids Res.* **49**, 4848–4863 (2021).

189. 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).

190. Zhang, X. *et al.* Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. *Nat. Commun. 2017 81* **8**, 1–12 (2017).

191. Herold, S. *et al.* Recruitment of BRCA1 limits MYCN-driven accumulation of stalled RNA polymerase. *Nat. 2019 5677749* **567**, 545–549 (2019).

192. Stein, H. & Hausen, P. Enzyme from Calf Thymus Degrading the RNA Moiety of DNA-RNA Hybrids: Effect on DNA-Dependent RNA Polymerase. *Science (80-.).* **166**, 393–395 (1969).

193. Hartmann, G. Nucleic Acid Immunity. *Adv. Immunol.* **133**, 121–169 (2017).

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

195. 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).

196. Suzuki, Y. *et al.* An Upstream Open Reading Frame and the Context of the Two AUG Codons Affect the Abundance of Mitochondrial and Nuclear RNase H1. *Mol. Cell. Biol.* **30**, 5123–5134 (2010).

197. 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).

198. Wu, H., Lima, W. F. & Crooke, S. T. Molecular cloning and expression of cDNA for human RNase H. *Antisense Nucleic Acid Drug Dev.* **8**, 53–61 (1998).

199. 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).

200. Lima, W. F. *et al.* Human RNase H1 Uses One Tryptophan and Two Lysines to

Position the Enzyme at the 3′-DNA/5′-RNA Terminus of the Heteroduplex Substrate. *J. Biol. Chem.* **278**, 49860–49867 (2003).

201. 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).

202. Wu, H., Lima, W. F. & Crooke, S. T. Properties of Cloned and Expressed Human RNase H1. *J. Biol. Chem.* **274**, 28270–28278 (1999).

203. 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).

204. Misic, J. *et al.* Mammalian RNase H1 directs RNA primer formation for mtDNA replication initiation and is also necessary for mtDNA replication completion. *Nucleic Acids Res.* **50**, 8749–8766 (2022).

205. Cerritelli, S. M. *et al.* Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol. Cell* **11**, 807–815 (2003).

206. Reyes, A., Rusecka, J., Tońska, K. & Zeviani, M. RNase H1 Regulates Mitochondrial Transcription and Translation via the Degradation of 7S RNA. *Front. Genet.* **10**, 500411 (2020).

207. Reyes, A. *et al.* RNASEH1 Mutations Impair mtDNA Replication and Cause Adult-Onset Mitochondrial Encephalomyopathy. *Am. J. Hum. Genet.* **97**, 186–193 (2015).

208. 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).

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

210. Pereira Bento, F. Towards understanding R-loop regulation by RNase H1 in saccharomyces cerevisiae. (Johannes-Gutenberg University Mainz, 2023). doi:http://doi.org/10.25358/openscience-9113.

211. Arora, M. *et al.* Endogenous Replication Stress in Mother Cells Leads to

Quiescence of Daughter Cells Article Endogenous Replication Stress in Mother Cells Leads to Quiescence of Daughter Cells. *CellReports* **19**, 1351–1364 (2017).

212. 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).

213. Andrs, M. *et al.* Excessive reactive oxygen species induce transcription-dependent replication stress. *Nat. Commun. 2023 141* **14**, 1–15 (2023).

214. Rinaldi, C. & Wood, M. J. A. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat. Rev. Neurol. 2017 141* **14**, 9–21 (2017).

215. Liang, X. H., Sun, H., Nichols, J. G. & Crooke, S. T. RNase H1-Dependent Antisense Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the Nucleus. *Mol. Ther.* **25**, 2075–2092 (2017).

216. Cerritelli, S. M., Sakhuja, K. & Crouch, R. J. RNase H1, the Gold Standard for R-Loop Detection. *Methods Mol. Biol.* **2528**, 91–114 (2022).

217. Smolka, J. A., Sanz, L. A., Hartono, S. R. & Chédin, F. Recognition of rna by the s9.6 antibody creates pervasive artifacts when imaging rna:Dna hybrids. *J. Cell Biol.* **220**, (2021).

218. Wood, M. *et al.* TDP-43 dysfunction results in R-loop accumulation and DNA replication defects. *J. Cell Sci.* **133**, (2020).

219. 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).

220. Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nat. 2014 5117509* **511**, 362–365 (2014).

