**Investigation of the role of USP11 in regulating R-loop homeostasis**

By Mateusz Jan Jurga

Submitted for the degree of Doctor of Philosophy (PhD)

Department of Molecular Biology & Biotechnology

University of Sheffield

September 2019

## Acknowledgements

I would like to express gratitude to my supervisor Prof. Sherif El-Khamisy and co-supervisor Prof. Alastair Goldman for their support and guidance throughout these years. It has been a long journey with peaks and troughs and their input was invaluable.

I would also like to thank members of El-Khamisy group and C06 office for being a friendly and supportive bunch. Namely, I would like to thank Dr Swagat Ray for answering my endless questions; Dr Chunyan Liao for help with nickel pull-down assays; Dr Jude Chiang for providing GFP-RNase H1 construct; Dr Ryan Beveridge for teaching me immunofluorescence technique; Dr Eva Karyka for helping me with a KEAP1 lentivirus; to-be-Dr Julia Morris for testing USP11 plasmids, Dr Ringaile Zaksauskaite for helping with KEAP1 over-expression in aged USP11 sgRNA cell lines, to-be-Dr Nelma Pereira Palminha for testing KEAP1 siRNA, and to-be-Dr Arwa Abugable for conducting immunofluorescence assay in USP11-depleted U-2 0S cells.

I need to thank Nicol Garnon and the two cats for making sure my work-life balance was healthy, for feeding me, and making sure I did not go completely crazy. Finally, I must thank my parents because none of my accomplishments would ever happen without their hard work. I stand on their shoulders.

## Abstract

RNA:DNA hybrids (R-loops) are transcription by-products that carry out both physiological and deleterious roles. Physiological R-loops promote transcription termination and class-switch recombination, while pathological R-loops are linked to genome instability and human disease. R-loop homeostasis is tightly regulated by RNA binding proteins, R-loop-specific nucleases, such as RNase H1 and H2, and R-loop-specific helicases, such as senataxin (SETX) and DHX9. Even though R-loop metabolising enzymes are well characterised in the context of R-loop homeostasis, it is largely unknown how they are regulated in order to support physiological but prevent pathological R-loop functions. I performed an un-biased genetic screen which led to the identification of ubiquitin specific protease 11 (USP11) as a novel R-loop regulator. USP11 depletion triggers R-loop accumulation in nuclear and nucleolar loci, which is dependent on the catalytic activity of USP11. USP11 depletion triggers ubiquitination and subsequent degradation of SETX. Co-immunoprecipitation data suggest that USP11 and SETX interact via the N-terminal domain of SETX. Depletion of SETX or USP11 hypersensitises human cells to camptothecin and triggers R-loop accumulation. These phenotypes are not exacerbated further upon co-depletion of both SETX and USP11. Moreover, USP11 deletion using CRISPR/Cas9 leads to a degradation of SETX and R-loop accumulation. Surprisingly, aged USP11 knockout cells restore SETX levels and R-loop homeostasis, which is concomitant with downregulation of an E3 ubiquitin ligase KEAP1. Therefore, the data described in this thesis suggest that USP11 regulates R-loop homeostasis via controlling SETX proteostasis and that aged CRISPR/CAS9 cell lines are capable of culture adaptation.

Table of Contents

[Acknowledgements 2](#_Toc26445203)

[Abstract 3](#_Toc26445204)

[List of Figures 8](#_Toc26445205)

[Abbreviations 10](#_Toc26445206)

[Chapter 1: Introduction 13](#_Toc26445207)

[1.1. Introduction 13](#_Toc26445208)

[1.2. Transcription-induced DNA breaks 16](#_Toc26445209)

[1.3. R-loop biology 19](#_Toc26445210)

[1.3.1. R-loop formation 19](#_Toc26445211)

[1.3.2. R-loop prevention and resolution 22](#_Toc26445212)

[1.3.2.1. RNA-binding proteins 22](#_Toc26445213)

[1.3.2.2. Ribonuclease H proteins 24](#_Toc26445214)

[1.3.2.3. R-loop helicases 25](#_Toc26445215)

[1.4. R-loops are linked to genome instability 27](#_Toc26445216)

[1.4.1. Single-strand breaks 28](#_Toc26445217)

[1.4.2. Double-strand breaks 32](#_Toc26445218)

[1.5. R-loops control gene expression and regulate epigenetic markers 35](#_Toc26445219)

[1.6. R-loops and human diseases 36](#_Toc26445220)

[1.7. Role of ubiquitin in DNA damage response pathways 37](#_Toc26445221)

[Chapter 2: Materials and Methods 42](#_Toc26445222)

[2.1. Standard Solutions 42](#_Toc26445225)

[2.2. Standard Equipment 45](#_Toc26445226)

[2.3. Mammalian Tissue Culture 46](#_Toc26445227)

[2.3.1. Media 46](#_Toc26445228)

[2.3.2. Cell Splitting 46](#_Toc26445229)

[2.3.3. Cell Counting 47](#_Toc26445230)

[2.3.4. siRNA transfection 47](#_Toc26445231)

[2.3.5. Plasmid transfection 47](#_Toc26445232)

[2.3.6. Double transfection 48](#_Toc26445233)

[2.3.7. siRNA sequences 48](#_Toc26445234)

[2.3.8. Plasmids 48](#_Toc26445235)

[2.4. Molecular biology techniques 49](#_Toc26445236)

[2.4.1. Immunofluorescence 49](#_Toc26445237)

[2.4.2. Generation of stable cell lines 50](#_Toc26445238)

[2.4.3. Microscopy 50](#_Toc26445239)

[2.4.4. Clonogenics 51](#_Toc26445240)

[2.4.5. Western Blotting 51](#_Toc26445241)

[2.4.5.1. Cell lysis 51](#_Toc26445242)

[2.4.5.2. Bradford assay 51](#_Toc26445243)

[2.4.5.3. SDS-PAGE 51](#_Toc26445244)

[2.4.5.4. Loading of samples 52](#_Toc26445245)

[2.4.5.5. Transfer 52](#_Toc26445246)

[2.4.5.6. Blocking 52](#_Toc26445247)

[2.4.5.7. Antibodies 52](#_Toc26445248)

[2.4.6. DRIP 53](#_Toc26445249)

[2.4.7. DRIP (adopted from Sanz and Chedin, 2019) 55](#_Toc26445250)

[2.4.8. Cross-linking Chromatin Immunoprecipitation 56](#_Toc26445251)

[2.4.9. Pull-down assay under denaturing conditions 57](#_Toc26445252)

[2.4.10. GFP pull-down assay under native conditions 58](#_Toc26445253)

[2.4.11. Slot blot 59](#_Toc26445254)

[2.4.12. PCR 59](#_Toc26445255)

[2.4.13. PCR purification kit 60](#_Toc26445256)

[2.4.14. RNA extraction 61](#_Toc26445257)

[2.4.15. cDNA synthesis 61](#_Toc26445258)

[2.4.16. RT-qPCR and qPCR 61](#_Toc26445259)

[2.4.17. Primers 62](#_Toc26445260)

[2.4.18. Bacterial transformation 64](#_Toc26445261)

[2.4.19. Liquid bacterial culture 64](#_Toc26445262)

[2.4.20. Glycerol stock 64](#_Toc26445263)

[2.4.21. Mini-prep 64](#_Toc26445264)

[2.4.22. Midi-prep 64](#_Toc26445265)

[2.4.23. DNA quantification 64](#_Toc26445266)

[2.4.24. DNA agarose gel electrophoresis 64](#_Toc26445267)

[2.4.25. DNA fragment purification 65](#_Toc26445268)

[2.4.26. DNA sequencing 65](#_Toc26445269)

[2.4.27. Gibson assembly 65](#_Toc26445270)

[2.4.28. Statistical analysis 65](#_Toc26445271)

[2.5. Screening procedures 65](#_Toc26445272)

[2.5.1. Master Plates 65](#_Toc26445273)

[2.5.2. DUB Screen reverse siRNA transfection 66](#_Toc26445274)

[2.5.3. Immunofluorescence for the DUB Screen 66](#_Toc26445275)

[2.5.4. High-throughput microscopy 66](#_Toc26445276)

[2.5.5. High-throughput analysis of the DUB siRNA screen 67](#_Toc26445277)

[Chapter 3: DUB siRNA screen to identify new R-loop regulators 68](#_Toc26445278)

[3.1. Introduction 68](#_Toc26445281)

[3.2. Aims and objectives 69](#_Toc26445282)

[3.3. Methodology 69](#_Toc26445283)

[3.4. Results 70](#_Toc26445284)

[3.4.1. Pre-screen pilot experiments 70](#_Toc26445285)

[3.4.2. DUB siRNA screen data 78](#_Toc26445286)

[3.4.3. USP11 validation as an R-loop regulator 80](#_Toc26445287)

[3.5. Discussion 85](#_Toc26445288)

[3.5.1. MG132 treatment leads to R-loop accumulation phenotype 85](#_Toc26445289)

[3.5.2. Evaluation of the DUB siRNA screen 85](#_Toc26445290)

[3.5.3. USP11 might be a novel R-loop regulator 87](#_Toc26445291)

[Chapter 4: USP11 might regulate the stability of senataxin 89](#_Toc26445292)

[4.1. Introduction 89](#_Toc26445294)

[4.1.1. Overview of USP11 89](#_Toc26445295)

[4.1.2. Overview of senataxin 91](#_Toc26445296)

[4.2. Aims and objectives 93](#_Toc26445297)

[4.3. Results 94](#_Toc26445298)

[4.3.1. USP11 and SETX are epistatic 94](#_Toc26445299)

[4.3.2. USP11 might regulate SETX stability post-translationally 99](#_Toc26445300)

[4.3.3. USP11 depletion leads to elevated ubiquitination of the N-terminal domain of SETX 101](#_Toc26445301)

[4.4. Discussion 107](#_Toc26445302)

[4.4.1. USP11 and SETX are found in nucleoli 107](#_Toc26445303)

[4.4.2. USP11 controls stability of SETX through post-translational modifications 108](#_Toc26445304)

[4.4.3. Insights from USP11 and SETX clonogenic cell survival assays 108](#_Toc26445305)

[Chapter 5: USP11 knockout CRISPR cell lines display R-loop accumulation phenotype and increased senataxin turnover 110](#_Toc26445306)

[5.1. Introduction 110](#_Toc26445311)

[5.1.1. CRISPR-Cas9 110](#_Toc26445312)

[5.2. Aims and objectives 111](#_Toc26445313)

[5.3. Results 111](#_Toc26445314)

[5.3.1. Generation of USP11 knockout and USP11-complemented cell lines 111](#_Toc26445315)

[5.3.2. Characterization of early passages of CRISPR cell lines 114](#_Toc26445316)

[5.3.2.1. USP11 catalytic activity is required for normal R-loop homeostasis 114](#_Toc26445317)

[5.3.2.2. USP11 knockout clones display SETX knockdown 119](#_Toc26445318)

[5.3.3. Characterization of late passages of USP11 knockout cells 122](#_Toc26445319)

[5.4. Discussion 127](#_Toc26445320)

[Chapter 6: Possible complementation mechanism in late passages of USP11 knockout cells 129](#_Toc26445321)

[6.1. Introduction 129](#_Toc26445327)

[6.1.1. Overview of KEAP1 129](#_Toc26445328)

[6.2. Aims and objectives 130](#_Toc26445329)

[6.3. Results 130](#_Toc26445330)

[6.3.1. Olaparib sensitivity of USP11-depleted cells is rescued by RNase H1 130](#_Toc26445331)

[6.3.2. KEAP1 depletion rescues the R-loop accumulation phenotype in USP11-depleted cells 133](#_Toc26445332)

[6.4. Discussion 135](#_Toc26445333)

[Chapter 7: General discussion 136](#_Toc26445334)

[7.1. Overview 136](#_Toc26445335)

[7.2. Future directions 139](#_Toc26445336)

[7.3. USP11 as a drug target for novel therapies 142](#_Toc26445337)

[7.4. Summary 143](#_Toc26445338)

[References 144](#_Toc26445339)

[Appendix 170](#_Toc26445340)

# List of Figures

**Chapter 1: Introduction**

* 1. DNA repair mechanisms.
  2. RNA Polymerases induce torsional stress.
  3. Thread-back model of R-loop formation.
  4. R-loops drive single-strand breaks.
  5. R-loops and G-quadruplexes.
  6. Transcription-Replication collisions.
  7. Modes of Ubiquitination.

**Chapter 3: DUB siRNA screen reveals a novel R-loop regulator**

* 1. R-loops co-localise with nucleoli.
  2. CPT-treatment induces transient nucleolar R-loops.
  3. *In vitro* ec-RNase H treatment abolishes R-loops.
  4. MG132 treatment induces R-loops.
  5. Results from S9.6 immunofluorescence DUB siRNA screen.
  6. USP11-depletion triggers R-loop accumulation.
  7. USP11-depletion leads to R-loop accumulation at nucleolar and nuclear loci.
  8. USP11-depletion leads to R-loop accumulation at nucleolar and nuclear loci according to Sanz and Chedin (2019) DRIP protocol.
  9. USP11 localises to R-loop forming loci according to cross-linking ChIP.

**Chapter 4: USP11 might regulate the stability of senataxin**

* 1. eGFP-SETX1-667 and USP11 might interact in nucleoli.
  2. Depletion of USP11 and SETX.
  3. USP11 and SETX are epistatic.
  4. USP11 regulates SETX at a post-translational level.
  5. USP11 depletion leads to enrichment of ubiquitinated eGFP-SETX1-667.
  6. Recombinant USP11 deconjugates ubiquitin from ubiquitinated eGFP-SETX1-667.
  7. USP11 does not remove K48-linked ubiquitin chains from eGFP-SETX1-667.
  8. USP11 and eGFP-SETX1-667 co-immunoprecipitate.

**Chapter 5: USP11 knockout CRISPR cell lines display R-loop accumulation phenotype and increased senataxin turnover**

* 1. USP11 CRISPR knockout clones.
  2. USP11 sgRNA clones display R-loop accumulation phenotype according to immunofluorescence data.
  3. USP11 knockout clones display R-loop accumulation phenotype according to DRIP-qPCR data.
  4. USP11 knockout clones display R-loop accumulation phenotype – raw DRIP-qPCR data.
  5. USP11C318S knockout clones display R-loop accumulation phenotype.
  6. The catalytic activity of USP11 is required to maintain SETX proteostasis.
  7. USP11C318S mutants accumulate ubiquitinated eGFP-SETX1-667.
  8. Aged USP11 sgRNA Clones revert to wild-type R-loop homeostasis.
  9. Aged USP11 sgRNA Clones do not accumulate R-loops according to DRIP data.
  10. Aged USP11 sgRNA Clones over-express senataxin.
  11. SETX mRNA levels do not change over time as USP11 sgRNA clones are passaged.

**Chapter 6: Possible complementation mechanism in late passages of USP11 knockout cells**

* 1. hs-RNase H1 rescues sensitivity to olaparib of USP11-depleted cells.
  2. KEAP1 knockdown rescues R-loop accumulation phenotype of USP11-depleted cells.

**Chapter 7: General discussion**

* 1. Proposed model of R-loops resolution via SETX.

# Abbreviations

(-) – Negative

(+) – Positive

A – Adenine

AGS – Aicardi - Goutières syndrome

AID – Activation Induced Deaminase

ALS – Amyotrophic Lateral Sclerosis

Alt-NHEJ – Alternative NHEJ

ASM – Active Site Mutant

ATM – Ataxia Telangiectasia Mutated

ATR – Ataxia Telangiectasia and Rad3 related

BER – Base Excision Repair

C – Cytosine

*C9orf72* – chromosome 9 open reading frame 72

CFS – Common Fragile Site

ChIP – Chromatin Immunoprecipitation

CHX – Cyclohexamide

CPEO – Chronic progressive external ophthalmoplegia

CpG – Cytosine-Guanine dinucleotide

CPT – Camptothecin

CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats

DDR – DNA Damage Response

DNA – Deoxyribonucleic Acid

DNAP β – DNA polymerase β

DNA-PK – DNA-dependent Protein Kinase

DRIP – DNA:RNA Immunoprecipitation

DSB – Double Strand Break

dsRNA – double-stranded RNA

DUB – deubiquitylase

E1 – ubiquitin-activating enzyme

E2 – ubiquitin-conjugating enzyme

E3 – ubiquitin ligase

ec – *E.coli*

eIF4B – Eukaryotic Initiation Factor 4B

FA – Formaldehyde

FACT – Facilitates Chromatin Transcription

FRDA – Friedreich ataxia

FXN – frataxin

G – Guanine

G4 – G-quadruplex

GFP – Green Fluorescent Protein

GG-NER – Global Genome NER

GQN1 – G Quartet Nuclease 1

gRNA – guide RNA

H2A – Histone 2A

H3K9me2 – demethylated Histone 3 Lysine 9

H3S10P – phoshoporylated Histone 3 Serine 10

HR – Homologous Recombination

hs – *Homo Sapiens*

IF – Immunofluorescence

ITS – Internal transcribed spacer

J2 – anti-dsRNA antibody

K11 – Lysine 11

K48 – Lysine 48

KD – Knock-down

kDa – kilo Daltons

KO – Knockout

mCP – mRNA polyadenylation and cleavage complex

MDS – Myelodysplastic syndrome

MINDY– Motif interacting with Ub-containing novel DUB family

MMR – Mismatch Repair

mRNA – messenger RNA

ncRNA – non-coding RNA

NEM – N-Ethylmaleimide

NER – Nucleotide Excision Repair

NHEJ – Non-homologous End Joining

Ni – Nickel

NLS – Nuclear Localization Signal

NMD – Non-sense Mediated mRNA Decay

OAO2 – Oculomotor Apraxia Type 2

PAM – Protospacer Adjacent Motif

PARP – poly-ADP Ribose Polymerase

PARPi – PARP inhibitor

PCNA – Proliferating Cell Nuclear Antigen

PIKK – Phosphatidylinositol 3-kinase-related Kinase

PPR – Promoter proximal region

PRC – Polycomb Repressive Complex 1

pre-crRNA - precursor CRISPR RNA

PTM – post-translational modification

qPCR – quantitative Polymerase Chain Reaction

rDNA – ribosomal DNA

REZ – R-loop Elongation Zone

RIZ – R-loop Initiation Zone

R-loop – RNA:DNA hybrid

RNA – Ribonucleic Acid

RNAi – RNA interference

RNAP – RNA Polymerase

RNase H – Ribonuclease H

RPA – Replication Protein A

RT – room temperature

S9.6 – anti-R-loop antibody

SCAN1 – Spinocerebellar ataxia with axonal neuropathy

SETX – senataxin

sgRNA – single guide RNA

shRNA – small hairpin RNA

siRNA – small interfering RNA

SMA – Spinal Muscular Atrophy

SMN – survival of motor neuron

SSA – Single Strand Annealing

SSB – Single Strand Break

SSBR – Single Strand Break Repair

ssDNA – single-stranded DNA

SUMO – Small Ubiquitin-like Modifier

tDNA – transfer DNA

T – Thymine

TDP – Tyrosyl-DNA Phosphodiesterase

TGFβ – Transforming Growth Factor β

THO – Transcription Elongation Complex

THO/TREX – Transcription/Export Complex

TLS – Trans-lesion synthesis

TOP – Topoisomerase

tracrRNA – trans-activating crRNA

TRC – Transcription Replication Collision

TR-HR – Transcription-associated Homologous Recombination

tRNA – transfer RNA

Ub – Ubiquitin

USP – Ubiquitin-Specific Protease

UTR – Untranscribed region

UV – Ultraviolet

UVB – Ultraviolet B

WAS – Wiskott-Aldrich Syndrome

WRN – Werner Protein

WS – Werner Syndrome

WT – Wild Type

X-ChIP – cross-linking ChIP

ZUFSP – Zn-finger and UFSP domain proteins

α-AMN – α-Amanitin

γH2AX – phosphorylated histone 2

# Chapter 1: Introduction

## Introduction

Health and survival of all species is dependent on their ability to maintain genetic material intact and pass it to the future generations. Even though DNA is a stable molecule, it can be damaged by an array of sources of either exogenous or endogenous origin (Ciccia and Elledge, 2010). Exogenous DNA damaging agents are to some extent avoidable, as they include ionising and ultraviolet radiations, tobacco smoke, alcohol, and chemotherapy drugs like etoposide and camptothecin (CPT). However, endogenous agents pose a greater threat to DNA, as they are by-products of normal cellular reactions and thus cannot be avoided. For instance, reactive oxygen and nitrogen species, which are by-products of electron-transport chain, can cause DNA adducts and single-strand breaks (SSBs). Transcription can lead to formation of protein-linked DNA breaks and replication can result in DNA mismatches or double-stand breaks (DSBs) (Jackson and Bartek, 2009).

Each of ~3.72 x1013 ­­cells in a human body can experience 49 DNA lesions per minute (Bianconi et al., 2013; Tubbs and Nussenzweig, 2017). Each of these lesions has a potential to trigger tumourigenesis, cell death or fast-ageing. As a result, human cells evolved elaborate DNA damage response (DDR) pathways that recognise and repair different types of DNA lesions. Some of these DDR pathways are briefly outlined below (Fig. 1.1).





**Figure 1.1 DNA repair mechanisms.**

**a)** BER (Base Excision Repair) – G base is damaged to D, which is excised to abasic site. Next, a nick is generated, filled by a polymerase, and ligated by a ligase. Adapted from David et al., 2007.

**b)** MMR (Mismatch Repair) – upon detection of a mismatch, an endonuclease generates a nick, which is extended by an exonuclease and a helicase. The gap is filled and ligated. Adapted from Jiricny, 2006.

**c)** NER (Nucleotide Excision Repair) – upon detection of a large lesion, a helicase unwinds the DNA duplex and endonucleases generate nicks, which lead to formation of a ssDNA stretch. The gap is filled and ligated**.**

**d)** SSBR (Single Strand Break Repair) – upon detection of a DNA nick, DNA ends are processed to restore apriopriate 5’ and 3’ chemistry. Next, the gap is filled and ligated**.**

**e)** NHEJ – upon detection of a DSB, DNA ends are processed to restore appriopriate 5’ and 3’ chemistry. Next, the ends are ligated. Adapted from Ciccia and Elledge, 2010.

**f)** HR – upon detection of a DSB, the DNA ends are resected. Next, recombination between homologous sequences occur. Gaps are filled and ligated. Adapted from Filippo et al., 2008.

Base Excision Repair (BER) removes chemically modified nitrogenous DNA bases, which do not support standard A-T (adenine – thymine) and G-C (guanine – cytosine) base pairing (Fig. 1.1a). This leads to a subsequent excision of a remaining deoxyribose sugar and reinstallation of an intact nucleotide by a DNA polymerase (David et al., 2007). Mismatch Repair (MMR) is a pathway that ensures fidelity of replication. It recognises and removes misincorporated nucleotides in newly synthesised DNA stretches (Fig. 1.1b). The excision of nucleotides leaves a gap in DNA, which is subsequently filled by DNA polymerase δ. The remaining nick is sealed by DNA ligase I (Jiricny 2006). Nucleotide Excision Repair (NER) recognises and removes large DNA lesions induced by reactive-oxygen species or ultraviolet radiation. NER orchestrates excision of 22-30 nucleotides encompassing the lesion leading to a formation of a gap that is filled by a DNA polymerase (Fig. 1.1c). Global-genome NER subpathway probes genome for lesions that distort the structure of DNA helix, while transcription-coupled NER subpathway ensures smooth progression of RNA polymerases during transcription (Marteijn et al., 2014). Single-strand break repair (SSBR) recognises and repairs the most abundant lesions - single-strand DNA breaks (SSBs) (Fig. 1.1d). Depending on the chemical nature of the break, an appriopriate end-processing factor is recruited to the site of the breakage to restore 5’-phosphate and 3’-hydroxyl group at DNA termini. Once the 5’ and 3’ chemistry is correct, the gap is filled by DNA Polymerase β and the leftover nick is closed by DNA ligases (Caldecott, 2008).

If two SSBs are in near proximity, a double-strand break (DSB) is formed. Although DSBs are relatively rare, they are highly toxic and pose a serious threat to genome stability. They lead to a dissociation of broken DNA termini, which triggers formation of acentric or dicentric (via unscheduled recombination) chromosomal fragments (Jackson, 2002). In order to tackle DSBs, cells evolved four distinct pathways: homologous recombination (HR), non-homologous end-joining (NHEJ), alternative-NHEJ (alt-NJEH), and Single Strand Annealing (SSA). The two most researched DSB repair pathways are HR and NHEJ (Fig. 1.1e-f). Both pathways involve end-processing in order to restore appropriate 5’ and 3’ chemistry at DNA termini. HR uses a sister chromatid as a template for the gap-filling step, which renders this pathway error-free. On the other hand, NHEJ does not facilitate the gap-filling step rendering the process error-prone. Therefore, HR is only employed in cycling cells in G2 and S phases when sister chromatids are present, whereas NHEJ is employed in non-cycling cells, such as neurons, and cells in G1 phase (Ciccia and Elledge, 2010).

As all repair pathways have an ability to not only repair, but also harm DNA, cells need to ensure that lesions are repaired by the most suited pathways. This is accomplished primarily with help of proteins belonging to PIKK and PARP families (phosphatidylinositol 3 kinase-like kinase; poly(ADP)ribose polymerase), which are activated by specific types of DNA lesions (Harper and Elledge, 2007). Exposure to ionizing radiation, reactive oxygen species, or DNA processing by BER may all result in a formation of SSBs. These activate PARP proteins, which orchestrate single-strand break repair process via PARylation of their substrates. PIKK proteins that regulate DDR are ATR, ATM and DNA-PK. They are activated by DSBs and orchestrate appropriate DSB repair pathways via phosphorylation of their substrates. DNA-PK facilitates NHEJ in the absence of sister chromatids, ATM promotes HR in the presence of sister chromatids, while ATR is responsible for repair of stalled replication forks (Ciccia and Elledge, 2010). Moreover, DDR also employs other post-translational modifications like acetylation, methylation, ubiquitination and SUMOylation, as they are fast and reversible mechanisms of signal transduction. They are used to regulate the activation and transition of DNA repair pathways from the first to the final stages of repair (Dantuma and Attikum, 2017).

## Transcription-induced DNA breaks

Transcription can be another potential source of DNA damage. As RNA Polymerases (RNAPs) transcribe DNA into RNA, they create torsional stress. Positive supercoiling is generated downstream of RNAPs, while negative supercoiling is generated upstream (Fig. 1.2a). In order to restore normal chromatin supercoiling and facilitate the progression of RNAPs, a class of enzymes called topoisomerases cleave and re-ligate DNA strands to remove the torsional stress. Human cells express two types of topoisomerases – TOP1 and TOP2. TOP1 cleaves a single strand of DNA to form a transient SSB, while TOP2 cleaves both DNA strands to form a transient DSB. Upon cleavage, TOP1 becomes covalently bound to 3’ DNA termini, while TOP2 becomes bound to 5’ termini (Bermejo et al., 2007).

The topoisomerase-DNA intermediate can be stabilised by endogenous and exogenous agents. Proximal DNA lesions and collisions with a transcription bubble or replication machinery can abort normal topoisomerase cleavage & re-ligation cycle, leading to a stabilisation of the topoisomerase-DNA intermediate. As a result, topoisomerase-linked DNA breaks are formed (Fig. 1.2b). (Alagoz et al., 2013).

Exogenously, topoisomerase-DNA intermediates can be stabilised by a drug called camptothecin (CPT), and its derivatives used in chemotherapy. CPT disturbs topoisomerase cleavage & re-ligation cycle by stabilising the topoisomerase-DNA intermediate after the cleavage step but before the re-ligation (Fig. 1.2b) (Veloso et al., 2013). As a result, topoisomerase-linked DNA break is formed, which leads to genome instability (Householder et al., 2015). Human patients who cannot resolve topoisomerase-DNA intermediates were found to develop neurodegenerative disorders manifesting themselves in ataxias, seizures, cognitive defects and axonal neuropathies (Takashima et al., 2002; Gomez-Herreros et al., 2014). Furthermore, transcription can cause genome instability not only through promoting topoisomerase-linked DNA breaks, but also through formation of RNA:DNA hybrids (R-loops). As R-loops are the subject of this thesis, their formation and links to genome instability & human disease are described in detail the next section.





**Figure 1.2 RNA Polymerases induce torsional stress.**

**a)** Advancing RNA Polymerase generates positive (+) supercoiling downstream of the polymerase and negative (-) upstream. TOP1 resolves the positive supercoiling, while TOP2 resolves the negative supercoiling.

**b)** Cleavage/re-ligation cycles of topoisomerases can be distorted by nearby lesions or topoisomerase inhibitors like CPT. As a result, TOP1 and TOP2 might become covalently trapped to DNA. In order to be released, these intermediates need to be processed by TDP1 or TDP2.

## R-loop biology

R-loops are natural by-products of transcription that require tight regulation (Fig. 1.3a). Gathered evidence suggest that physiological R-loops promote class-switch recombination, transcription termination, transcription-associated homologous recombination (Yo et al., 2003; Skourti-Stathaki et al., 2011; Yasuhara et al., 2018), modify gene expression (Sun et al., 2013), regulate DNA methylation (Ginno et al., 2013), and stabilise chromatin loops (Pezone et al., 2019). On the other hand, pathological R-loops are strongly linked to genome instability in *Saccharomyces cerevisiae* and human cells (Huertas & Aguilera, 2003; Li & Manley, 2005; Tuduri et al., 2009; Wahba et al., 2011; Sollier et al., 2014; Garcia-Rubio et al., 2015, Chuang et al., 2019).

## R-loop formation

Structural studies of transcription bubble suggested a “thread-back model” where a nascent messenger RNA (mRNA) strand invades DNA duplex upstream of the RNA polymerase and anneals to the template strand (Westover et al., 2004) (Fig. 1.3a). This leads to a formation of an RNA:DNA hybrid (R-loop) and displacement of the non-template DNA strand, which remains unannealed (Skourti-Stathaki and Proudfood, 2014). R-loop formation is a natural by-product of transcription and indeed, it was estimated that R-loops cover up to 5% of a human genome at any time (Wongsurawat et al., 2011; Chakraborty et al., 2018). As RNA polymerases advance on DNA during transcription, they trigger positive supercoiling downstream and negative supercoiling upstream of the transcription bubble (Liu and Wang, 1987). Negative supercoiling stalls the transcription bubble in S.*cerevisiae* (Joshi et al., 2010), which in turn promotes R-loop formation by giving more time to mRNA to successfully invade DNA duplex. It was demonstrated in S.*cerevisiae* that negative supercoiling is resolved by Top1p (French et al., 2011). In line with this, Top1p-deleted cells were shown to display R-loop accumulation phenotype in highly transcribed genes (El-Hage et al., 2010). Positive supercoiling was shown to trigger melting of double-stranded DNA (Vologodskii et al., 1992) and hence, might promote mRNA invasion and nucleation according to current statistical models (Stolz et al., 2019). For an R-loop to be formed, the annealing of mRNA and the template DNA must be thermodynamically favourable over the reannealing of the DNA duplex. Therefore, R-loop-forming sequences were analysed and split into R-loop Initiation Zones (RIZs) & R-loop Elongation Zones (REZs) (Fig. 1.3b). It was found that guanine-rich (G) non-template strands within RIZ promote R-loop formation, while GGGG cluster(s) within REZ promote nucleation events that elongate the R-loop (Roy et al., 2008, Roy and Lieber, 2009). In line with these findings, it was suggested that G-clustering promote class-switch recombination, which strengthened the role of R-loops in this process (Zhang et al., 2014).

Furthermore, it was demonstrated *in vitro* that R-loop formation can occur in the absence of G-clusters if a ssDNA nick is present, suggesting that SSBs might induce R-loop formation *in vivo* (Roy et al., 2010) (Fig. 1.3c). Alternatively, SSBs could stall RNA polymerases increasing the chances for R-loop formation. To date, most of the studies described R-loop formation in *cis*, where the R-loop is formed co-transcriptionally by a nascent mRNA. However, R-loops can also be generated in *trans*. Kogoma lab provided evidence that RecA can facilitate R-loop formation post-transcriptionally, i.e. in *trans*, in *E.coli* (Kasahara et al., 2000). Next, Koshland lab demonstrated that Rad51, *Saccharomyces cerevisiae* homologue of RecA, can promote R-loop formation in *trans* as well (Fig. 1.3d) (Wahba et al., 2013). The authors demonstrated that *trans* R-loops triggered genome instability, which was suppressed by over-expression of R-loop nuclease – RNase H (Wahba et al., 2013). Based on evidence from *E.coli* and *S.cerevisiae*, it is plausible to say that human cells might also have a capacity to form *trans* R-loops. Since human genome is highly repetitive, it might contain multiple loci that support RNA nucleation and subsequent R-loop formation *in trans*. Indeed, recently Arab et al. (2019) demonstrated that a long non-coding RNA TARID creates an R-loop in *trans* over *TCF21* gene in human cells.



**Figure 1.3. Thread-back model of R-loop formation.**

**a)** Structural model of R-loop formation. A nascent messenger RNA (mRNA) strand invades DNA duplex upstream of the RNA polymerase and anneals to the template strand. Due to R-loop formation, non-template strand is displaced (ssDNA). Adapted from Skourti-Stathaki & Proudfoot, 2014.

**b)** R-loop initiation zone (RIZ) and R-loop elongation zone (REZ) are found on the DNA non-template strands. RIZ sequences contains GGGG cluster(s), while REZ sequences are G-rich.

**c)** R-loop formation *in trans* might be promoted *in vitro* by ssDNA nicks (Roy et al., 2013). ncRNA – non-coding RNA

**d)** Rad51 facilitates R-loop formation *in-trans* *in vivo* (Wahba et al., 2013).







## R-loop prevention and resolution

Perhaps unsurprisingly, harmful effects of R-loops led to evolution of an array of proteins that resolve them or prevent their formation. These proteins are described below.

## RNA-binding proteins

Only protein-unbound RNA is able to invade DNA and form an R-loop. Therefore, prokaryotic transcription-coupled translation does not permit R-loop formation, as nascent mRNAs are immediately covered with proteins (Gowrishankar and Harinarayanan, 2004). In eukaryotic cells however, a nuclear membrane spatially separates transcription from translation, so separate mechanisms had to evolve to cover RNA strands and prevent R-loop formation.

In their work, Chavez and Aguilera described *hpr1Δ* strain of *S.cerevisiae* where *hpr1Δ* cells displayed transcription elongation defects and genome instability (Chavez and Aguilera, 1997). Next, they demonstrated through a series of pull-down experiments that Hpr1p belonged to a transcription elongation complex THO. Knockout of any THO factor was epistatic with *hpr1Δ* (Chavez et al., 2000). Two years later, THO complex was co-purified with mRNA export factors Yra1 & Sub2, and subsequently renamed as THO/TREX complex (transcription/export) (Sträßer et al., 2002). Finally, Aguilera lab provided evidence showing that THO/TREX mutants can be rescued by over-expression of R-loop specific nuclease – RNase H1 (Ribonuclease H1), linking R-loops to RNA-binding proteins and genome instability (Huertas and Aguilera, 2003). Further research into *Saccharomyces cerevisiae* demonstrated that deletions of genes involved in transcription termination, RNA export and RNA degradation triggered chromosomal instability, which was reduced by RNase H1 over-expression in most of the cases (Wahba et al., 2011). Therefore, the authors suggested that RNA-binding proteins guard *S.cerevisiae* genome from a pathological R-loop accumulation and subsequent genome instability.

Hieter lab conducted a screen in *S.cerevisiae* to identify novel factors in R-loop homeostasis (Stirling et al., 2012). They employed an anti-R-loop antibody, S9.6, to visualise R-loops and used RNase H to validate the signal. The screen yielded seven subunits of the mRNA polyadenylation and cleavage (mCP) complex. Depletions of these factors triggered R-loop accumulation. One of them, Fip1p, is particularly interesting as it is an orthologue *of* *Homo sapienshuman* FIP1L1 protein, which is mutated in eosinophilic leukaemia. Accordingly, siRNA depletion of FIP1L1 in HCT116 cells triggered a 3-fold enrichment of DNA damage markers (Stirling et al., 2012). Therefore, it is plausible to hypothesise that eosinophilic leukaemia might be partially caused by pathological R-loop accumulation, due to loss-of-function mutations of FIP1L1.

In agreement with *S.cerevisiae* data presented above, Li and Manley demonstrated i*n* vivo and *in* vitro that mammalian cells also accumulate R-loops when RNA-binding proteins are depleted. First, they showed that chicken DT40 cells display genome instability phenotype upon depletion of RNA splicing factor ASF1/SF2. Next, they rescued the ASF1/SF2-depleted cells by over-expressing RNase H1, suggesting R-loop accumulation as the cause of the genome instability (Li and Manley, 2004).

SLU7, a marker of hepatocarcinoma, is a splicing regulator that is essential for normal liver homeostasis (Elizalde et al., 2014). Recently, SLU7 was demonstrated to indirectly maintain wild-type splicing patterns of the R-loop regulator ASF1/SF2 (Jimenez et al., 2019). The authors showed that upon depletion of SLU7, ASF1/SF2 was mis-spliced and downregulated in three human cell lines, which in turn triggered R-loop accumulation and DNA damage. Data gathered from a genome-wide siRNA screen demonstrated that down-regulation of 86 mRNA processing factors increased γH2AX staining, which is an early marker of DNA damage, in human HeLa cells (Paulsen et al., 2009). Interestingly, DNA damage phenotypes triggered by seven, out of 86, mRNA processing factors were reduced by RNase H over-expression. The seven factors were: proposed RNA helicase aquarius; splicing factors Crnkl1, Cdc40, and Skiip; small nuclear ribonucleoproteins Srnpa1, Snrpd1, and Snrpd3 (Paulsen et al., 2009). The role of aquarius in R-loop homeostasis was later showed by Sollier et al., (2014). Palancade lab noticed that intronless genes accumulate R-loops, while intron-rich genes are protected from R-loop accumulation (Bonnet et al., 2017). They demonstrated that introns recruit the spliceosome, which inhibits R-loop formation by covering the nascent mRNA (Bonnet et al., 2017).

The research presented in this subsection supports the notion that RNA-binding proteins inhibit R-loop formation by covering mRNA. However, there might be an additional method of R-loop prevention, which is described below (Gonzalez-Aguilera et al., 2008). THSC complex facilitates rapid mRNA export to cytoplasm, as it bridges transcription machinery and nuclear pores. *S.cerevisiae* knockout strains of THSC complex display a genome instability phenotype, which was reduced by over-expression of RNase H1 (Gonzalez-Aguilera et al., 2008). Therefore, THSC complex prevents R-loop formation by spatially separating RNA and DNA strands, which might be crucial for R-loop formation *in trans*.

## Ribonuclease H proteins

In addition to mechanisms that prevent R-loop formation, cells evolved proteins that resolve already existing R-loops. These are R-loop-specific nucleases and helicases. This subsection will discuss the nucleases.

Ribonuclease H (RNase H) enzymes specifically bind to R-loops and cleave the RNA moiety (Stein and Hauser, 1969). Usually two distinct RNase H enzymes are found in eukaryotic and prokaryotic species. RNase H1 and H2 are found in eukaryotes. RNase H1 is a monomer that only cleaves R-loops that are longer than 4 base pairs in nucleus and mitochondria. RNase H2 however is a heterotrimeric protein that mainly excises misincorporated single ribonucleotides from nuclear DNA (Cerritelli and Crouch, 2009; Wanrooij et al., 2017). Double RNase H1/H2Δ in *S.cerevisiae* displayed a median 10-fold enrichment in nuclear R-loop intensity, supporting the role of ribonucleases in R-loop homeostasis (Wahba et al., 2011). Homozygous *Rnaseh1* deletion was found to be embryonic lethal in mice. Quantitative PCR (qPCR) analysis of mouse mitochondrial DNA revealed a replication failure that led to the developmental arrest by day 9 (Cerritelli et al., 2003). In human patients, chronic progressive external ophthalmoplegia (CPEO) was shown to be caused by a loss-of-function of RNase H1 (Reyes et al., 2015). The authors suggested that R-loop accumulation in mitochondria might result in multiple mitochondrial DNA deletions, leading to a muscle wasting and brain dysfunctions (Reyes et al., 2015). Recently, an *in vitro* study demonstrated that mitochondrial DNA synthesis in humans starts with an R-loop that must be processed by RNase H1 in order to commence the replication process (Posse et al., 2019). Additionally, an *in vivo* study showed that R-loop accumulation occurring in mitochondria can be reduced by RNase H1 over-expression, demonstrating that RNase H1 resolves both nuclear and mitochondrial R-loops (Silva et al., 2018). Kunkel lab gathered data showing that *S.cerevisiae* missing RNase H2 (*rnh201Δ* strain) adapted to the mutation by altering transcription levels of 349 genes by at least 1.5-fold (Arana et al., 2012). Transcription-, stress response-, and genome maintenance-associated genes were among the 349 misregulated genes, strengthening the role of RNase H2 in DNA damage response (Arana et al., 2012). Homozygous *Rnaseh2* deletion, as *Rnaseh1* deletion, leads to mice embryonic lethality by day 11 (Reijns et al., 2012). RNase H2 null mouse embryos were found to possess over a million misincorporated ribonucleotides per cell. This triggered chromosome instability and subsequently cell death (Reijns et al., 2012).