221. Crossley, M. P. *et al.* Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA–DNA hybrid imaging. *J. Cell Biol.* **220**, (2021).

222. Yan, Q. *et al.* Proximity labeling identifies a repertoire of site-specific R-loop modulators. *Nat. Commun. 2022 131* **13**, 1–17 (2022).

223. 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).

224. 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).

225. Li, J. *et al.* TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat. Methods 2020 174* **17**, 399–404 (2020).

226. Zhang, L. & Elias, J. E. Relative protein quantification using tandem mass tag mass spectrometry. in *Methods in Molecular Biology* vol. 1550 185–198 (Methods Mol Biol, 2017).

227. Borisova, M. E., Wagner, S. A. & Beli, P. Mass Spectrometry-Based Proteomics for Quantifying DNA Damage-Induced Phosphorylation. *Methods Mol. Biol.* **1599**, 215–227 (2017).

228. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc. 2007 28* **2**, 1896–1906 (2007).

229. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc. 2018 141* **14**, 68–85 (2018).

230. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

231. Cox, J. *et al.* Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

232. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. **4**, 207–214 (2007).

233. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nat. Methods 2017 145* **14**, 513–520 (2017).

234. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).

235. Casado, P. *et al.* Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci. Signal.* **6**, 1–14 (2013).

236. Wiredja, D. D., Koyutürk, M. & Chance, M. R. The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* **33**, 3489–3491 (2017).

237. Linding, R. *et al.* NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* **36**, D695–D699 (2008).

238. Lachmann, A. & Ma'ayan, A. KEA: kinase enrichment analysis. *Bioinformatics* **25**, 684–686 (2009).

239. Brionne, A., Juanchich, A. & Hennequet-Antier, C. ViSEAGO: A Bioconductor package for clustering biological functions using Gene Ontology and semantic similarity. *BioData Min.* **12**, 1–13 (2019).

240. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J. & Gevaert, K. Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* **6**, 786–787 (2009).

241. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

242. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

243. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nat. 2024 6308016* **630**, 493–500 (2024).

244. González-Acosta, D. & Lopes, M. DNA replication and replication stress response in the context of nuclear architecture. *Chromosom. 2023 1331* **133**, 57–75 (2023).

245. Lebrec, V., Poteau, M., Morretton, J. P. & Gavet, O. Chk1 dynamics in G2 phase

upon replication stress predict daughter cell outcome. *Dev. Cell* **57**, 638-653.e5 (2022).

246. Minocherhomji, S. *et al.* Replication stress activates DNA repair synthesis in mitosis. *Nature* **528**, 286–290 (2015).

247. Ray Chaudhuri, A. *et al.* Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat. Struct. Mol. Biol. 2012 194* **19**, 417–423 (2012).

248. Segeren, H. A., van Liere, E. A., Riemers, F. M., de Bruin, A. & Westendorp, B. Oncogenic RAS sensitizes cells to drug-induced replication stress via transcriptional silencing of P53. *Oncogene 2022 4119* **41**, 2719–2733 (2022).

249. Wilhelm, T., Said, M. & Naim, V. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. *Genes 2020, Vol. 11, Page 642* **11**, 642 (2020).

250. Palumbieri, M. D. *et al.* Nuclear actin polymerization rapidly mediates replication fork remodeling upon stress by limiting PrimPol activity. *Nat. Commun. 2023 141* **14**, 1–15 (2023).

251. Wu, W. *et al.* Mitotic DNA synthesis in response to replication stress requires the sequential action of DNA polymerases zeta and delta in human cells. *Nat. Commun. 2023 141* **14**, 1–17 (2023).

252. Böhly, N. *et al.* Increased replication origin firing links replication stress to whole chromosomal instability in human cancer. *Cell Rep.* **41**, 111836 (2022).

253. Mazouzi, A. *et al.* A Comprehensive Analysis of the Dynamic Response to Aphidicolin-Mediated Replication Stress Uncovers Targets for ATM and ATMIN. *Cell Rep.* **15**, 893–908 (2016).

254. Beli, P. *et al.* Proteomic Investigations Reveal a Role for RNA Processing Factor THRAP3 in the DNA Damage Response. *Mol. Cell* **46**, 212–225 (2012).

255. Songyang, Z. *et al.* Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* **4**, 973–982 (1994).

256. Macůrek, L. *et al.* Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nat. 2008 4557209* **455**, 119–123 (2008).

257. Brown, V. E. *et al.* CDK2 regulates collapsed replication fork repair in CCNE1-amplified ovarian cancer cells via homologous recombination. *NAR Cancer* **5**, (2023).

258. Valenzisi, P., Marabitti, V., Pichierri, P. & Franchitto, A. WRNIP1 prevents transcription-associated genomic instability. *Elife* **13**, (2024).

259. Okamoto, Y. *et al.* FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress. *FEBS J.* **286**, 139–150 (2019).

260. Liu, M. Y. *et al.* ATR phosphorylates DHX9 at serine 321 to suppress R-loop accumulation upon genotoxic stress. *Nucleic Acids Res.* **52**, 204–222 (2024).

261. Lin, R. *et al.* R-loopBase: a knowledgebase for genome-wide R-loop formation and regulation. *Nucleic Acids Res.* **50**, D303–D315 (2022).

262. Hasanova, Z., Klapstova, V., Porrua, O., Stefl, R. & Sebesta, M. Human senataxin is a bona fide R-loop resolving enzyme and transcription termination factor. *Nucleic Acids Res.* **51**, 2818–2837 (2023).

263. Das, R., Dey, A., Misra, H. S. & Uppal, S. RNase H1 exists as phase-separated assemblies in association with elongating RNA Polymerase II during active transcription. *bioRxiv* 2022.04.08.487625 (2022) doi:10.1101/2022.04.08.487625.

264. Siddiqui-Jain, A. *et al.* CX-4945, an Orally Bioavailable Selective Inhibitor of Protein Kinase CK2, Inhibits Prosurvival and Angiogenic Signaling and Exhibits Antitumor Efficacy. *Cancer Res.* **70**, 10288–10298 (2010).

265. Grygier, P. *et al.* Silmitasertib (CX-4945), a Clinically Used CK2-Kinase Inhibitor with Additional Effects on GSK3β and DYRK1A Kinases: A Structural Perspective. *J. Med. Chem.* **66**, 4009–4024 (2023).

266. Debdab, M. *et al.* Leucettines, a class of potent inhibitors of cdc2-like kinases and dual specificity, tyrosine phosphorylation regulated kinases derived from the marine sponge leucettamine B: Modulation of alternative pre-RNA splicing. *J. Med. Chem.* **54**, 4172–4186 (2011).

267. Ogawa, Y. *et al.* Development of a novel selective inhibitor of the Down syndrome-

related kinase Dyrk1A. *Nat. Commun. 2010 11* **1**, 1–9 (2010).

268. Adayev, T., Wegiel, J. & Hwang, Y. W. Harmine is an ATP-competitive inhibitor for dual-specificity tyrosine phosphorylation-regulated kinase 1A (Dyrk1A). *Arch. Biochem. Biophys.* **507**, 212–218 (2011).

269. Alvarez, M., Altafaj, X., Aranda, S. & De La Luna, S. DYRK1A Autophosphorylation on Serine Residue 520 Modulates Its Kinase Activity via 14-3-3 Binding. *Mol. Biol. Cell* **18**, 1167 (2007).

270. Sridharan, S. *et al.* Systematic discovery of biomolecular condensate-specific protein phosphorylation. *Nat. Chem. Biol. 2022 1810* **18**, 1104–1114 (2022).

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

272. Khosraviani, N. *et al.* Nucleolar Pol II interactome reveals TBPL1, PAF1, and Pol I at intergenic rDNA drive rRNA biogenesis. *bioRxiv* 2023.12.10.570972 (2023) doi:10.1101/2023.12.10.570972.

273. Roukens, M. G., Alloul-Ramdhani, M., Moghadasi, S., Brouw, M. O. den & Baker, D. A. Downregulation of Vertebrate Tel (ETV6) and Drosophila Yan Is Facilitated by an Evolutionarily Conserved Mechanism of F-Box-Mediated Ubiquitination. *Mol. Cell. Biol.* **28**, 4394–4406 (2008).

274. Ring, D. B. *et al.* Selective Glycogen Synthase Kinase 3 Inhibitors Potentiate Insulin Activation of Glucose Transport and Utilization In Vitro and In Vivo. *Diabetes* **52**, 588–595 (2003).

275. MacAulay, K. *et al.* Glycogen Synthase Kinase 3α-Specific Regulation of Murine Hepatic Glycogen Metabolism. *Cell Metab.* **6**, 329–337 (2007).