Aicardi - Goutieres syndrome (AGS) is caused by loss-of-function mutations within subunits of RNase H2. Cerritelli lab provided evidence that R-loops do not trigger the AGS phenotype, using *S.cerevisiae* as a model organism (Chon et al., 2013). However, accumulation of ribonucleotides within the human genome was shown to trigger over-expression of Interferon α, which led to a chronic inflammation (Rabe, 2013). As a result, AGS patients are born with severe mental and physical impairments (Rice et al., 2007).

Additionally, it is worth to mention that RNase H enzymes are commonly used to validate R-loop signals and are considered a reliable control for R-loop-based studies (Li and Manley, 2004; Paulsen et al., 2009; Stirling et al., 2012; Sanz et al., 2016; Silva et al., 2018).

## R-loop helicases

In addition to employing R-loop nucleases, eukaryotic cells also use R-loop specific helicases to tackle R-loop accumulation. Zakian lab provided evidence that *S.cerevisiae* Pif1p helicase favourably unwinds forked R-loops *in vitro* (Boule and Zakian, 2007). Then, they showed that Pif1p binds to tRNA genes (tDNA) *in vivo*, where it unwinds pathological R-loops formed during RNA Polymerase III transcription (Tran et al., 2017). As Pif1p is highly conserved from yeast to humans (Bochman et al., 2010), it was suggested that human Pif1 might also be unwinding R-loops at tRNA genes (Tran et al., 2017). Therefore, future studies await to establish the role of human Pif1 in R-loop homeostasis and its potential role in human disease.

*S.cerevisiae* Sen1p, a putative helicase, was shown to be essential for transcription termination of small nucleolar and nuclear genes (Steinmetz et al., 2001). As the molecular role of Sen1p in transcription termination remained unknown, Mischo et al., (2011) carried out a series of immunoprecipitation experiments to characterise Sen1p. They employed S9.6 immunoprecipitation protocol (DRIP) followed by quantitative PCR to show that loss of Sen1p triggered R-loop accumulation at RNA Polymerase II – transcribed loci. Next, a follow-up study revealed how human Sen1p homologue, senataxin (SETX), is linked to transcription termination and R-loop homeostasis in HeLa cells (Skourti-Stathaki et al., 2011). It was found that RNA Polymerase II transcription termination occurs over G-rich pause sites where R-loop formation is favourable. R-loops are next resolved by SETX in a way that the released mRNA can be cleaved by a 5’-3’exonuclease XRN2 at its 3’ poly(A)site. As a result, RNA Polymerase II is released and the transcription bubble terminates (Skourti-Stathaki et al., 2011). Notably, mutated *SETX* was associated with neurodegenerative disorders such as a juvenile type of amyotrophic lateral sclerosis (ALS4) and ataxia with oculomotor apraxia type 2 (OAO2) (Chen et al., 2004; Moreira et al., 2004). ALS4 is an autosomal dominant disorder characterised by missense gain-of-function mutations of *SETX*, whilst AOA2 is an autosomal recessive disorder characterised by non-sense loss-of-function mutations of *SETX* (Arning et al., 2013, Bennett et al., 2013). The AOA2 mutations of SETX tend to cluster within its N-terminal protein interaction and C-terminal R-loop helicase domains, while ALS4 mutations are spread across the gene body (Arning et al., 2013).

Interestingly, one of the possible roles of R-loops in ALS4 neuropathy was recently identified (Grunseich et al., 2018). It was shown that ALS4 patients with a specific heterozygous gain-of-function SETXL389S mutation exhibit fewer R-loops. As a result, DNA loci that normally form R-loops remain unmasked and can be bound by proteins. Consequently, Grunseich and colleagues demonstrated that in ALS4 cells, DNA-methyltransferase 1 is able to bind to and methylate a promoter region of *BAMBI*, which leads to a suppression of its transcription. As BAMBI is a negative regulator of TGF-β pathway, its reduction triggers the activation of the pathway. Additionally, the authors conducted a genome-wide analysis in ALS4 cells to find that R-loop reduction triggers methylation and subsequent gene silencing of 1200 promoters (Grunseich et al., 2018). Therefore, genome-wide transcription alteration caused by a gain-of-function SETX in ALS4 might, at least partially, contribute to the development of the disorder.

DHX9 is another human helicase that was shown to preferentially unwind R-loops and G-quadruplexes *in vitro* (Chakraborty and Grosse, 2011). Erkizan et al., (2015) demonstrated *in vitro* that the activity of DHX9 is reduced by Ewing sarcoma oncoprotein - EWS-FLI1. Therefore, the authors hypothesised that growth of Ewing sarcoma might be partially dependent on reduced activity of DHX9 (Erkizan et al., 2015). In line with this, Gorthi et al., (2018) demonstrated *in vivo* that Ewing sarcoma cells accumulate R-loops, which sequester BRCA1 and subsequently abolish homologous recombination pathway. However, the authors did not assess the role of DHX9 in R-loop homeostasis in these cells (Gorthi et al., 2018). In a recent study, Chakraborty et al., (2018) demonstrated that DHX9 unwinds self-folded nascent mRNAs in order to facilitate loading of splicing factors onto the mRNA. However, when the splicing factors are depleted, DHX9 helicase activity promotes R-loop formation, as the unfolded mRNA is able to hybridise with the template strand. On the other hand, when only DHX9 is depleted in otherwise unperturbed cells, the global R-loop levels are reduced according to two studies (Chakraborty et al., 2018; Cristini et al, 2018). Therefore, DHX9 is considered to promote formation of physiological R-loops in healthy cells and pathological R-loops in disorders where RNA processing is disregulated (Chakraborty et al., 2018). Interestingly, Cristini et al., (2018) also gathered evidence that upon CPT-treatment, DHX9 suppresses R-loop formation and promotes transcription termination, mirroring the role of SETX. Thus, future studies await to discern the roles of DHX9 in either promotion or resolution of R-loops.

FANCM is an ATPase/translocase that was shown to unwind R-loops *in vitro* (Schwab et al., 2015). Recently, Silva et al., (2019) added a layer of complexity to FANCM-related research by demonstrating that FANCM resolves telomeric R-loops *in-vitro­­* and FANCM-depletion triggers accumulation of telemetric R-loops in U2-OS cells (Silva et al., 2019). Unfortunately, the authors did not provide data on whether FANCM specifically or preferentially resolves telomeric R-loops.

Finally, three RNA DEAD-box helicases, DDX5, DDX19 and DDX21, were linked to R-loop homeostasis (Mersaoui et al., 2019; Hodroj et al., 2017; Song et al., 2017). Upon DNA damage, DDX19 was shown to transiently translocate from nucleopores to the nucleus to resolve pathological R-loops. *In vitro*, DDX19 unwinds R-loops via its helicase activity; and *in vivo*, DDX19-depletion triggers R-loop accumulation and genome instability (Hodroj et al., 2017). DDX21 was demonstrated to resolve R-loops *in vitro*. However, its activity was dependent on its acetylation status. Permanently acetylated DDX21 was shown to be inactive and unable to resolve R-loops i*n vivo*, triggering R-loop accumulation phenotype and subsequent genome instability (Song et al., 2017). Moreover, DDX21 was shown to be recruited to a transcription start site of *ENPP2* to remove R-loops and facilitate transcription. The recruitment was facilitated by a demethylase JMJD3 in a non-enzymatic manner (Argaud et al., 2019). DDX5 was shown *in vitro* and *in vivo* to unwind R-loops at transcription termination sites. Once liberated, the RNA moiety was degraded by an exoribonuclease XRN2 in a manner resembling SETX-XRN2 interaction. This degradation event was essential as XRN2-depleted cells displayed R-loop accumulation phenotype. Additionally, XRN2 recruitment to DDX5 was dependent on methylation status of DDX5 (Mersaoui et al., 2019), suggesting a role for post-translational modifications in fine-tuning R-loop homeostasis.

## R-loops are linked to genome instability

Even though R-loops do not directly damage DNA, they are strongly linked to genome instability according to studies described below. Due to R-loop formation, the displaced non-template strand is left unannealed and therefore vulnerable to DNA damaging agents. Additionally, the R-loop structures stall transcription bubbles, leading to deleterious collisions between the RNA and DNA polymerases. In line with this, abnormal expression of R-loop binding proteins was demonstrated to correlate with cancer prognosis and drug response (Boros-Olah, et al., 2019). Therefore, the next two subsections will review the current state of knowledge on R-loops and their contribution to genome instability.

## Single-strand breaks

Transcription-associated mutagenesis was found to usually target single, non-template DNA strands (Beletskii and Bhagwat, 1996), suggesting that ssDNA is more prone to damage than the DNA duplex (Hamperl and Cimprich, 2014). Therefore, numerous studies focused on the consequences of ssDNA displacement concomitant with R-loop formation.

Currently, the consensus among scientists is that the R-loop formation in B-cells promotes the first step of class-switch recombination – the AID (Activation Induced Deaminase) attack on ssDNA (Fig. 1.4a) (Muramatsu et al., 2000; reviewed by Casellas et al., 2016). Dickenson et al., (2003) found that AID deaminates cytosines to uracils in the ssDNA, which are then excised by uracil DNA glycosylases to abasic sites (Shalhout et al., 2014). Abasic sites are further repaired by single-strand break repair pathway. Perhaps unsurprisingly, increased protein levels of AID were linked to oncogenesis in B cells (Shalhout et al., 2014). As AID expression is confined to B-cells, the research focus shifted to other cytosine deaminases like APOBEC3B, which was shown to cause cytosine to uracil deaminations at estrogen receptor-responsive genes (Periyasamy et al., 2015). Therefore, future research might shed the light on the link between cytosine deaminases, R-loops and DNA damage.



**Figure 1.4. R-loops drive single-strand breaks.**

**a)** AID/APOBEC family of enzymes deaminates cytosines to uracils, which are then recognised as substrates for BER pathway. A BER enzyme, DNA glycosylase, excises uracils to abasic sites.

G-quadruplexes (G4) were shown to promote R-loop formation and contribute to genome instability (De Magis et al., 2018). G4s are DNA loops comprised of 4 guanines, which are bonded by Hogsteen base pairing (Lam et al., 2013; Bochman et al., 2012) (Fig. 1.5a). Electron microscopy studies provided evidence that G4s and R-loops are found at the same loci. G4s form on non-template strands, while R-loops occupy the corresponding template strands (Fig. 1.5b) (Duquette et al., 2004). Therefore, R-loops and G4s might stabilise each other and together contribute to genome instability (De Magis et al., 2018). G4s are substrates for, among others, G quartet nuclease 1 (GQN1), which specifically cleave G4 structures to SSBs (Fig. 1.5c) (Sun et al., 2001). GQN1 nucleases are found to be elevated in B-cells and cervical adenocarcinoma, suggesting a role in class-switch recombination and tumourigenesis (Sun et al., 2001). In line with this, *S.cerevisiae* G4-nuclease Kem1p was shown to promote telomere maintenance, as homozygous *kem1* deletion triggered telomere shortening (Liu et al., 1995). Therefore, future studies await to investigate the role of G4-specific nucleases in promoting R-loop-associated genome instability in eukaryotic cells. Moreover, both G4s and R-loops were shown to be detected by Replication Protein A (RPA) through the displaced ssDNA (Fig. 1.5d) (Prakash et al., 2011; Kabeche et al., 2018). As RPA was shown to recruit AID *in vitro* (Chaudhuri et al., 2004), it might contribute to genome instability in B-cells. Consequently, future studies *in vivo* might shed light on whether RPA recruits cytosine deaminases to R-loops in various cell lines.



**Figure 1.5. R-loops and G-quadruplexes (G4s).**

**a)** G4s on a ssDNA. Guanines are paired by Hogsteen base pairing.

**b)** G4s and R-loops are found within the same locus. R-loops occupy template strands whereas G4s occupy non-template strands.

**c)** GQN1 endonucleases cleave 5’ ssDNA upstream of G4s, leading to formation of a SSB.

**d)** RPA specifically binds to the displaced ssDNA. Next, RPA recruits AID, which deaminates cytosines to uracils within the ssDNA. As a result, DNA glycosylase excises newly formed uracils to abasic sites.





## Double-strand breaks

Double-strand breaks (DSBs) are formed when two single-strand breaks are in proximity (Stavnezer et al., 2008). Therefore, as SSBs stall RNA polymerases, which in turn promotes R-loop formation, R-loops might contribute to DSB formation.

As R-loops are unusual three-stranded nucleic acid structures, it was shown i*n vitro* that structure-specific nucleases, XPF-ERCC1 and XPG, can cleave double-stranded DNA in the proximity of R-loops (Tian and Alt, 2000). Later, these findings were recapitulated *in vivo* by Sollier et al., (2014). The authors found that unprocessed R-loops formed during transcription are converted to double-strand breaks by the NER factors XPF and XPG. Therefore, transcription-coupled NER might drive genome instability when pathological R-loops accumulate. Furthermore, a growing body of evidence implies that R-loops take part in DSB formation via inducing replication stress (Hamperl and Cimprich, 2014). Current state of knowledge on this topic is reviewed below. Evidence gathered by Zakian lab suggests that DNA polymerase 2 stalls at the open reading frames of highly transcribed genes. Crucially, the stalling was found to be transcription-dependent (Azvolinsky et al., 2009), suggesting a possible role for R-loops in replication stress. In line with this, TOP1-depleted cells were demonstrated to exhibit slow replication fork progression and DSB accumulation at highly transcribed genes. RNase H1 over-expression reinstated the wild-type phenotype in TOP1-depleted cells, implying the role of R-loops in causing the replications stress (Tuduri et al., 2009).

Helmrich et al., (2011) demonstrated that transcription of >800kb-long human genes takes more time than a replication of the whole human genome. Therefore, transcription-replication collisions are inevitable. The sites prone to transcription-replication collisions (TRCs) were mapped to specific loci called Common Fragile Sites (CFSs). Notably, the authors showed that R-loops formed at the CFSs and triggered genome instability (Helmrich et al., 2011). In line with this, SETX was shown to localise to replication forks where it facilitates replication progression (Alzu et al., 2012). Herrera-Moyano et al., (2014) gathered data suggesting that *S.cerevisiae* *fact-/-* strain (FAcilitates Chromatin Transcription) exhibits genome instability and replication deficiency phenotypes, which are reduced by either transcription inhibition or RNase H1 over-expression. Consequently, the authors believe that *fact-/-* strain accumulates R-loops, which lead to transcription-replication collisions and eventually to genome instability (Herrera-Moyano et al., 2014). This in line with data obtained from *S.cerevisiae npl3-/-* and *S.pombe* *dcr1Δ* strains (Santos-Pereira et al., 2013; Castel et al., 2014), which are described below.

Npl3 is a heterogenous ribonuclear protein that binds to highly transcribed genes. *npl3-/-* strain exhibits genome instability and replication stress phenotypes, which are reduced by RNase H1 over-expression (Santos-Pereira et al., 2013). Dcr1 is a component of RNA interference pathway, which was shown to facilitate transcription termination of highly transcribed genes including tDNA and rDNA genes. As all these loci were correlated to replication pausing, they are likely to be sites of TRCs. In line with this, *S.pombe* *dcr1Δ* strain was shown to display: 1) R-loop accumulation phenotype at rDNA loci and 2) loss of rDNA copy number; corroborating the notion of R-loops driving replication stress and genome instability (Castel et al., 2014).

TRCs can occur in head-on or co-directional orientations (Fig. 1.6). In *S.cerevisiae*, it was shown that head-on, but not co-directional, collisions significantly contribute to genome instability phenotypes (Prado and Aguilera, 2005; Desphande and Newlon, 1996). However, in human cells, it was demonstrated that most TRCs occur in a co-directional orientation rather than the head-on (Petryk et al.,2016). This led Hamperl et al., (2017) to discern between the effects of co-directional and head-on collisions on human genome. First, they showed that R-loop-forming sequences exhibit TRCs, while R-loop-free sequences did not. Then, they demonstrated that co-directional TRCs resolve R-loops, while head-on collisions stabilise R-loops. The head-on collisions triggered ATR activation and subsequent phosphorylation of H2A.X (γH2A.X), while the co-directional collisions triggered ATM activation, which did not lead to subsequent phosphorylation of H2A.X. The reasons for contrasting H2A.X status are unclear. However, as ATM was activated, the co-directional collisions could also lead to double-strand breaks (Hamperl et al.,2017). Therefore, future studies are required to understand 1) structural intermediates occurring upon a TRC, 2) how R-loops are converted into DSBs during TRC, and 3) what additional factors are implicated in the signalling pathways downstream of TRCs.



**Figure 1.6. Transcription-Replication collisions.**

**a)** Representation of a head-on collision between a replication fork and a transcription bubble.

**b)** Representation of a co-directional collision between a replication fork and a transcription bubble. Adapted from Hamperl & Cimprich, 2014.



## R-loops control gene expression and regulate epigenetic markers

Recently, R-loop accumulation was shown to inhibit looping between an enhancer region and a number of promoters by a chromatin conformation capture assay (Chiang et al., 2019). The authors demonstrated that BRCA1-regulated R-loop at the transcriptional enhancer region upstream of *ESR1* negatively modulates expression of neighbouring genes involved in differentiation. If the R-loop is unresolved, the luminal progenitor cells do not differentiate and may trigger development of a breast cancer (Chiang et al., 2019).

As described in section 1.3.1, G-rich sequences promote R-loop formation (Roy et al., 2008, Roy and Lieber, 2009). Approximately 60% of human promoters are enriched for cytosine-guanine dinucleotides called CpG islands (Ginno et al., 2012). Most of the CpG-containing promoters were found to be unmethylated and enriched for R-loops. As promoter methylation is associated with gene silencing, the authors suggested that R-loops mask dsDNA from methyltransferases at promoter regions and thus facilitate further transcription (Ginno et al., 2012). This is in line with a recent study where R-loops were shown to inhibit DNA methylation of 1200 gene promoters and subsequently promote their transcription (Grunseich et al., 2018). CpG islands are also common over transcription termination pause-sites, where they trigger R-loop formation (Ginno et al., 2013). Interestingly, the termination-associated R-loops recruit G9a histone lysine methyltransferase, which promotes silencing and transcription termination (Skourti-Stathaki et al., 2014).

Recently, R-loops were shown to form over a subset of CpG-rich & Polycomb-repressed developmental genes in mouse embryonic stem cells (Skourti-Stathaki et al., 2019). Interestingly, the authors demonstrated that these R-loops facilitate recruitment of Polycomb Repressive Complex 1 (PRC1), which in turn maintains repressive chromatin marks (monoubiquitinated H2A) over the transcription initiation sites. Upon RNase H1 over-expression, loss of repressive marks and transcriptional derepression were observed, validating the role of R-loops in regulating expression of these genes (Skourti-Stathaki et al., 2019).

Arab et al. (2019) characterised GADD45A protein as a novel reader of *trans* R-loops. The authors suggested that GADD45A binds to a small subsection of *trans* R-loops where it recruits TET1, which in turn facilitates transcriptional derepression by DNA demethylation. Interestingly, the R-loop formation, TET1 recruitment, and derepression were found to be cell-cycle dependent (Arab et al., 2019).

R-loops and repressive histone modifications (H3K9me2) were mapped to the same locus in a neurological disorder Friedreich ataxia (Groh et al., 2014). In line with this, it was shown in HeLa cells, *C.elegans* and *S.cerevisiae* that a repressive phosphorylated histone mark H3S10P coincides with R-loops at actively-transcribed genes, pericentromeric and centromeric regions (Castellano-Pozo et al., 2013). H3S10P mark, which is associated with gene silencing, was reduced by RNase H1 over-expression, implying that R-loops recruit unknown factors, which promote phosphorylation of H3S10 and subsequent gene silencing (Castellano-Pozo et al., 2013). This agrees with a more recent study where it was shown that pathological R-loops promote formation of H3S10P marks (Garcia-Pichardo et al., 2017). The authors demonstrated that inhibition of phosphorylation of H3S10P in sen1p mutant *S.cerevisiae* strain reduced the genome instability phenotype, suggesting that R-loops do not drive DNA damage on their own.

Altogether, the gathered data strongly suggest that physiological R-loops can trigger chromatin modifications to either silence or promote transcription at terminator and promoter regions respectively. Pathological R-loops, however, seem to drive gene silencing and even genome instability through repressive histone marks and aberrant chromatin looping. Therefore, future studies await to shed the light on R-loops and their role in epigenetics and chromatin conformation.

## R-loops and human diseases

Perhaps unsurprisingly, perturbed R-loop homeostasis was shown to be at the root of a few human congenital disorders. Friedreich ataxia (FRDA) and Amyotrophic Lateral Sclerosis (ALS) are two neurological disorders that are characterised by nucleotide repeat expansion-mutations, which promote R-loop formation (Groh et al., 2014; Walker et al., 2017). In case of FRDA, GAA triplet expansion within frataxin gene (*FXN*) causes R-loop accumulation and subsequent silencing of *FXN*, which leads to development of FRDA syndrome (Groh et al., 2014). In majority of ALS cases, GGGGCC repeat expansions in the non-coding *C9orf72* (chromosome 9 open reading frame 72) promote R-loop formation, which was shown to contribute to genome instability phenotype and development of ALS symptoms (Walker et al., 2017).

Werner syndrome (WS) and Spinal Muscular Atrophy (SMA) are characterised by loss-of-function mutations, which alter R-loop homeostasis and trigger subsequent genome instability. WS develops due to a loss of Werner protein (WRN), which is involved in preventing replication stress by stabilising replication forks. In the absence of WRN, WS cells accumulate R-loops upon mild replication stress, which are then processed to DSBs by an endonuclease XPG. As a result, WS cells exhibit genome instability and WS patients are more likely to develop tumours (Marabitti et al., 2019). Spinal Muscular Atrophy develops due to a loss of Survival of Motor Neuron 1 (*SMN1*), which often results in infant mortality by the age of two. Staropoli lab suggested a model where a loss of SMN triggers spliceosomal dysfunction in motor neurons, which leads to R-loop accumulation. Subsequently, pathological R-loops drive genome instability and neuronal cell death (Jangi et al., 2017).

Wiskott-Aldrich Syndrome (WAS) patients exhibit R-loop driven genome instability in their blood and lymph cells (Sarkar et al., 2018). WAS is caused by a loss-of-function mutations of WAS protein (WASp). The authors gathered evidence suggesting that WASp is essential for TOP1 localisation to chromatin, where the latter protein resolves torsional stress. In the absence of WASp, TOP1 is not enriched on the chromatin, consequently transcription bubbles are stalled due to negative supercoiling, and pathological R-loops are formed (Sarkar et al., 2018). Mutations in multiple splicing factors were linked to myelodysplastic syndrome (MDS), which is a blood disorder (Kim et al., 2015; Shirai et al., 2015; Obeng et al., 2016). Recently, Fu and Graubert labs independently showed that MDS cells accumulate R-loops (Chen et al., 2018; Nguyen et al., 2018). Moreover, both groups demonstrated that RNase H1 over-expression partially rescued MDS cells, suggesting that R-loops, to some extent, drive progression of the disease. Fanconi Anemia syndrome is characterised by a loss of the Fanconi Anemia repair pathway that primarily resolves DNA crosslinks (Liang et al., 2019). Fanconi Anemia syndrome patients suffer from congenital abnormalities and are prone to tumourigenesis. Kupfer lab gathered evidence that a component of Fanconi Anemia pathway FANCI-FANCD2 binds to the displaced ssDNA or the unhybridized RNA in the R-loop structure, in order to facilitate swift R-loop resolution. Therefore, the authors concluded that Fanconi Anemia syndrome might be partially driven by its inability to resolve pathological R-loops (Liang et al., 2019).

## Role of ubiquitin in DNA damage response pathways

As described above, eukaryotic cells developed a number of ways to prevent or resolve R-loops in order to maintain correct chromatin marks and genome integrity. However, how the R-loop metabolising enzymes are controlled is mostly unclear. The processes that regulate protein homeostasis include post-translational modifications, for instance: ubiquitination and deubiquitination. Since this thesis describes the role of (de)ubiquitination in fine-tuning R-loop homeostasis, an overview on ubiquitin and related processes is provided below.

Ubiquitin (Ub) is a highly conserved 8.5kDa peptide that is rapidly conjugated or deconjugated from proteins. Enzymes might be mono or polyubiquitinated. Ubiquitin contains seven lysine residues, which can be ubiqutinated too in order to form linear chains or branches of ubiquitin peptides (Sahtoe & Sixma, 2015) (Fig. 1.7). Therefore, ubiquitination is a modular process, which enables a high degree of regulation of various cellular processes including protein degradation, transcription and DDR (Wolberger, 2014). Ubiquitin is conjugated via consecutive actions of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin ligases (E3). There are two E1, 41 E2 and about 750 E3 enzymes in human cells (Hutchins et al., 2013). E3 ligases confer specificity to ubiquitination process and were shown to be essential components of DDR pathways. For instance, E3 ligases RNF8 and RNF168 facilitate cellular response to double-strand breaks (Panier and Duricher, 2009); KEAP1 fine-tunes homologous recombination by ubiquitinating PALB2 (Orthwein et al., 2015); monoubiquitination of FANCI/FANCD2 complex activates the Fanconi Anemia repair pathway; and monoubiquitinated Proliferating Cell Nuclear Antigen (PCNA) promotes trans-lesion synthesis (TLS) while polyubiquitinated PCNA triggers error-free lesion bypass (reviewed by Chen et al., 2011). Notably, as the process of conjugating ubiquitin is reversible, actions of E3 ligases are counteracted by deubiquitinases.



**Figure 1.7. Modes of ubiquitination**

**a)** Monoubiquitination.

**b)** Dimonoubiquitination.

**c)** Mono- and polychain-ubiquitination.

**d)** Branched polyubiquitination.

Deubiquitinases (DUBs) are cysteine or metalloproteases that hydrolyse ubiquitin-protein bonds. DUBs are also able to resolve bonds within ubiquitin peptides to modulate Ub chains and branches. Therefore, the role of DUBs is to regulate the free pool of ubiquitin peptides, degree of proteins’ ubiquitination, and the nature of this process. It was shown that there are 107 DUBs encoded by the human genome (Hutchins et al., 2013), which fall into seven distinct structural subclasses: Machado-Joseph Disease protein domain proteases (JMDs), the JAMM motif proteases (JAMMs), the Otubain proteases (OTUs), the ubiquitin C-terminal hydrolyses (UCHs), motif interacting with Ub-containing novel DUB family (MINDY), Zn-finger and UFSP domain proteins (ZUFSP), and the ubiquitin-specific proteases (USPs) (Nijman et al., 2005, Abdul Rehman et al., 2016, Hermanns et al., 2018). Even though DUBs constitute a rather small protein family, they can be specific for protein substrates and/or for particular Ub modifications (Komander et al., 2009). To ensure high fidelity, deubiquitinases are tightly controlled by post-translational modifications (PTMs) (Kessler & Edelmann, 2011), partner proteins (Kee et al., 2010), proteolytic cleavage (Huang et al., 2006), and redox regulation (Komander et al., 2009).

DUBs are essential for accurate DNA repair mechanisms. USP1 was suggested to be an inhibitor of DNA translesion synthesis (TLS). Upon UV-induced damage, USP1 deubiquitylates its own C-terminal domain, which results in an autocleavage event and degradation of the protein. Proteolysis of USP1 stabilizes levels of monoubiquitinated PCNA (Proliferating Cell Nuclear Antigen protein), which triggers TLS (Huang et al., 2006). USP47 regulates steady-state levels of DNA polymerase β (DNAP β), which is essential for Base Excision Repair (BER) pathway. The knockdown of USP47 was demonstrated to trigger proteolysis of DNAP β and accumulation of DNA breaks upon hydrogen peroxide treatment (Parsons et al., 2011). USP3 and USP16 control cell cycle progression via deubiquitination of histone 2A (H2A) (Nicassio et al., 2007; Joo et al., 2007). Moreover, USP3 inhibits localization of DDR E3 ligase RNF168 to sites of DNA damage, while USP16 was suggested to partially suppress ATM-driven transcriptional induction upon presence of double-stranded breaks (DSBs) (Shanbhag and Rafalska-Metcalf, 2010). USP28 was suggested to be a tumour suppresser, as it is often deleted in cancer cells. Upon DNA damage, USP28 stabilizes CHK2, which modulates activity of DDR transcription factor p53 (Zhang et al., 2006). UCHL3 was shown to control steady-state levels of TDP1. Once deubiquitinated, TDP1 is stabilised and facilitates protein-linked DNA break repair (Liao et al., 2018). Downregulation of UCHL3 was demonstrated to cause TDP1 degradation and subsequent development of a neurological disorder Spinocerebellar ataxia with Axonal Neuropathy (SCAN1) (Liao et al., 2018).

Notably, DUBs were suggested to be relatively more druggable than E3 ligases due to their reactive cysteine residues within active site pockets (D’Arcy et al., 2014). In line with this, DUB inhibitors were developed for use in anti-cancer treatments (McClurg & Robson, 2015), corroborating their role in DDR. However, the role of DUBs in fine-tuning R-loop homeostasis is unknown. Since DUBs can stabilise their target substrates (Zhang et al., 2006; Liao et al., 2018), it is plausible to hypothesise that they could ensure swift prevention and resolution of pathological R-loops.

# Chapter 2: Materials and Methods



## Standard Solutions

10x Tris Buffered Saline (TBS)

24.2 g Tris base and 80 g NaCl dissolved in 900 ml of double distilled water (ddH20). Hydrogen chloride (HCl) or sodium chloride (NaOH) were added to pH 7.6. Then, ddH20 was added to total volume of 1000 ml. 10x TBS was kept at room temperature (RT).

1x TBST

100 ml 10x Tris Buffered Saline (TBS) diluted in 900 ml ddH20 to make 1x TBS. Next, using a cut tip 1 ml Tween® 20 (Sigma) was added to 1x TBS to make 1x TBST, which was was kept at RT.

10x Running Buffer

144 g glycine, 30.3 g Tris Base & 10 g SDS dissolved in ddH20 to make 1000 ml 10x Running Buffer, which was kept at RT.

1x Transfer Buffer

200 ml 5xTrans-Blot® Turbo™ transfer buffer, 200 ml 100% ethanol & 600 ml ddH20. 1x Transfer Buffer was kept at RT.

1M Tris

131.14 g Tris base dissolved in 900 ml ddH20. HCl or NaOH were added to pH 6.8 or 8.8. Then, ddH20 was added to total volume of 1000 ml. 1M Tris solutions were kept at RT.

10% (w/v) Ammonium Persulphate (APS)

APS powder was dissolved 10 times (w/v) in ddH20. 400 μl aliquotes were kept at -20ºC.

10% (w/v) Sodium Dodecyl Sulphate (APS)

SDS powder was dissolved 10 times (w/v) in ddH20 and kept at RT.

10x Tris-Borate-EDTA (TBE) buffer

108 g Tris base, 55 g Boric acid & 40 ml 0.5M EDTA were dissolved in total volume of 1000 ml ddH2O. The solution was kept at RT.

1x Phosphate Buffered Saline (PBS)

A PBS tablet added to 500 ml ddH20 % autoclaved at 121ºC for 20 min. Next, kept at RT.

0.25% (w/v) Trypsin

1g Trypsin 1:250 Powder (Gibco, #27250-018) in 20 ml PBS & filter sterlised. Next, 380 ml autoclaved PBS was added and kept at 4 ºC.

4% (w/v) Ethylenediaminetetraacetic Acid (EDTA)

27.397 ml 0.5 M filter sterilised EDTA in autoclaved 72.603 ml ddH20. Kept at 4 ºC.

Trypsin/EDTA

7.5 ml 4% (w/v) EDTA, 30 ml 0.25% (w/v) Trypsin & 37.5 ml autoclaved PBS Kept at 4 ºC.

5x Protein Loading Buffer (PLB)

1 g SDS powder, 771.25 mg DTT powder, 2.5 ml 250 mM Tris-Hcl pH6.8 solution, 500 mg Bromophenol Blue powder, 5 ml 50% (w/v) Glycerol & 2.5 ml ddH20 combined under a fume hood. 5x PLB was kept at RT.

5x DNA Loading Buffer (DLB)

0.2 % (w/v) bromophenol blue powder, 25 mM EDTA solution, 39% (v/v) glycerol solution & 0.2 % (w/v) xylene cyanol combined and dissolved in ddH20 under a fume hood. 5x DLB was kept at RT.

3% (w/v) Bovine Serum Albumin (BSA)

0.3 g BSA powder in 10 ml sterile PBS. Vortexed and filter sterilised. Always made fresh.

1:1 Methanol:Acetone (v/v) Solution

25 ml 100% methanol mixed with 25ml 100% acetone under a fume hood and kept in -20 ºC.

LB Agar

14 g LB Agar dissolved in 400 ml ddH2O and autoclaved at 121 ºC for 15 min. Kept a RT.

LB Broth

8 g LB Agar dissolved in 400 ml ddH2O and then autoclaved at 121 ºC for 15 min. Kept a RT.

Lysis Base Buffer

20 mM HEPES pH 7.4, 40 mM NaCl, 2 mM MgCl2 & 1 % Triton (v/v). Kept at 4 ºC or on ice at all times.

NP-40 Cell Lysis Buffer

50 mM Tris-HCL pH 8, 150 mM NaCl & 1 % (v/v) NP-40. Kept at RT.

1% (w/v) Methylene Blue

Methylene Blue hydrate powder dissolved 100x in ddH2O to 1 % (w/v). Kept at RT.

37% (w/v) Formaldehyde

1.85 g paraformaldehyde was dissolved in 10 µl 10M KOH and 3.5 ml ddH2O under a fume hood. Once dissolved, the volume was adjusted to 5 ml using ddH2O and used immediately.

ChIP Lysis Buffer

50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH8, 1 % triton X-100, 0.1 % Sodium Deoxycholate, 0.1 % SDS dissolved in ddH20. Kept in dark & at RT.

10% (w/v) Sodium deoxycholate

2 g Sodium deoxycholate diluted in 20 ml ddH20. Kept in dark & at RT.

RIPA buffer

50 mM Tris-HCl pH8, 150 mM NaCl, 2 mM EDTA pH8, 1 % NP-40, 0.5 % Sodium Deoxycholate & 0.1 % SDS were dissolved in ddH20. Kept in dark & at RT.

Low salt wash buffer

0.1 % SDS, 1 % triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH8 & 150 mM NaCl were dissolved in in ddH20. Kept at RT.

High salt wash buffer

0.1 % SDS, 1 % triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH8 & 300 mM NaCl were dissolved in in ddH20. Kept at RT.

Lithium chloride wash buffer

0.25 M LiCl, 1 % NP-40, 1 % Sodium Deoxycholate, 1 mM EDTA & 10mM Tris-HCl pH8 were dissolved in ddH20. Kept in dark & at RT.

Elution buffer

1 % SDS & 100 mM NaHCO3 dissolved in ddH20. Kept at RT.

GFP dilution/wash buffer

10mM Tris-HCl pH7.5, 150mM NaCl & 0.5mM EDTA dissolved in ddH20. Kept at RT.

## Standard Equipment

Beckman GP Centrifuge

BioMAT2 Class II Biological Safety Cabinet

BioRad ChemiDoc™ MP Imaging System

BioRad Trans-Blot® Turbo™ Transfer System

BioRad PowerPac™ HC

BioRad Gel Electrophoresis Tanks

BMG Labtech FLUOstar Omega

Clean Air Limited Fume Cupboard – 1220/E5a/FC1

Fisher Scientific UV Crosslinker FB-UVXL-1000

Galaxy R CO2 Incubator (Wolf Laboratories)

Gilson Pipetman: P2, P20, P200, P1000

Gilson GVLab Vortex

Grant Sub-Aqua 5 Waterbath

Jenway 3510 pH Meter

Jenway 1002 Stirred

Jenway Genova Spectrophotometer

Labnet Prism R Cooling Microcentrifuge

Mettler AE 163 Scales

New Brunswick Scientific U57085-85ºC Ultra Low Freezer

Nikon Eclipse E4000 Inverted Microscope Equipped with a Nikon DXM1200 Digital Camera

Nikon Eclipse TE300 Trinocular Inverted Fluorescence Microscope

Olympus CK40 Inverted Microscope equipped with an Olympus U-LS30-3 Digital Camera

Sanyo Falcon 6/300 Refrigerated Centrifuge

Stuart Mini See-Saw Rocker SSM4

Stuart Block Heater SBH130

Techne TC-3000X Thermocycler

Thermo Scientific ND-1000 Nanodrop Spectrophotometer

Thermo Scientific Heraeus Pico 17 Microcentrifuge

## Mammalian Tissue Culture

Adherent cells were used:

- MRC-5 – Human Foetal Lung Fibroblasts

- HEK-293 – Human Embryonic Kidney cells

- U-2 OS – Human Bone Osteosarcoma cells

## Media

MRC-5, HEK-293T and HEK-293T-based CRISPR cell lines were cultivated in Minimum Essential Medium (MEM, Sigma Aldrich). The media were supplemented with Fetal Calf Serum (10 % v/v, Sigma Aldrich), L-Glutamine (1 % v/v, Gibco) and Pen Strep (1 %, ThermoFisher Scientific). U-2 OS cell line was cultivated in McCoy’s 5a media (modified), which was supplemented with 10 % FCS. Cells were grown in 5 % CO2 incubators at 37 ºC. Media bottles were stored at 4 ºC and heated up to 37 ºC before use in a water bath.

## Cell Splitting

>75 % confluent cells were split once a week 1:10 or 1:5. Media was removed from the T75 flask and adherent cells were washed twice with 15 ml PBS. Then, 2 ml of Trypsin/EDTA was added to the cells and left for 5 min. 8 ml of suitable medium was added to wash off the cells. Solution was transferred to Universal tube and spun at 1000 rpm (revolution per minute) for 5 min. Pelleted cells were resuspended in fresh medium. 1 ml (1:10) or 2 ml (1:5) of cell suspension was aspirated and added to the flask and topped up with fresh medium to 15 ml. In case of T175 flasks, 30 ml of PBS was used per wash, 4 ml of Trypsin/EDTA and cells were kept in total volume of 30 ml. All cells in flasks were kept in 37 ºC.

## Cell Counting

10 μl of cell suspension was aspirated and the cells were counted with Improved Neabauer Haemocytometer under an inverted microscope.

## siRNA transfection

Filter tips were used only. Cells were seeded on day 1 and left in a 37 ºC incubator until 60 % confluent. Media was aspirated and 1000 μl of fresh media was added. Next, 200 μl of transfection mix was added per well. Concentration of siRNA was 50 nM in final volume of 1200 μl. Transfection mix was made using Serum Free Media (SFM). 100 μl of SFM was mixed with 3.0 μl of DharmaFECT1 (Dharmacon) in a fresh Eppendorf tube and left for 5 min. Meanwhile, 100 μl of SFM was mixed with desired siRNA in another fresh Eppendorf tube and left for 5 min. After the incubation time, both solutions were mixed and incubated for another 20 min at room temperature. 200 μl of a transfection mix was added to a single well drop by drop. The plate was left for 48 h in 37 ºC incubator. The volumes were scaled up or down depending on the size of a plate. Total 600 μl volume was used for 24-well plate, 1200 μl for 6-well plate, 9600 μl for 15 cm dishes.

## Plasmid transfection

Non-filter tips were used. Cells were seeded on day 1 and left in a 37 ºC incubator until 60 % confluent. Media was aspirated and 1000 μl of fresh media was added. 200 μl of transfection mix was added per well on Day 2. Transfection mix was made using Serum Free Media (SFM). 100 μl of SFM was mixed with 1000 ng of plasmid DNA in another fresh Eppendorf tube and left for 5 min. Meanwhile, 100 μl of SFM was mixed with 2.0 μl of 1 mg/ml pH 7 PEI (Polyethylenimine; Polysciences, 23966-1) in a fresh Eppendorf tube and left for 5 min. After the incubation time, both solutions were mixed and incubated for another 15 min at room temperature. 200 μl of a transfection mix was added to a single well drop by drop. The plate was left for 48 h in 37 ºC incubator. The volumes were scaled up or down depending on the size of a plate. DNA to PEI ratio was always 1 µg of DNA : 2 µl of PEI. Total 600 μl volume was used for 24-well plate, 1200 μl for 6-well plate, 20 ml for 15 cm dishes.

## Double transfection

First, siRNA transfection (2.3.4) was carried out. Cells were left for 4 h in 37 ºC incubator after transfection. Then, plasmid transfection (2.3.5) was carried out. 4 h post second transfection, media was aspirated and fresh media was added in order to remove toxic transfection reagents.