276. Hoeflich, K. P. *et al.* Requirement for glycogen synthase kinase-3β in cell survival and NF-κB activation. *Nat. 2000 4066791* **406**, 86–90 (2000).

277. Tirman, S. *et al.* Temporally distinct post-replicative repair mechanisms fill PRIMPOL-dependent ssDNA gaps in human cells. *Mol. Cell* **81**, 4026-4040.e8 (2021).

278. Lezaja, A. *et al.* RPA shields inherited DNA lesions for post-mitotic DNA synthesis. *Nat. Commun. 2021 121* **12**, 1–15 (2021).

279. Lamm, N. *et al.* Nuclear F-actin counteracts nuclear deformation and promotes fork repair during replication stress. *Nat. Cell Biol.* **22**, 1460–1470 (2020).

280. Koundrioukoff, S. *et al.* Stepwise Activation of the ATR Signaling Pathway upon Increasing Replication Stress Impacts Fragile Site Integrity. *PLoS Genet.* **9**, (2013).

281. Camilla Rega, A. *et al.* High resolution profiling of cell cycle-dependent protein and phosphorylation abundance changes in non-transformed cells. *bioRxiv* 2024.06.20.599917 (2024) doi:10.1101/2024.06.20.599917.

282. Hydbring, P., Malumbres, M. & Sicinski, P. Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nat. Rev. Mol. Cell Biol. 2016 175* **17**, 280–292 (2016).

283. Chi, Y. *et al.* Identification of CDK2 substrates in human cell lysates. *Genome Biol.* **9**, 1–12 (2008).

284. Moses, A. M., Hériché, J. K. & Durbin, R. Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. *Genome Biol.* **8**, 1–14 (2007).

285. Aleem, E., Kiyokawa, H. & Kaldis, P. Cdc2–cyclin E complexes regulate the G1/S phase transition. *Nat. Cell Biol. 2005 78* **7**, 831–836 (2005).

286. Moiseeva, T. N. *et al.* An ATR and CHK1 kinase signaling mechanism that limits origin firing during unperturbed DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 13374–13383 (2019).

287. Recasens, A. *et al.* Global phosphoproteomics reveals DYRK1A regulates CDK1 activity in glioblastoma cells. *Cell Death Discov. 2021 71* **7**, 1–16 (2021).

288. Borisova, M. E. *et al.* p38-MK2 signaling axis regulates RNA metabolism after UV-light-induced DNA damage. *Nat. Commun. 2018 91* **9**, 1–16 (2018).

289. Walter, C. *et al.* Global kinome profiling reveals DYRK1A as critical activator of the human mitochondrial import machinery. *Nat. Commun. 2021 121* **12**, 1–12 (2021).

290. Song, W. J. *et al.* Phosphorylation and Inactivation of Glycogen Synthase Kinase 3β (GSK3β) by Dual-specificity Tyrosine Phosphorylation-regulated Kinase 1A (Dyrk1A). *J. Biol. Chem.* **290**, 2321–2333 (2015).

291. Najas, S. *et al.* DYRK1A-mediated Cyclin D1 Degradation in Neural Stem Cells Contributes to the Neurogenic Cortical Defects in Down Syndrome. *EBioMedicine* **2**, 120–134 (2015).

292. Bekker-Jensen, D. B. *et al.* Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat. Commun. 2020 111* **11**, 1–12 (2020).

293. Sun, K. *et al.* The splicing factor DHX38 enables retinal development through safeguarding genome integrity. *iScience* **26**, 108103 (2023).

294. Li, J. & Kalev-Zylinska, M. L. Advances in molecular characterization of myeloid proliferations associated with Down syndrome. *Front. Genet.* **13**, (2022).

295. Murray, A. *et al.* Dose imbalance of DYRK1A kinase causes systemic progeroid status in Down syndrome by increasing the un-repaired DNA damage and reducing LaminB1 levels. *eBioMedicine* **94**, (2023).

296. Doble, B. W. & Woodgett, J. R. GSK-3: tricks of the trade for a multi-tasking kinase. *J. Cell Sci.* **116**, 1175–1186 (2003).

297. Saha, A. & Deshaies, R. J. Multimodal Activation of the Ubiquitin Ligase SCF by Nedd8 Conjugation. *Mol. Cell* **32**, 21–31 (2008).