## siRNA sequences

siRNA name: Sequence:

USP11 siRNA #1 5’ GCG CAC AGC UGC AUG UCA U 3’

USP11 siRNA #2 5’ GAG AAG CAC UGG UAU AAG C 3’

USP11 siRNA #3 5’ GGA CCG UGA UGA UAU CUU C 3’

USP11 siRNA #4 5’ GAA GAA GCG UUA CUA UGA C 3’

SETX siRNA #1 5’ GCA CGU CAG UCA UGC GUA A 3’

SETX siRNA #2 5’ GCA AUA AGC UCA UCC UAG U 3’

SETX siRNA #3 5’ GCU CAA CUC UCC AAA UAG A 3’

SETX siRNA #4 5’ UAG CAC AGG UUG UUA AUC A 3’

Scrambled/control siRNA 5’ UAA UGU AUU GGA ACG GAU 3’

UCHL3 siRNA #1 5’ CAGCAUAGCUUGUCAAUAA 3’

UCHL3 siRNA #2 5’ GCAAUUCGUUGAUGUAUAU 3’

## Plasmids

Plasmid name: Source:

pIRES-FLAG-HA-USP11 Wade Harper

pIRES-FLAG-HA-USP11(C318S) this thesis

pIRES-10xHis Sherif El-Khamisy

pIRES-Ubiquitin-10xHis Sherif El-Khamisy

peGFP-C1 Clonetech

pCMV6-XL-SETX OriGene

peGFP-SETX(1-667aa) this thesis

peGFP-RNaseH1 Sherif El-Khamisy

## Molecular biology techniques

## Immunofluorescence

60 000 to 80 000 cells per well were seeded for immunofluorescence (IF) on glass coverslips in 24-well plates on day 1. If required, cells were treated with 50 μM α-amanitin (AMN) for 15 h, 50 μM MG132 for 2 h, or 25 μM CPT for 10 and 30min on day 2. DMSO was used as a negative control. When the incubation time was over, plates were put on ice, media was aspirated and cells were washed twice with 500 μl ice-cold PBS. Next, 200 μl / well of ice-cold methanol:acetone was added for 10 min and incubated in -20 ºC. Methanol:acetone was aspirated and cells were washed three times with 500 μl ice-cold PBS. 200 μl / well of filtered 3 % (w/v) BSA was added for 30 min and incubated at room temperature. Next, BSA was aspirated and 170 μl / well of primary antibody diluted in filtered 3 % (w/v) BSA was added:

- S9.6 antibody (α-R-loop): 1:500 dilution, Mouse, Kerafast (ENH001),

- α-eGFP antibody: 1:2000 dilution, Rabbit, Abcam (ab6556),

- α-Nucleolin antibody: 1:2000 dilution, Mouse, Abcam (ab136649),

- α-Nucleolin antibody: 1:2000 dilution, Rabbit, Abcam (ab22758),

- α-USP11 antibody: 1:2000 dilution, Rabbit, Bethyl Laboratories (A-301-613A),

Cells were incubated for 1h with primary antibodies. Antibodies were recycled and cells were washed three times with 500 μl PBS. Next, each well was incubated with 170 μl of secondary antibody (diluted in 3 % BSA with DAPI) for an hour in dark:

- Alexa Fluor® 555, α-mouse antibody, 1:500 dilution, ThermoFisher (a21428),

- Alexa Fluor® 594, α-mouse antibody, 1:500 dilution, ThermoFisher (a11005),

- Alexa Fluor® 488, α-rabbit antibody, 1:500 dilution, ThermoFisher (a11008),

- Alexa Fluor® 594, α-rabbit antibody, 1:500 dilution, ThermoFisher (a110012),

- DAPI (Sigma, D9542), 1ngml-1 final concentration.

Antibodies were removed, cells were washed three times with 500 μl PBS. After the final wash, cover slips were taken out of the wells, rinsed in sterile water, wiped and fixed on microscope slides using Thermo Scientific Shandon™ Immu-Mount™ (9990402). Slides were stored at 4 ºC in the dark.

## Generation of stable cell lines

CRISPR-Cas9 system was used to generate homogeneous stable cell lines following protocol described by Ran et al., 2013. Small guiding RNAs against exon 1 and exon 3 of USP11 were designed using crispr.mit.edu engine and cloned into Cas9 vector. HEK-293T cells were transfected with cocktail of two (exon 1 and exon 3) sgRNA(USP11)-eGFP-Cas9 plasmids. Transfection efficiency was examined under inverted fluorescent microscope by observing how many cells were expressing eGFP. Next, transfected cells were serially diluted on a 96-well plate. After few days, single-cell colonies were examined again for eGFP-expression. Positive cells were then propagated.

Small guiding RNAs: Sequence:

USP11 Exon 1 forward 5’ CAC CGA GAA CGG ACG GCG ATG GCG A 3’

USP11 Exon 1 reverse 5’ AAA CTC GCC ATC GCC GTC CGT TCT C 3’

USP11 Exon 3 forward 5’ CAC CGT GAG ATA AAC TGG CGC CTC A 3’

USP11 Exon 3 reverse 5’ AAA CTG AGG CGC CAG TTT ATC TCA C 3’

CRISPR-Cas9 plasmids: Source:

pSpCas9(BB)-2A-Puro\_PX459 Feng Zhang Lab

pSpCas9(BB)-2A-eGFP\_PX458 Feng Zhang Lab

To generate heterogeneous stable cell lines, protein of interest was transfected on a plasmid as described in section 2.3.5 into HEK-293T cells. The cells were cultivated in presence of appropriate selection marker. The concentration of the selection marker was increased steadily over weeks and untransfected cells were used as a control. Finally, a Western Blot was run to confirm expression of the protein of interest.

## Microscopy

63X objective of Leica FW4000 Fluorescent Microscope (Leica Microsystems) or Nikon confocal microscope system A1 (Nikon Instruments, Tokyo, Japan) was used to take images after immunofluorescence. Alexa Fluor 555 and 594 (400ms exposure time) were imaged using DsRed RMD filter (wavelength 558nm). Alexa Fluor 488 (400ms exposure time) was imaged using Alexa Fluor 488 filter (wavelength 488nm). DAPI (exposure time: 1-10ms) was imaged using DAPI filter (wavelength 359nm). The image acquisition was double-blinded. Lyca software was used to save images, which were analysed and cropped in ImageJ. S9.6 signal intensity was measured using ImageJ. The intensity signal was always corrected for the background noise and nucleolar masks were used.

## Clonogenics

200 000 MRC-5 cells per well were seeded on a day 1 in a 6-well plate. Cells were double transfected on a day 2. Cells were trypsinized and counted on day 3. Next, 4000 cells were seeded in 10 ml of media in 10 cm dish. Each condition was done in triplicate. Drugs were added on day 4. Olaparib was added for 24 h, CPT for 1.5 h, while FA for 3 h. Media was gently changed after drug incubation and plates were left for 7 days in 37 ºC incubator. Media was poured off, plates were left to air dry and 10 ml of 80 % ethanol was added to the plates for 15 min. Next, ethanol was removed and plates were left to completely air dry. 8 ml per plate of 1 % Methylene Blue was added for 1 h. Methylene Blue was recycled and plates were gently washed few times with ddH2O to remove the background staining. Plates were left to air dry and colonies were counted.

## Western Blotting

## Cell lysis

All work was carried out on ice. Cells were washed twice with 2ml ice-cold PBS. 50μl/well of lysis buffer (50 mM Tris-HCl (pH 8), 40 mM NaCl, 2 mM MgCl2, Triton (0.5 % v/v), 1x Protease Inhibitor, 20 mM NEM, 250 units of Basemuncher (Expedeon)) was added. Cells were scraped and lysates were transferred to Eppendrof tubes. These were vortexed every 4 min during a 20 min incubation. Lysates were centrifuged at 13200 rpm for 20 min at 4 ºC. The supernatants were aspirated and stored at -20 ºC.

## Bradford assay

1μl of lysate from 2.4.5.1 was mixed with 999 μl of Coomassie Protein Assay Regent in a spectrophotometer cuvette. Next, Bradford assay was followed on Jenway Genova Spectrophotometer. Absorbance values were normalized to the lowest one and SDS-PAGE samples were prepared.

## SDS-PAGE

Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis (SDS-PAGE) was used to separate proteins from whole cell extracts. 10 ml of 10 % resolving gel (3.3 ml 30 % Acrylamide, 4 ml ddH2O, 2.5 ml 1 M Tris (pH 8.8), 100 μl SDS, 100 μl APS and 4 μl UltraPure™ TEMED) was poured into a BioRad cassette, covered with 0.5 ml of isopropanol and left to polymerize at room temperature. Next, isopropanol was removed and stacking gel (3.736 ml ddH2O, 680 μl 30% Acrylamide, 500 μl 1 M Tris (pH 6.8), 40 μl 10 % SDS, 40 μl 10 % APS and 10 μl UltraPure™ TEMED) was poured on top of the resolving gel. Finally, BioRad 10-well combs were inserted into the cassette. The polymerized gel was inseted in the BioRad tank for SDS-PAGE and covered with 1xRunning Buffer. The comb was removed.

## Loading of samples

Equal amount of protein per sample was loaded into the SDS-PAGE gel using data from Bradford assay. Therefore, specific volume of whole-cell extract was mixed with 6 μl of 5x PLB and topped up with ddH2O to 30 μl. Samples were boiled at 95 ºC for 7 min and spun down briefly at high speed. Next, samples were loaded into the gel. 5 μl of Precision Plus Protein Dual Color Standard (BioRad) ladder was used and 30 μl of 1x PLB was added to empty wells. The gel was run at 120 V for the first 10-15 min and then at 200 V for next 60-90 min.

## Transfer

BioRad Trans-Blot® Turbo™ Transfer System was used to transfer proteins from SDS-PAGE gel to a BioRad nitrocellulose. Blotting papers and the nitrocellulose membrane were left 1xBioRad Trans-Blot® Turbo™ transfer buffer for few minutes to absorb the buffer. BioRad Trans-Blot® Turbo™ Transfer System was set to High Molecular Weight to ensure transfer of all proteins.

## Blocking

After carrying out the transfer, the nitrocellulose membrane was incubated in 5 % milk (w/v in 1x TBST) for 1 h on a rocker at room temperature.

## Antibodies

After blocking, the membrane was incubated with primary antibody overnight on a rocker at 4 ºC. Antibodies:

- α-eGFP (Rabbit, 1:2000 in 5 % milk, Abcam, ab6556)

- β-Actin (Mouse, 1:1000 in 5 % milk, Sigma, A5316)

- α-USP11 (Rabbit, 1:2000 in 5 % milk, Bethyl Laboratories, A-301-613A)

- α-KEAP1 (Rabbit, 1:1000 in 5 % milk, Abcam)

- α-SETX (Rabbit, 1:1000 in 5 % milk, Bethyl Laboratories, A-301-104A)

- α-TDP1 (Rabbit, 1:1000 in 5 % milk, Abcam, ab4166)

- α-SPRTN (Mouse, 1:1000 in 5 % milk, Atlas Antibodies, HPA025073)

- α-His (Mouse, 1:1000 in 5 % milk, Abcam, ab18184)

- α-GAPDH (Mouse, 1:5000 in 5 % milk, Calbiochem, CB1001)

- α-RPA 194 (Mouse, 1:1000 in 5 % milk, Santa Cruz, sc-48385)

- α-RNA POL II (Mouse, 1:1000 in 5 % milk, Santa Cruz, sc-55492)

- α-K48 ubiquitin linkage (Rabbit, 1:1000 in 5 % milk, Abcam, ab140601)

- α- β-Tubulin (Mouse, 1:1000 in 5 % milk, Abcam, ab7792)

Then, the membrane was washed three times with 10 ml of TBST for 10 minutes. A secondary antibody was added for one hour incubation on a rocker at room temperature (IgG (H+L)-HRP-Conjugate, 1:4000 in 5 % milk, BioRad, anti-mouse 170-6516, anti-rabbit 17-6522). The membrane was washed again three times for 10 min with 10 ml of TBST and then it was covered with 1 ml of Electrochemiluminescence (ECL) – Clarity™ Western ECL (BioRad, 1705060). The membrane was visualized using BioRad ChemiDoc™ MP Imaging System (BioRad, 1708280).

## DRIP

DRIP-qPCR was essentially performed as ChIP with minor modifications as described by Ray et al., 2013.

Cell harvesting

Two confluent 15 cm dishes of HEK2-293T cells (per condition) were washed once with ice-cold PBS. Then, 5 ml of ice-cold PBS were added to each dish, the cells were scraped and transferred to universal tubes. The cells were spun for 5 min at 4 °C at 1000 rpm. The supernatant was removed, the cells were resuspended in 500 µl of ice-cold PBS and transferred into fresh Eppendorf tubes. The cells were spun for 5 min at 4 °C at 1000 rpm again. The supernatant was removed, the cells were resuspended in 600 µl ChIP lysis buffer and left for 10 min on ice.

Sonication

The lysates were sonicated for 7 min (30 s on, 30 s off) in The Bioruptor® Pico (Diagenode, B01060010). Next, cell debris was pelleted by centrifugation for 10 min at 4 °C at 8000 g. The supernatant (containing sheared chromatin) was transferred to a fresh Eppendorf tube. 19 µl of each sonicated sample was snap frozen and set aside as an input (stored at -80 °C).

Immunoprecipitation

Remaining sheared chromatin was used for immunoprecipitation. Three Eppendorf tubes (A, B and C) were filled with 190 µl of sheared chromatin and 1710 µl of RIPA buffer each resulting in 1:10 dilution (per condition). Tube A would be used for mouse IgG isotype control (ThermoFisher, 10400C), while tubes B and C would be used for S9.6 antibody (Kerafast, ENH001). 28 µg of S9.6 antibody were added to tubes B & C each and 28 µg of mouse IgG control to the tube A. Next, the tubes were left on a rotator at 4 °C for 1 h. Meanwhile, 3x 30 µl Dynabeads™ protein G beads (ThermoFisher, 10004D) were washed twice in 400 µl RIPA buffer. Once washed, the beads were added to the tubes A,B,C and left overnight at 4 °C. The next day, tube C was removed from the rotator, supernatant was aspirated and 85 µl of ddH20, 10 µl of 10x RNase H reaction buffer and 25 units of RNase H were added (NEB, M0297L). The tube C was left for 90 min on a shaker at 37 °C. Next, all the tubes were placed in DynaMag™ magnet (ThermoFisher, 12320D) and the supernatants were discarded. The beads were washed once in a low salt wash buffer, once in a high salt wash buffer, and once in a Lithium Chloride wash buffer. Then, the tubes were spun for 1 min at 1300 rpm and the residual supernatant was removed. 150 µl elution buffer were added to each tube and the tubes were left on a shaker for 15 min at 30 °C. The eluates were transferred to fresh Eppendorf tubes. Meanwhile, 19 µl of the input chromatin were thawed on ice and diluted in the elution buffer to 150 µl. Next, 4.8 µl of 5 M NaCl, 2 µl RNase A (10 mg / ml) and 2 µl proteinase K (20 mg / ml) were added to the tubes and left shaking at 37 °C for an hour.

Phenol-chloroform extraction

1 volume of UltraPure™ Phenol:Chloroform: Isoamyl Alcohol (ThermoFisher, 15593031) solution was added to each tube, inverted gently and spun at 15000 rpm for 5 min at 4 °C. Top layer was extracted and transferred to a fresh tube. Next, 1 volume of Chloroform (Sigma, 288306) solution was added to each tube, inverted gently and spun at 15000 rpm for 5 min at 4 °C. Top layer was extracted and transferred to a fresh tube.

Ethanol precipitation

1/10 volume of 3 M Sodium Acetate (Sigma, S2889), 2.4 volumes 100 % ethanol and 1 μl of glycogen (ThermoFisher, 10814010) were added. Solution was inverted gently to precipitate DNA left to incubate at -80 °C for 45 min. Next, DNA was pelleted at 15000 rpm for 30 min at 4 °C. The supernatant was aspirated and DNA was washed 2 times with 70 % EtOH (5 min, 15000 rpm, 4 °C each wash). Then, the DNA was left to air dry and resuspended in 50 μl of ddH20. DNA samples were examined by qPCR.

## DRIP (adopted from Sanz and Chedin, 2019)

DRIP-qPCR was conducted following a recently published protocol (Sanz and Chedin, 2019).

Two confluent 15 cm dishes of HEK2-293T cells (per condition) were washed once with ice-cold PBS. Then, 5 ml of ice-cold PBS were added to each dish, the cells were scraped and transferred to universal tubes. The cells were spun for 5min at 4 °C at 1000 rpm. The supernatant was removed, the cells were resuspended in 500 µl of ice-cold PBS and transferred into fresh Eppendorf tubes. The cells were spun for 5min at 4 °C at 1000 rpm again. The supernatant was removed and the cells were resuspended in 1.6 TE pH8.0. SDS and proteinase K were added to 0.625 % SDS and 62.5 µg / ml concentrations respectively. Next, cells were left to lyse overnight in 37 °C.

The following day, DNA was extracted following a standard phenol-chloroform extraction and ethanol precipitation as described above. However, after ethanol precipitation, DNA was not spun down but spooled using a cut 1 ml tip. Once spooled, DNA was washed with 70 % EtOH but not spun down. Instead, 70 % EtOH was gently aspirated. Next, DNA was air-dried and resuspended in 125 µl of TE buffer. 75 µl of the DNA suspension in TE were taken for subsequent overnight digest with 30 units of *BsrGI, EcoRI, HindIII, SspI* and *XbaI* restriction enzymes.

The following day, 8 µg of digested DNA was incubated with *E.coli* RNase H enzyme for 4-6 h. Next, 8 µg of digested DNA (from overnight digest) and 8 µg of RNase-H-treated DNA were incubated overnight with 20 µl of S9.6 antibody in the DRIP binding buffer (10 mM sodium phosphate pH 7.0, 140 mM NaCl, 0.05 % (v/v) Triton X-100, TE).

The following day, 100 µl of protein G beads (ThermoFisher, 10003D) were washed three times in 500 µl of DRIP binding buffer by vortexing and added to the S9.6-containing DNA digests from the previous step. The tubes were left to incubate on an end-to-end rotor for 2 h in 4 °C. Next, the beads were washed three times by vortexing in the DRIP binding buffer. The wash buffer was aspirated and R-loops were eluted from the beads in 50 mM Tris pH 8.0, 10 mM EDTA pH 8.0, 0.5 % (v/v,) 0.456 mg / ml proteinase K. The elution was done on an end-to-end rotor at 55 °C for 45 min. Then, DNA was extracted from the eluate by standard phenol-chloroform extraction and ethanol precipitation. Once air-dried, DNA was resuspended in 50 µl of ddH20. DNA samples were examined by qPCR.

## Cross-linking Chromatin Immunoprecipitation

It was performed as described by Ray et al., 2013.

Formaldehyde cross-linking

Freshly made 37 % (w/v) formaldehyde was added drop-wise to two confluent 15 cm dishes containing HEK-293T to a final concentration of 1 %. The cells were left on a rotor at RT for 10 min. Next, glycine was added drop-wise to the cells to a final concentration of 125 mM. The cells were left on a rotor at RT for 5 min.

Cell harvesting

Cross-linked cells were washed once with ice-cold PBS. Then, 5 ml of ice-cold PBS were added to each dish, the cells were scraped and transferred to a universal tube. The cells were spun for 5 min at 4 °C at 1000 rpm. The supernatant was removed, the cells were resuspended in 500 µl of ice-cold PBS and transferred into a fresh Eppendorf tube. The cells were spun for 5 min at 4 °C at 1000 rpm again. The supernatant was removed, the cells were resuspended in 600 µl ChIP lysis buffer and left for 10 min on ice.

Sonication

The lysates were sonicated for 7 min (30 s on, 30 s off) in The Bioruptor® Pico (Diagenode, B01060010). Next, cell debris was pelleted by centrifugation for 10 min at 4 °C at 8000 g. The supernatant (containing sheared chromatin) was transferred to a fresh Eppendorf tube. 19 µl of each sonicated sample was snap frozen and set aside as an input (stored at -80 °C).

Immunoprecipitation

Remaining sheared chromatin was used for immunoprecipitation. Two Eppendorf tubes (A and B) were filled with 190 µl of sheared chromatin and 1710 µl of RIPA buffer (supplemented with protease inhibitor) each resulting in 1:10 dilution (per condition). Tube A was used for rabbit IgG isotype control (ThermoFisher, 10500C), while tube B was used for USP11 antibody. 2 µg of USP11 antibody were added to the tube B. 2 µg of rabbit IgG control were added to the tube A. Next, the tubes were left on a rotator at 4 °C for 1h. Meanwhile, 2x 30 µl Dynabeads™ protein A beads (ThermoFisher, 10002D) were washed twice in 400 µl RIPA buffer. Once washed, the beads were added to the tubes A and B, and left overnight at 4 °C. The next day, all the tubes were placed in DynaMag™ magnet (ThermoFisher, 12320D) and the supernatants were discarded. The beads were washed once in a low salt wash buffer, once in a high salt wash buffer, and once in a Lithium Chloride wash buffer. Then, the tubes were spun for 1 min at 1300 rpm and the residual supernatant was removed. 150 µl elution buffer were added to each tube and the tubes were left on a shaker for 15 min at 30 °C. The eluates were transferred to fresh Eppendorf tubes. Meanwhile, 19 µl of the input chromatin were thawed on ice and diluted in the elution buffer to 150 µl. Next, 2 µl of RNase A (10 mg / ml) were added to each tube and left shaking at 37 °C for 30 min. Then, 6 µl of 5 M NaCl and 2 µl proteinase K (20 mg / ml) were added to the tubes and left shaking at 65 °C for two hours.

Phenol-chloroform extraction

1 volume of UltraPure™ Phenol:Chloroform: Isoamyl Alcohol (ThermoFisher, 15593031) solution was added to each tube, inverted gently and spun at 15000 rpm for 5 min at 4 °C. Top layer was extracted and transferred to a fresh tube. Next, 1 volume of Chloroform (Sigma, 288306) solution was added to each tube, inverted gently and spun at 15000 rpm for 5 min at 4 °C. Top layer was extracted and transferred to a fresh tube.

Ethanol precipitation

1/10 volume of 3M Sodium Acetate (Sigma, S2889), 2.4 volumes 100 % ethanol and 1 μl of glycogen (ThermoFisher, 10814010) were added. Solution was inverted gently to precipitate DNA left to incubate at -80 °C for 45 min. Next, DNA was pelleted at 15000 rpm for 30 min at 4 °C. The supernatant was aspirated and DNA was washed 2 times with 70 % EtOH (5 min, 15000 rpm, 4 °C each wash). Then, the DNA was left to air dry and resuspended in 50 μl of ddH20. DNA samples were examined by qPCR.

## Pull-down assay under denaturing conditions

Cell lysis

Transfected HEK-293T cells were grown in a 15 cm dish. 48 h post transfection, 25 µM MG132 was added to the cells for 2 h. Next, the cells were washed twice in 10 ml of ice-cold PBS. PBS was removed and cells were scraped in residual PBS and pelleted by centrifugation at 1000 rpm for 5 min. PBS was removed and cells were re-suspended in 500 µl of lysis base buffer (50 mM Tris-HCl (pH 8), 40 mM NaCl, 2 mM MgCl2, Triton (0.5 % v/v), 1x Protease Inhibitor, 20 mM NEM, and 250 units of Basemuncher (BM0025m, Expedeon)). Then, the cells were incubated on ice for 30 min with periodic vortexing every 5 min. Next, cells were spun down at 15000 rpm for 20 min at 4 °C. The supernatant was collected (whole cell lysate). Appropriate volume of whole cell lysate, equal to 25 µg of protein, was kept aside for Input sample. Next, 1 volume of equilibration buffer (6 M guanidine hydrochloride, 0.05 % (v/v) Tween-20, 40 mM imidazole) was added to the lysate.

Bead preparation

30 µl of GFP-Trap®\_MA beads (Chromotek, Planegg-Martinsried, Germany) or His-Pur™ Ni-NTA magnetic beads (ThermoFisher) were washed twice in a wash buffer (8 M Urea, 0.05 % (v/v) Tween-20, 50 mM imidazole, PBS pH 8.0). During each was, the beads were rotating for 5min at room temperature.

Pull down

Diluted cell lysate was added to the washed beads, and left to incubate for 1 h on an end-to-end rotor at room temperature. Next, the beads were washed three times in 1 ml of the wash buffer (8 M Urea, 0.05 % (v/v) Tween-20, 50 mM imidazole, PBS pH 8.0). During each was, the beads were rotating for 5 min at room temperature. Then, the beads and the input samples were boiled in 40 µl of 1x SDS protein loading buffer for 10 min with periodic vortexing. Next, the beads and the input samples were spun down by centrifugation at full speed for 1 min and the supernatants were loaded on a 4-15 % precast gel (BioRad, 4561093).

## GFP pull-down assay under native conditions

Cell lysis

Transfected HEK-293T cells were grown in a 15 cm dish. 48 h post transfection, 25 µM MG132 was added to the cells for 2 h. Next, the cells were washed twice in 10 ml of ice-cold PBS. PBS was removed and cells were scraped in residual PBS and pelleted by centrifugation at 1000 rpm for 5 min. PBS was removed and cells were re-suspended in 200 µl of lysis buffer (20 mM HEPES (pH 7.4), 40 mM (NaCl), 2 mM MgCl2, Triton (1 % v / v), 1x Protease Inhibitor, 40 mM NEM, and 250 units of Basemuncher). The cells were incubated on ice for 30 min with periodic vortexing every 5 min. Next, cells were spun down at 20000 g for 10 min at 4 °C. The supernatant was collected (whole cell lysate). Appropriate volume of whole cell lysate, equal to 25 µg of protein, was kept aside for Input sample. The remaining supernatant was diluted to 1 ml in GFP dilution/wash buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA).

Bead preparation

eGFP-Trap®\_MA beads (Chromotek, gtma-10) were used. 25 µl of the beads were washed three times in 500 µl of the GFP dilution/wash buffer. After each wash, the beads were separated from the supernatant on a magnetic stand.

Pull down

Diluted whole cell lysate was added to the washed beads and left to rotate for 2 h at 4 °C. After incubation, the lysate was removed and beads were washed three times with the GFP dilution/wash buffer. Next, the beads were spun down for 1 min at 1000 rpm and the residual buffer was aspirated. 40 µl of 1xSDS loading buffer was added to the beads. 5x SDS loading buffer and ddH20 were added to the input sample to 40 µl final volume. The beads and input samples were boiled at 95 °C for 10 min with periodic vortexing. Next, the beads and the input samples were spun down by centrifugation at full speed for 1 min and the supernatants were loaded on a 4-15% precast gel (BioRad, 4561093).

## Slot blot

Slot blot was performed as described by Stork et al (2016). Genomic DNA was extracted using High Pure PCR Template Preparation Kit (Roche) and quantified by Nanodrop. 1 µg of DNA from each sample was spotted in duplicates on nitrocellulose membrane using a vacuum pump and a slot blot apparatus. The membrane was air dried and cut into two pieces – one for S9.6 (R-loop) antibody and the other one for ssDNA antibody (loading control). For ssDNA, the blot was denatured for 10 min in 1.5 M NaCl, 0.5 M NaOH solution and then neutralised for 10 min in 0.5 M Tris-HCl pH 7.0, 1 M NaOH solution. Next, the blots were UV-crosslinked (120000 µJ / cm2) and subsequently blocked in 5 % milk for 1 h. The blots were incubated with S9.6 antibody (1:500, Kerafast) and ssDNA antibody (1:10 000, #MAB3868, EMD Millipore) over night on a shaker. The next day, they were washed three times with 1x TBST, incubated with secondary anti-mouse antibody (1:4000 in 5 % milk, BioRad) for 1 h at room temperature on a shaker, washed three times again with 1x TBST and visualised using BioRad ChemiDoc™ MP Imaging System.

## PCR

PCR reactions for Site-directed Mutagenesis and for amplification of specific sequences were done using the conditions described below. PCR kit was supplied by Merck (71805-3).

|  |  |
| --- | --- |
| Reagent | Volume per reaction (µl) |
| 10x KOD Hot Start Polymerase Buffer | 5 |
| 2 mM dNTPs | 5 |
| 25 mM MgSO4 | 2 |
| DMSO | 1 |
| Forward Primer (10 µM) | 1.5 |
| Reverse Primer (10 µM) | 1.5 |
| Template DNA (20 ng / µl) | 2 |
| KOD Hot Start Polymerase (1 unit / µl) | 1 |
| PCR Grade Water | 31 |

Reaction times:

1. Initial denaturation step: 5 min at 94 °C,
2. 34 cycles of:

- Denaturation: 1 min at 94 °C,

- Annealing: 1min at temperature 5 ºC lower than the lowest primer melting temperature,

-Elongation: 1 min / kb at 68 °C,

3) Final extension: 10 min at 68 °C,

4) Infinite hold at 4 °C.

In case of SDM, DPN1 restriction enzyme (NEB, R0176S) was added to the complete PCR reaction in order to digest template DNA. This was done according to the manufacturer’s protocol.

## PCR purification kit

Complete PCR reaction mix was purified using QIAquick PCR purification kit (Qiagen, 28104). Subsequently, purified DNA was used for bacterial transformation (plasmids) or sent for sequencing to GATC.

## RNA extraction

RNeasy Mini Kit (Qiagen, 74104) was used to extract and purify total RNA according to the manufacturer’s instruction manual. Immediately after the extraction, the cDNA synthesis protocol was followed.

## cDNA synthesis

cDNA synthesis was carried out using Applied Biosystems™ High Capacity cDNA Reverse Transcription Kit (ThermoFisher, 10400745). The protocol supplied by the manufacturer was followed.

## RT-qPCR and qPCR

Sample and standard preparation from cDNA samples (RT-qPCR)

cDNA samples were divided to make standards and DNA samples. To prepare standards for qPCR, samples were mixed 1:1 to make S1 (Standard 1, 100 %). Next, S2 (10 %), S3 (1 %) and S4 (0.1 %) were prepared using serial dilution method and ddH2O. DNA samples were made by diluting cDNA sample 1:10 in ddH2O. 5 μl of DNA sample and 5 μl of a standard were enough for a single run with one primer mix.

Sample and standard preparation from DRIP of ChIP samples

Input DNA samples were serially diluted to make S1 (Standard 1, 100 %), S2 (10 %), S3 (1 %) and S4 (0.1 %) in ddH2O. Immunoprecipitation samples were not mixed together and were used as DNA samples.

Primer mix preparation

2.8 μl of 5 μM primer mix were mixed with 10 μl of 2xSensimix™ (Bioline, QT615-05) and 2.8 μl of ddH2O. This solution was enough for one DNA sample or standard sample. Therefore, primer mix solutions were made up to match the number of standards and DNA samples. Then, primer mix solutions were mixed with 5 μl of standard or DNA solutions using filter tips. Next, qPCR was run on Rotor-Gene 6000 (Corbett Life Science). Initial step was carried out at 95 °C for 10 min. Next, 45 cycles were repeated of 95 °C for 15 s, 55 °C for 15 s and 72 °C for 30 s. Data was scored using Rotor-Gene 6000 series Software 1.7 (Corbett Life Science).

Quantification of RT-qPCR data

Rotor-Gene 6000 series Software 1.7 (Corbett Life Science) was used to determine the quantification cycle (Cq) value. Cq is a number at which the fluorescence signal is significantly above the background noise. Next, Cq values were converted to copy numbers using standard curve method (four standard dilutions in duplicates). Gene expression was normalised to Actin β mRNA levels. Values in the control cells were treated as 100 %.

Quantification of DRIP-qPCR data

R-loop accumulation in each gene loci was determined as percentage of input chromatin and then normalised to controls.

## Primers

Detailed primer data can be found in the appendix-1.

cDNA-qPCR primers:

USP11 forward 5’ TGG AAG GCG AGG ATT ATG TGC 3’

USP11 reverse 5’ ATG ACC TTG CGT TCA ATG GGT 3’

SETX forward 5’ CTT CAT CCT CGG ACA TTT GAG 3’

SETX reverse 5’ TTA ATA ATG GCA CCA CGC TTC 3’

KEAP1 reverse 5’ GGC TGA TGA GGG TCA CCA GTT 3’

KEAP1 forward 5’ TGG CCA AGC AAG AGG AGT TC 3’

Actin β forward 5’ GCT CAG GGC TTC TTG TCC TT 3’

Actin β reverse 5’ TCG ATG GGG TAC TTC AGG GT 3’

DRIP-qPCR primers

Actin 5’pause forward 5’ TTA CCC AGA GTG CAG GTG TG 3’

Actin 5’pause reverse 5’ CCC CAA TAA GCA GGA ACA GA 3’

ING3 forward 5’ TTT TTC TTC TCT AAC TAC CCT CCC C 3’

ING3 reverse 5’ GTG CCC TAA TCT GAA TGA CTA CA 3’

R#7 forward 5’ GAC ACT TCG AAC GCA CTT G 3’

R#7 reverse 5’ CTC AGA CAG GCG TAG CCC CG 3’

28S forward 5’ CAG GGG AAT CCG ACT GTT TA 3’

28S reverse 5’ AT GAC GAG GCA TTT GGC TAC 3’

EGR1 forward 5’ CATAGGGAAGCCCCCTCTTTC 3’

EGR1 reverse 5’ CTTGTGGTGAGGGGTCACTT 3’

SNRPN-neg 5’ forward 5’ GCCAAATGAGTGAGGATGGT 3’

SNRPN-neg reverse 5’ TCCTCTCTGCCTGACTCCAT 3’

MYADM-neg forward 5’ TGCATCTACATCCGCAAAAG 3’

MYADM-neg reverse 5’ AGAGTGGACGCTGCAGAAAT 3’

The USP11C318S mutation was made using following primers based on Wiltshire et al., 2010.

USP11 ASM forward 5’ CAA TCT GGG CAA CAC G**A**G CTT CAT GAA CTC GGC 3’

USP11 ASM reverse 5’ GCC GAG TTC ATG AAG C**T**C GTG TTG CCC AGA TTG 3’

USP11 ASM sequencing forward 5’CAT GTC GGA AGA GGA TGA GGA CTT C 3’

Primers used to generate peGFP-SETX(1-667aa):

SETX (1-667) forward 5’ ACA AGT CCG GAC TCA GAT CTA TGA GCA CAT GTT GTT GG 3’

SETX (1-667) reverse 5’CAC TCA CTG GTA CTG GCC CTT TAG TCA CCT TCT ATA GTG TTA TC 3’

peGFP-C1 forward 5’ AGG GCC AGT ACC AGT GAG 3’

peGFP-C1 reverse 5’ AGA TCT GAG TCC GGA CTT G 3’

Sequencing primers:

USP11 exon 1 foward 5’ CAT GCC ACC TGT GCC CTA AGT CTA AGT CTA C 3’

USP11 exon 1 reverse 5’ CTA TCT GGC GCC ACT GGC TGT C 3’

USP11 exon 3 foward 5’ CCA AAG TCT GAC CAG TGC CTT CTC AGT CAG C 3’

USP11 exon 3 reverse 5’ GGT ACA CAT ATG TCT GCT CTC TTC CC 3’

peeGFP-SETX(1-667) sequencing forward 5’CTG CTG CCC GAC AAC CAC TA 3’

## Bacterial transformation

50 μl of competent DH5α *E.coli* were thawed. 1 μl of DNA prep was added to cells and they were left on ice for 30 min before being heat-shocked for 45 s at 42 °C. 200 μl of LB broth was added and the cells were incubated for 1 h at 37 °C on a shaker. Cells were pelleted at 4000 rpm for 2 min, excess of supernatant was aspirated and cells were resuspended in the residual LB broth. Next, cells were plated on an LB agar plate with antibiotic selection using strict aseptic technique. The plate was left overnight at 37 °C.

## Liquid bacterial culture

LB broth containing appropriate antibiotic selection was prepared. 5 ml was prepared for Mini-prep, 100 ml for Midi-prep. A single colony was added to the broth and left overnight to grow on a shaker at 37 °C.

## Glycerol stock

750 µl of liquid bacterial culture was mixed with 250 µl 70 % (v / v) glycerol and frozen at -80 °C.

## Mini-prep

QIAprep® Spin-Mini-prep kit (Qiagen) was to used extract plasmid DNA and supplied protocol was followed. Plasmid DNA was eluted in 50 μl of elution buffer.

## Midi-prep

QIAGEN® Plasmid *Plus* Midi-prep kit (Qiagen) was to used extract plasmid DNA and supplied protocol was followed. Plasmid DNA was eluted in 200 μl of elution buffer.

## DNA quantification

Thermo Scientific ND-1000 Nanodrop Spectrophotometer was used to quantify plasmid DNA concentration. Instructions on the screen were followed.

## DNA agarose gel electrophoresis

100 ml of 1 % (w / v) TBE buffer was prepared and mixed with 2 μg / ml ethidium bromide. The solution was poured into a gel tray with a comb. Once set, the comb was removed and DNA samples in 1x DNA loading buffer were loaded alongside 5 μl of 2-Log DNA ladder (NEB, N3200L). The samples were run for 45 min at 110 V in 1x TBE buffer. Imgaes of DNA bands were acquired using ChemiDoc MP imaging system (Bio-Rad, 1708280).

## DNA fragment purification

QIAquick® Gel Extraction Kit (Qiagen, 28704) was used to purify DNA fragments from agarose gels according to the manufacturer’s instruction manual.

## DNA sequencing

GATC Biotech (Cologne, Germany) LightRun Sanger sequencing service was used to validate generated changes in vectors and genomic DNA.

## Gibson assembly

Gibson assembly method was used to generate eGFP-SETX(1-667aa) plasmid. peGFP-C1 plasmid was used as a backbone. pCMV6-XL-SETX plasmid contained SETX(1-667aa) insert. Overlapping primers were designed (section 2.1.2 Primers) using online NEBuilder® Assembly Tool. The insert and the backbone were amplified by PCR described in the section 2.4.11 PCR. Next, the amplified fragments were purified by gel electrophoresis method (described in 2.1.9-10). Then, Gibson Assembly Cloning Kit (NEB, E5510) was used to subclone the insert into the backbone according to the manufacturer’s instructions. Next, Gibson Assembly Chemical Transformation Protocol was followed using 2μl of the assembly reaction and NEB5α chemically-competent cells. Single colonies were inoculated into antibiotic-containing LB media and left on a shaker at 37°C over-night. The next day, mini-preps were carried out (section 2.1.6 Mini-Prep) and plasmids were sent off for sequencing (section 2.1.11 DNA sequencing).

## Statistical analysis

Statistical differences were calculated using Student’s t-test for pair-wise comparisons. ‘Two-tailed distribution’ and ‘two-sample unequal variance’ were always chosen. All data are presented as the mean values ± standard deviations (SD) of at least three biological repeats. A p-value of <0.05 was considered to be statistically significant. A p-value of <0.05 was denoted as \*, <0.01 as \*\*, <0.001 as \*\*\*, <.0001. ‘ns’ denoted p values >0.05.

## Screening procedures

## Master Plates

An ON-TARGET plus siRNA library for human DUBs (G-104705-05, GE Life Sciences) was employed in a 96-well format. The library contained 99 siRNA pools. Each pool contained 4 siRNAs at 5 μM concentration. Hamilton Star Liquid Handling Robot was used to produce 96-well daughter plates, out of the master plates, which contained 1.5 μM pools. The daughter plates were subsequently used for reverse siRNA transfection, immunofluorescence and high-throughput microscopy.

## DUB Screen reverse siRNA transfection

Only filter tips were used. 50 μl / well of transfection mix (0.3 μl DharmaFECT1 (Dharmacon), 49.7 μl Serum Free Media) was added to daughter 96-well plates and left for 25 min at room temperature. Next, 4750 MRC-5 cells in 100 μl MEM suspension were added to each well and incubated for 48 h in a 37 ºC incubator.

## Immunofluorescence for the DUB Screen

4750 MRC-5 cells / well were reverse-transfected in a 96-well plate on day 1. On day 3, 25 μM CPT or 25 μM DMSO was administered to the cells for 10 or 30 min. When the incubation time was over, plates were put on ice, media was aspirated and cells were rinsed twice with 100 μl ice-cold PBS. Next, 40 μl / well of ice-cold methanol:acetone solution was added and incubated for 10 min in -20 ºC. Methanol:acetone was aspirated and cells were rinsed three times with 100 μl ice-cold PBS. Next, 40 μl / well of 3 % (w / v) BSA was added and incubated for 30 min at room temperature. BSA was aspirated and 40 μl / well of a primary antibody diluted in 3 % (w / v) BSA was added:

- S9.6 antibody (α-R-loop): 1:500 dilution, Mouse, Kerafast (ENH001).

Cells were incubated for 1 h with the S9.6 and then the antibody was recycled. Cells were rinsed three times with 100 μl PBS and 40 μl/well of secondary antibody (diluted in 3 % (w/v) BSA with DAPI) was added for an hour in dark:

- Alexa Fluor® 555, α-mouse antibody, 1:500 dilution, ThermoFisher (a21428),

- DAPI (Sigma, D9542), 1 ng ml-1 final concentration.

The secondary antibody was removed and cells were rinsed three times with 100 μl PBS. After the final rinse, 100 μl / well of PBS was added and the daughter 96-well plate was stored in dark in 4 °C.

## High-throughput microscopy

40X objective of ImageXpress® Micro XLS Widefield High-Content Analysis System (Molecular Devices) was used to acquire images of cells from 96-well plates. Pictures from fixed 36 sites / well were obtained using DAPI filter (exposure time 40 ms) and CY3 filter (for Alexa Fluor® 555; exposure time 1000 ms).

## High-throughput analysis of the DUB siRNA screen

Custom Module of MetaXpress® software was used to analyse images obtained via high-throuput microscopy. The software counted average S9.6 foci / cell scores for each biological condition. Counted S9.6 foci were nuclear, 2.5-5 μm in diameter and at least 40 shades of grey brighter than the nuclear background. The foci that were found on the nuclear/cytoplasmic border or on the not-fully-visible nuclei were excluded from the count

# Chapter 3: DUB siRNA screen to identify new R-loop regulators



## Introduction

When this study began, RNA:DNA hybrids (R-loops) were known to be transcription by-products that had negative and positive cellular roles. Chronic R-loops were linked to genome instability (Sollier et al., 2014), while naturally occurring R-loops regulated transcription (Sun et al., 2013). R-loop formation was shown to be regulated by multiple RNA-binding proteins (Li & Manley, 2005). R-loop resolution was conducted by nucleases RNase H1/2; and helicases senataxin and aquarius (Skourti-Stathaki et al., 2011; Sollier et al., 2014). Having this information, we hypothesised that tight R-loop homeostasis required another layer of regulation at transcriptional and post-translational levels to fine-tune these R-loop-regulating proteins; and hence R-loop homeostasis.

It has been discussed by Oberle and Blattner (2010) that various post-translational modifications are often employed in order to regulate DNA damage response. Following this logic, post-translational modifications could lead to down-regulation of proteins that resolve or prevent R-loop formation, named here as R-loop-associated proteins, and result in elevated genome instability and cell death. For instance, this could be attained by promoting degradation of R-loop-associated proteins via ubiquitination.

Ubiquitination is a process of conjugating a small ubiquitin peptide to target proteins, which may trigger an array of outcomes including protein degradation, modulation of activity, or conformational changes (reviewed by Al-Hakim et al., 2010). Deubiquitination is the opposite process, where ubiquitin peptides are deconjugated from target proteins by enzymes called deubiquitinases (DUBs). Inhibition of a DUB could result in its target proteins being highly ubiquitinated and potentially degraded. If a DUB X controlled a protein Y that resolves R-loops, then depletion of X could trigger degradation of Y, and hence elevate R-loop levels.

Notably, DUBs were suggested to be relatively easy protein targets for small molecule inhibitors, which made them promising candidates for novel cancer therapies (Al-Hakim et al., 2010; Mermerian et al., 2007; Ratia et al., 2008, D’Arcy et al., 2014). Therefore, it was decided to focus on DUBs and their potential role in R-loop homeostasis.

## Aims and objectives

The aim of this chapter was to carry out a DUB siRNA screen to find a novel R-loop regulator. The objectives were to:

* Design a system where R-loops could be visualised
* Test the specificity of the system
* Using the system, design a DUB siRNA screen
* Carry out the DUB siRNA screen
* Carry out secondary screens to test candidates from the DUB siRNA screen.

A part of research included in this chapter was performed during my pre-PhD master’s project in the El-Khamisy lab. To see which experiments where conducted during my master’s project, please see the appendix-2.

## Methodology

S9.6 is a mouse monoclonal antibody that was raised against an *in vitro* synthesised ΦX174 RNA-DNA antigen (Boguslawski et al, 1986). The antibody has a strong preference for R-loops (dissociation constant: 0.6nM) and 4.5 times weaker preference for AU-rich RNA:RNA duplexes (dsRNA) (dissociation constant 2.7nM) (Phillips et al., 2013). Due to this limitation, any S9.6 – based experiments need to be validated to prove that the antibody binds to R-loops rather than to dsRNA.

S9.6 immunoprecipitation experiments (DRIP) followed by qPCR require RNase A and RNase H control treatments. RNase A digests RNA contaminants, which could interfere with the qPCR reaction. *E.coli* RNase H (ec-RNase H) or human RNase H1 (hs-RNase H1) specifically cleaves R-loops, as opposed to dsRNA or dsDNA (Nowotny et al., 2007; Lima et al., 1997), serving as a signal-validation control.

S9.6 immunofluorescence (IF) experiments also require validation, which can be done through ablating the S9.6 signal by ec-RNase H treatment or transcription inhibition. Alternatively, the signal can be induced using a topoisomerase 1 (TOP1)-poison camptothecin (CPT). CPT stabilizes TOP1-cleavage complexes on DNA, which in turn stalls RNA Polymerases. The stalling events promote R-loop formation, as the nascent mRNAs have more time to invade the DNA duplex and form R-loops.

*In vitro* RNase III treatment could serve as another signal-validation control for S9.6 antibody. RNase III is an endoribonuclease that specifically cleaves dsRNAs. It was recently employed in an S9.6 IF protocol to validate the S9.6 signal (Silva et al., 2018).

DRIP-qPCR is a relatively expensive and time-consuming experiment, but it is quantitative and produces data at a resolution dependent on the shearing efficiency. It enables detection of R-loops at ~300bp resolution.

S9.6 IF is a relatively cheap and fast method, but it is semi-quantitative and produces data at a low resolution. It enables the comparison of R-loop signal at a cellular level.

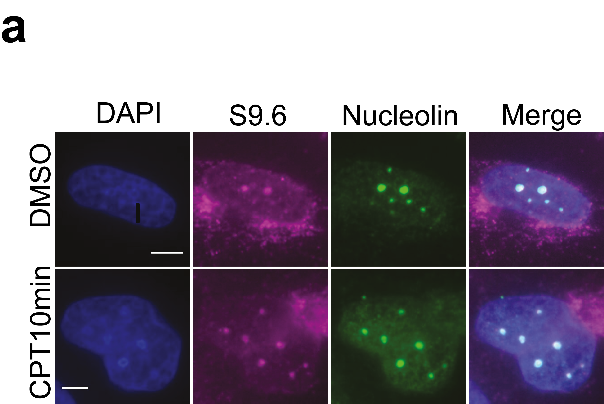
## Results

## Pre-screen pilot experiments

DRIP and S9.6 IF were the two most described methods to visualise R-loops at the beginning of this study. In order to carry out a DUB siRNA screen, a method to visualise and quantify R-loops in a high-throughput manner was required. DRIP cannot be scaled up for high-throughput screening due to the sheer amount of DNA needed for S9.6 immunoprecipitation (please see methods). However, the S9.6 IF method can be scaled up, as only 100 stained cells per repeat are required for quantification purposes. As a result, it was decided to optimise an S9.6 IF protocol for a DUB siRNA screen in a 96-well format.

It was decided to focus exclusively on S9.6 signal visible within the DAPI staining (nuclei staining). According to the state of the knowledge at the time, this approach excluded mitochondrial R-loops or cytoplasmic dsRNA, but focused on R-loops, and potentially dsRNA, found within nuclei (Skourti-Stathaki et al., 2014).

R-loops were previously described to localize to nucleoli (Marinello et al., 2013; Shen et al., 2017) and accumulate upon short exposure to topoisomerase 1-poison camptothecin (CPT) (Marinello et al., 2013). Therefore, MRC-5 cells were mock- or CPT-treated and immediately harvested for S9.6/nucleolin immunofluorescence (R-loop/nucleoli) (Fig. 3.1a). Confocal images revealed that R-loops observed in this experimental condition co-localised with nucleoli, implying their nucleolar origin.

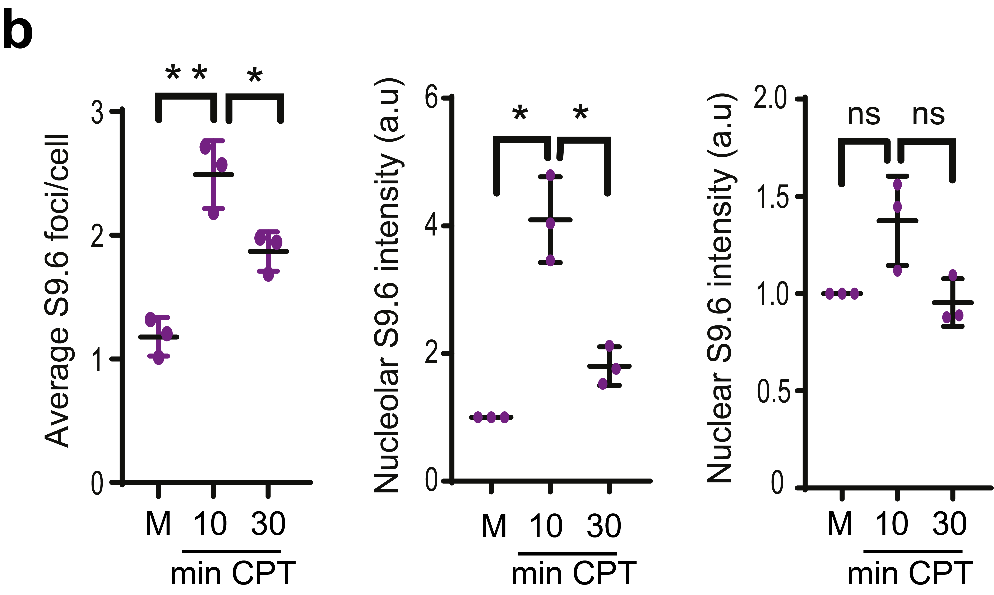
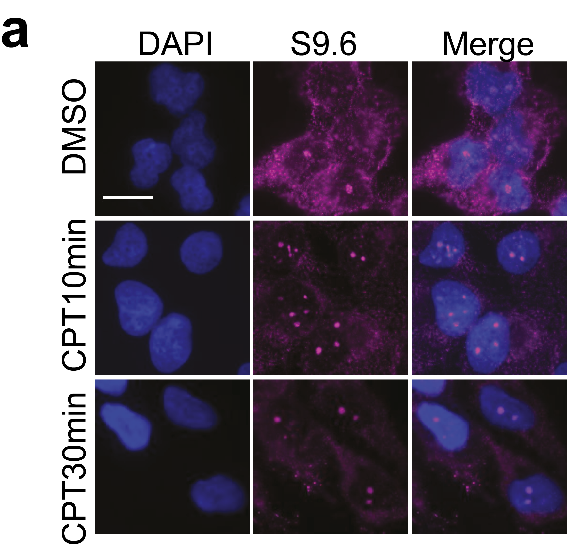


**Figure 3.1 R-loops co-localise with nucleoli.**

**a)** MRC-5 cells were treated with 25 µM CPT for 10 minutes (CPT10min) and immediately harvested for S9.6/nucleolin (R-loop/nucleoli) immunofluorescence. Representative confocal images are shown, scale bars are equal to 3 µm.

The Capranico lab demonstrated that 10min CPT treatment can induce transient R-loops. Evidence for transience came from the loss of the S9.6 signal after one-hour CPT treatment (Marinello et al., 2013). To replicate these findings, MRC-5 cells were CPT-treated for 10 or 30 min, and then immediately harvested for S9.6 immunofluorescence (Fig. 3.2a). Mock-treated cells displayed large, dim foci; while CPT-treated cells displayed small, bright, and highly abundant foci. Marinello et al., obtained similar S9.6 foci signal, but without the background staining observed here, which could be attributed to the use of different cell lines and CPT concentrations.

Nevertheless, the number and intensity of nucleolar S9.6 foci significantly increased after 10min CPT treatment and then decreased after 30min treatment, which was in line with Marinello’s findings (Fig. 3.2b). By using a nucleolar mask (obtained via immunostaining for nucleolin), a small but insignificant increase in the nuclear S9.6 intensity was observed after CPT treatment (Fig. 3.2b *right panel*), suggesting a potential limit of R-loop detection using our immunofluorescence protocol.



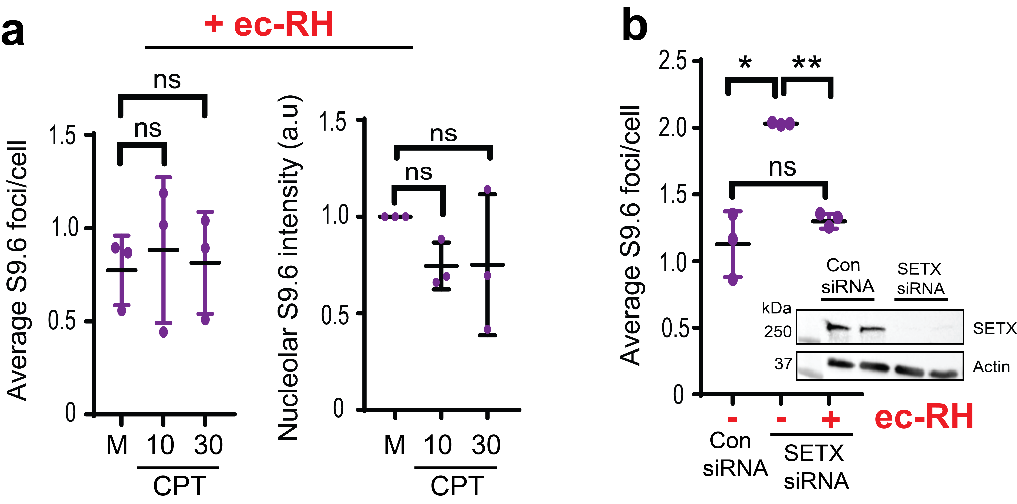
**Figure 3.2 CPT-treatment induces transient nucleolar R-loops.**

**a)** MRC-5 cells were treated with mock (DMSO) or 25 µM CPT for indicated times and immediately harvested for S9.6 immunofluorescence. Representative images are shown, scale bar = 10 µM.

**b)** Data are the average ±SD from 3 biological repeats of a), each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated (*left panel*). After correction for background noise and applying nucleolar masks, the total nucleolar (*middle panel*) and nuclear (*right panel*) S9.6 fluorescence was measured using ImageJ and normalised to mock (M, DMSO). ns; p>0.05, \* p <0.05, \*\* p<0.01, Student’s t-test.

Next, CPT-treated MRC-5s were *in vitro* incubated with *E.coli* RNase H and harvested for S9.6 immunofluorescence. ec-RNase H treatment ablated CPT-induced nucleolar S9.6 signal (Fig. 3.3a), which demonstrated the specificity of the antibody. Furthermore, siRNA depletion of R-loop helicase, senataxin (SETX), also triggered S9.6 foci accumulation, which was ablated by ec-RNase H treatment (Fig. 3.3b). This was consistent with data obtained by Skourti-Stathaki et al., 2011.

To summarise, an IF-based experimental system was established to visualise S9.6 foci. The foci were nucleolar, transiently inducible by CPT treatment, and digestible by ec-RNase H. These data suggested that S9.6 foci observed in this experimental setting matched the profile of nucleolar R-loops.



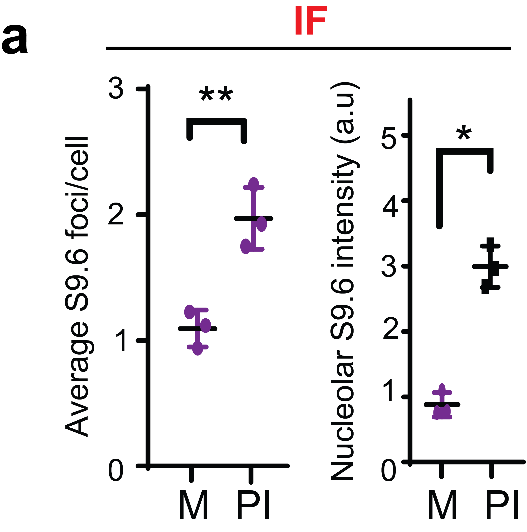
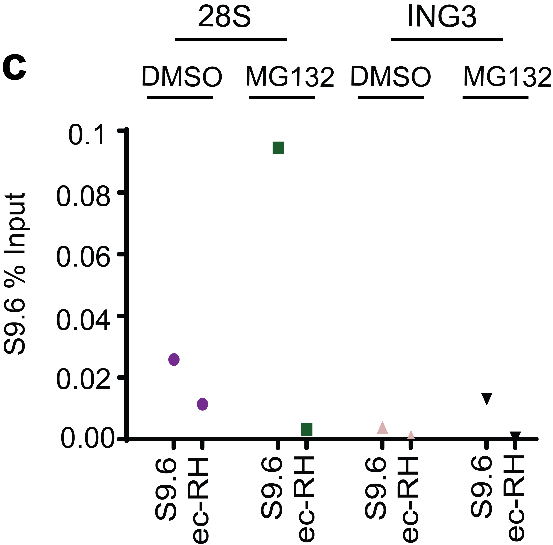
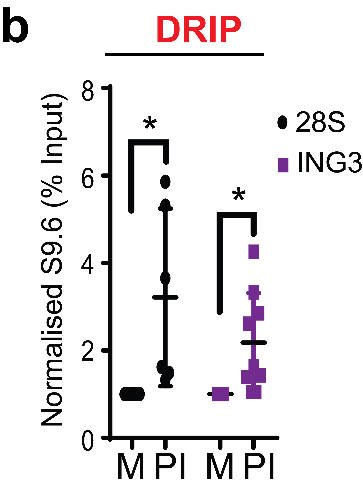
**Figure 3.3 *In vitro* ec-RNase H treatment abolishes R-loops.**

**a)** MRC-5 cels were treated with DMSO (M) or 25 µM CPT for 10 and 30 min, and then *in vitro* incubated with 25 units of *E.coli* RNase H for 1h at 37 ˚C. Next, cells were examined by S9.6 immunofluorescence. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated (*left panel*) and total nucleolar S9.6 fluorescence normalised to mock (*right panel*). ns; p>0.05, Student’s t-test.

**b)** MRC-5 cells were transfected with SETX or scrambled (Con) siRNAs, and *in vitro* incubated with 25 units of *E.coli* RNase H (ec-RH) for 1 h at 37 ˚C. Next, cells were examined by S9.6 immunofluorescence. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated. ns; p>0.05, \* p <0.05, \*\* p<0.01, Student’s t-test. *Insert*, Immunoblotting showing SETX protein expression with actin as a loading control.

Next, it was hypothesised that a DUB could be involved in R-loop homeostasis. Depletion of a DUB should distort ubiquitin levels and trigger unscheduled protein degradation, (de)activation, or relocation (Rape 2018). Therefore, global DUB inhibition followed by S9.6 immunofluorescence could shed light on the question of whether DUBs are involved in R-loop homeostasis. Unfortunately, a membrane-permeable pan DUB inhibitor (e.g. *N*-Ethylmaleimide – NEM; Tinnikov & Samuels, 2013) that could directly assess this question was not available, but a proteasomal inhibitor MG132 was (Lee and Goldberg, 1998). MG132 treatment was also predicted to distort ubiquitin levels, but rather than triggering unscheduled degradation, it would block degradation.

Interestingly, MG132 treatment alone induced a significant nucleolar R-loop accumulation phenotype, suggesting that ubiquitin recycling and protein degradation could be involved in R-loop homeostasis (Fig. 3.4a). This result was further tested by DNA/RNA immunoprecipitation followed by quantitative PCR (DRIP-qPCR) (Fig. 3.4b.c). On-bead *E.coli* RNase H treatment served as a signal-validation control. R-loops were quantified at published nucleolar (28S) and nuclear (ING3) loci (Karahan et al., 2015, Halász et al., 2017). 28S locus is transcribed by RNA Polymerase I, while ING3 locus is transcribed by RNA Polymerase II. RNA Polymerase III was not considered, as it only transcribes 5S ribosomal DNA and tRNAs (Jendrisak, 1980). Acquired DRIP-qPCR data suggested that MG132 pre-treatment triggered significant R-loop accumulation at nuclear and nucleolar loci in our experimental conditions (Fig 3.4b,c).



**Figure 3.4 MG132 treatment induces R-loops.**

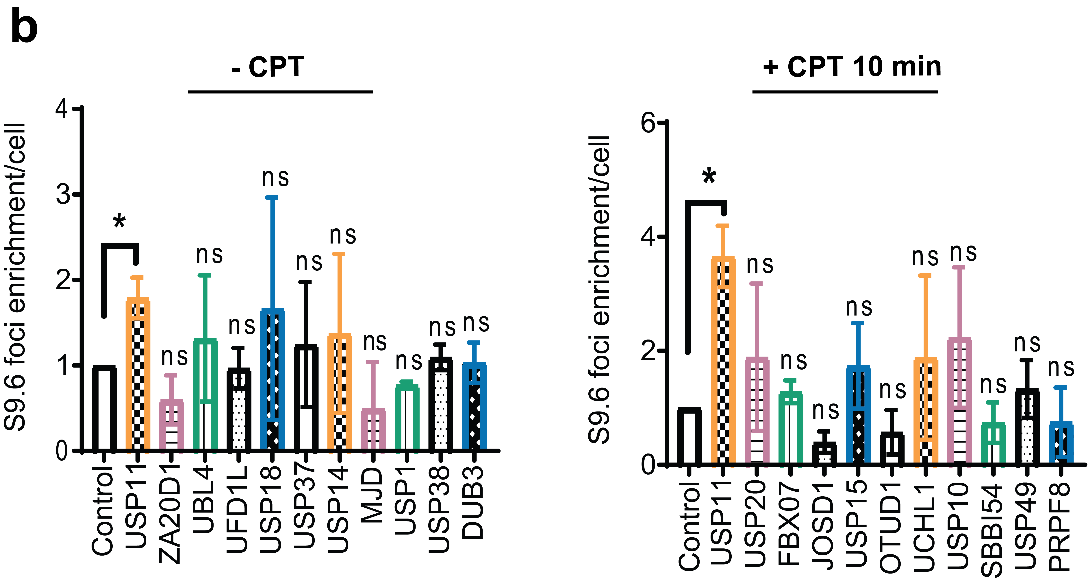
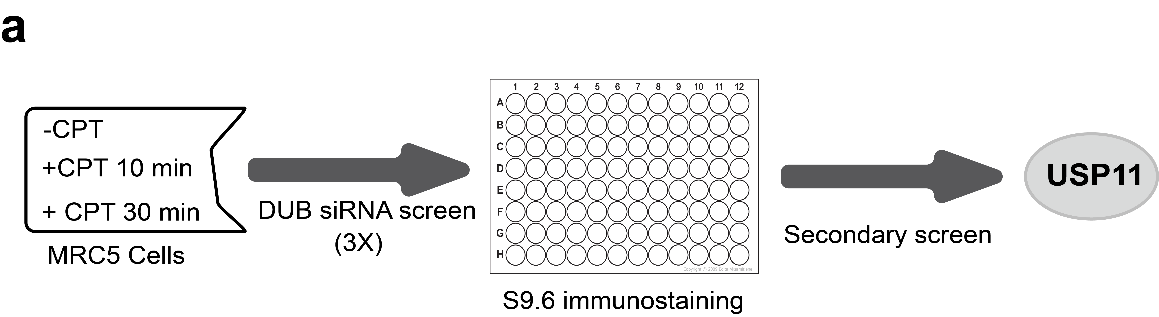
**a)** MRC-5 cells were incubated with mock (M, DMSO) or 25 µM MG132 (proteasomal inhibitor, PI) for 2 h and immediately harvested for S9.6 immunostaining. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated (*left panel*) and total nucleolar S9.6 fluorescence normalised to mock (*right panel*). \* p <0.05, \*\* p<0.01, Student’s t-test.

**b)** HEK-293T cells were incubated with 25 µM MG132 (PI, proteasomal inhibitor) for 2 h and harvested for DNA/RNA immunoprecipitation (DRIP). Quantitative PCR was conducted using primers targeting nucleolar (28S) and nuclear (ING3) loci. Normalised values to mock-treated controls are shown and represent the average ±SD from at least 5 biological repeats. \* p <0.05, Student’s t-test.

**c)** Raw data from a representative repeat of DRIP-qPCR from b). HEK-293T cells incubated with 25 µM MG132 for 2 h followed by DNA/RNA immunoprecipitation using S9.6 antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S) and nuclear (ING3) loci. *In vitro*, on-bead ec-RNase H (ec-RH) treatment served as a signal validation control.

## DUB siRNA screen data

We hypothesised that deubiquitinases (DUBs) could be involved in R-loops homeostasis. As DUBs remove conjugated ubiquitin from their target proteins, inhibition of a DUB could trigger excessive ubiquitination of an R-loop-associated protein and subsequently its degradation. To test this hypothesis, a high-throughput DUB siRNA screen was conducted using S9.6 immunofluorescence as a read out (Fig. 3.5a). 99 human DUBs were separately depleted in MRC-5 cells using siRNA pools in three different conditions: mock-treated and CPT-treated for 10 or 30min. Post CPT-treatment, the cells were immediately harvested for S9.6 immunofluorescence. Images of the cells were automatically taken by ImageXpress® Micro XLS Widefield High-Content Analysis System and processed by MetaXpress® software to yield average S9.6 foci/cell values. Three repeats of the screen revealed 13 putative candidates whose depletions triggered altered R-loop homeostasis, as compared to scrambled siRNA control cells (please see appendix-3). Among those thirteen, one deubiquitinase (USP11) stood out as its depletion triggered significant R-loop accumulation in both mock and 10min CPT-treated conditions (Fig. 3.5b).



**Figure 3.5 Results from S9.6 immunofluorescence DUB siRNA screen.**

**a)** Flow-chart depicting the design of S9.6 genetic screen. MRC-5 cells were reverse transfected on 96-well plates containing siRNA for 99 human deubiquitylases, mock-treated or treated with 25 µM CPT for 10 or 30 min and immunostainined with S9.6 antibodies. The S9.6 foci/cell values were quantified by Custom Module of MetaXpress® software and normalised to scrambled siRNA control. A secondary S9.6 screen was conducted in a 24-well format, which uncovered USP11 as a new R-loop regulator.

**b)** The DUB siRNA screen was performed as described above and data from selected DUBs are presented showing average S9.6 foci/cell normalised to scrambled siRNA control, from 3 biological repeats ± SD. \* p<0.05, ns; p **>** 0.05, Student t-test.

## USP11 validation as an R-loop regulator

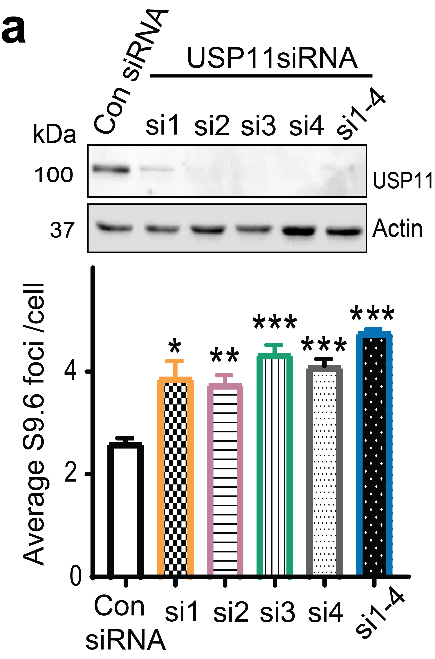
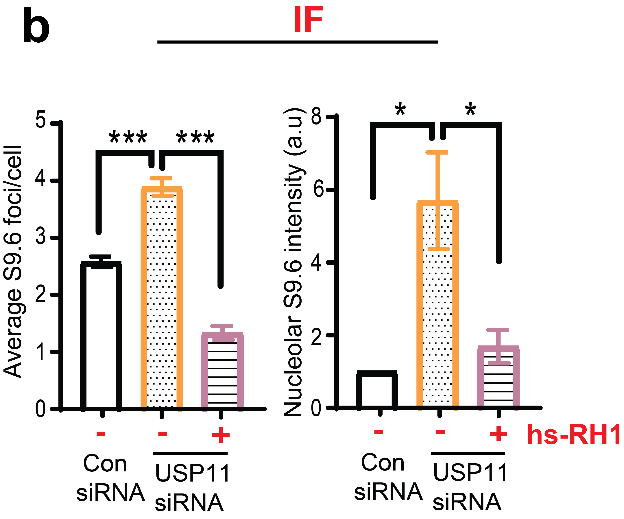
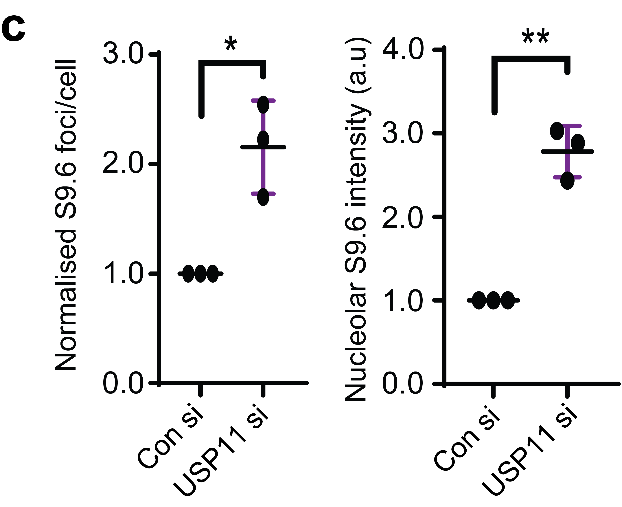
The USP11 siRNA pool was used in the DUB screen to deplete USP11 protein. The pool consisted of four siRNA oligonucleotides. Therefore, the first experiment was to test the individual siRNAs’ efficiency in depleting USP11 in MRC-5 cells (Fig. 3.6a, *top panel*). As all siRNA sequences resulted in USP11 knock-down, they were used in a subsequent S9.6 immunofluorescence experiment (Fig. 3.6a, *bottom panel*). The S9.6 foci/cell count suggested that USP11 siRNAs triggered significant S9.6 signal accumulation.

Next, the nucleolar S9.6 signal upon USP11-depletion was tested by over-expression of *Homo sapiens* eGFP-RNase H1 in MRC-5 cells (Fig. 3.6b) and then further tested in U-2 0S cell line (Fig. 3.6c).

Next, the R-loop accumulation phenotype in USP11-depleted cells was examined by DNA/RNA immunoprecipitation (DRIP-qPCR). Primer pairs targeting nucleolar (28S & R7, Shen et al., 2017) and nuclear (ING3 & actin, Groh et al., 2014) loci were used for quantitative PCR. The acquired data suggest that upon USP11-depletion, R-loops were enriched at both nucleolar and nuclear loci (Fig. 3.7).

Finally, nucleolar and nuclear R-loop accumulation phenotype in USP11-depleted cells was investigated using a recently published DRIP-qPCR protocol (Sanz and Chedin, 2019). Seven published primer pairs were used for quantitative PCR, two of which were negative controls (MYADM-neg, SNRPN-neg) (Fig 3.8a). Together, the acquired data suggest that USP11 is a novel nuclear- and nucleolar-R-loop regulator.

As it was shown using two different DRIP protocols that USP11 depletion triggers R-loop accumulation at 28S, R7, ING3, and actin loci, it was hypothesised that USP11 might bind to these loci to facilitate R-loop resolution. Cross-linking chromatin immunoprecipitation assay (X-ChIP) was performed using α-USP11 antibody followed by quantitiave PCR (Fig. 3.9a). In line with previous results, our data suggest that USP11 is significantly enrinched at the four-probed loci, suggesting that USP11 localises to chromatin to protect from abberant R-loop accumulation.

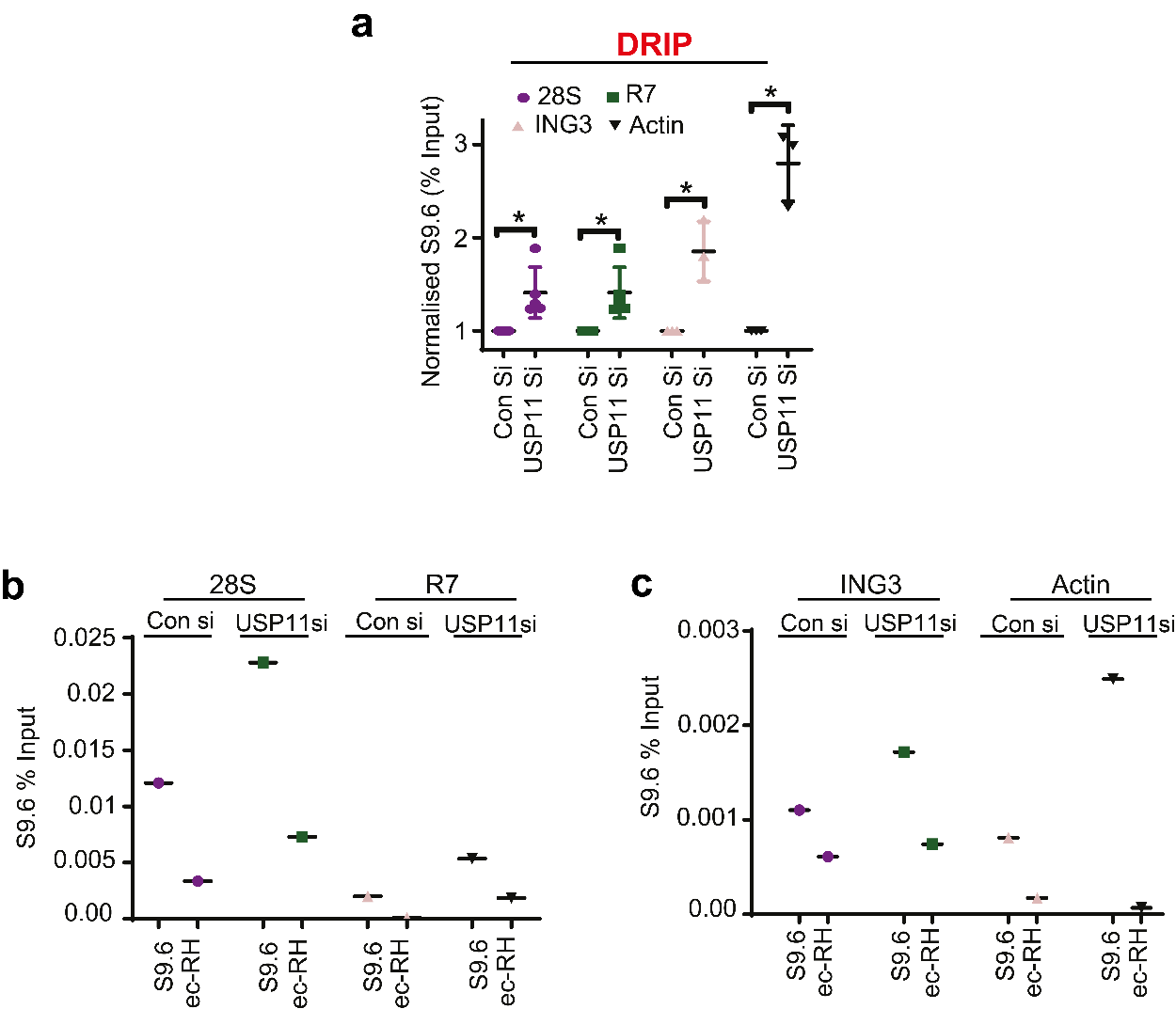


**Figure 3.6 USP11-depletion triggers R-loop accumulation.**

**a)** MRC-5 cells were transfected with scrambled siRNA (Con), 4 different siRNA oligos targeting different regions in *USP11* (si1-4) and pooled siRNA containing all 4 siRNAs (si1-4). Depletion of USP11 was examined by immunoblotting (*top panel*) and the average ±SD number of S9.6 foci/cell was calculated from 3 biological repeats, each consisting of at least 100 cells (*bottom panel*). \* p <0.05, \*\* p<0.01, \*\*\* p<0.001, Student’s t-test.

**b)** MRC-5 cells were transfected with scrambled (Con) or USP11 siRNAs and human eGFP-RNase H1 (hs-RH1). Cells were harvested for S9.6 immunostaining. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated (*left panel*) and total nucleolar S9.6 fluorescence normalised to scrambled (*right panel*). \* p <0.05, \*\*\* p<0.001, Student’s t-test.

**c)** U-2 0S cells were transfected with scrambled (Con) or USP11 siRNAs and harvested for S9.6 immunostaining. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell (*left panel*) and total nucleolar S9.6 fluorescence normalised to scrambled (*right panel*) were calculated. Normalisation was conducted due to observed high variability in S9.6 foci number and intensity in U-2 0S cells (as opposed to MRC-5 cells). \* p <0.05, \*\* p<0.01, Student’s t-test.



**Figure 3.7 USP11-depletion leads to R-loop accumulation at nucleolar and nuclear loci.**

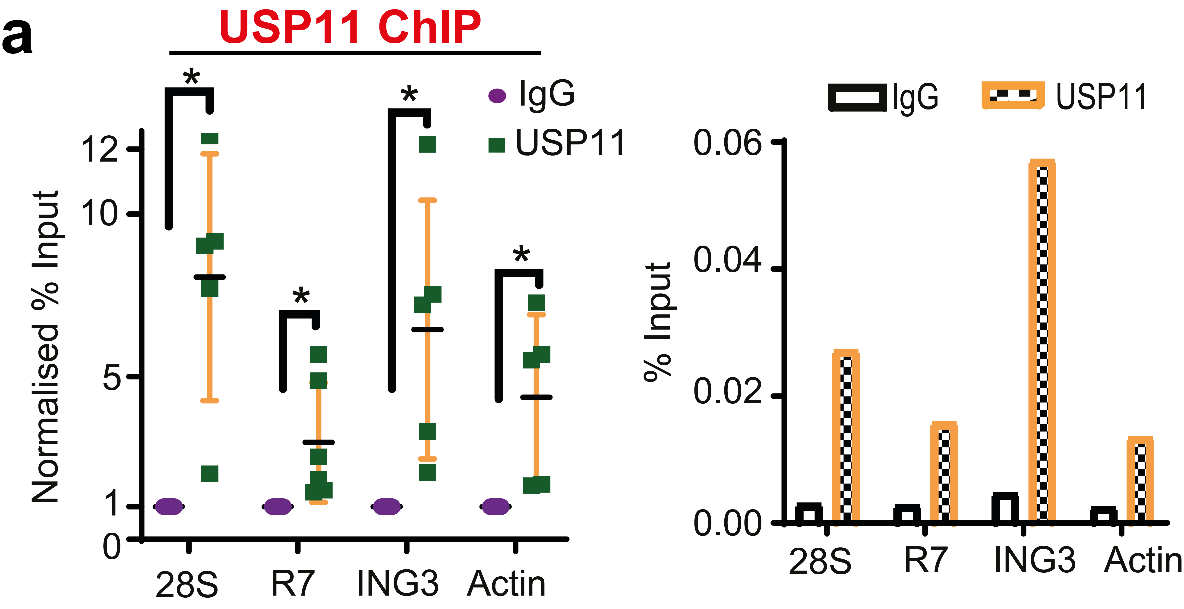
**a)** Lysates from Scrambled (Con Si) and USP11-depleted HEK-293T cells (USP11 Si) were subjected to DNA/RNA immunoprecipitation using S9.6 antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3 and actin) loci. Normalised values to mock-treated controls are shown and represent the average ±SD from at least 3 biological repeats. \* p <0.05, Student’s t-test.

**b)c)** Raw data from a representative repeat of DRIP-qPCR from a). *In vitro*, on-bead ec-RNase H (ec-RH) treatment served as a signal validation control.



**Figure 3.8 USP11-depletion leads to R-loop accumulation at nucleolar and nuclear loci according to Sanz and Chedin (2019) DRIP protocol.**

**a)** Lysates from Control (Con siRNA) and USP11 depleted HEK-293T cells (USP11 siRNA) were subjected to a more recently published DNA/RNA immunoprecipitation (DRIP) protocol using S9.6 antibodies (Sanz and Chedin, 2019). Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3, actin, EGR1, SNRPN-neg, MYADM-neg) loci. SNRPN-neg and MYADM-neg loci are negative controls. *In vitro*, ec-RNase H (ec-RH) treatment served as a signal validation control. Raw % input values from a representative experiment are shown (*left panel*) and the normalised values to mock-treated cells were quantified (*right panel*), and data represent the average ± SD from 3 biological repeats. ns; p **>** 0.05, \* p<0.05, \*\* p<0.01, Student’s t-test.



**Figure 3.9 USP11 localises to R-loop forming loci according to cross-linking ChIP.**

**a)** Lysates from HEK-293T cells were subjected to cross-linking chromatin immunoprecipitation using IgG and α-USP11 (USP11) antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3, actin) loci. Normalised % Input values to IgG-immunoprecipitated lysates were quantified (*left panel*), and data represent the average ± SD from 5 biological repeats. The raw % input values from a representative experiment are shown in the *right panel*. \*p<0.05, Student’s t-test.

To summarise data from this chapter:

* In spite of growing concerns regarding the S9.6 antibody, the obtained data suggest that an S9.6 IF protocol was optimised; where transient, nucleolar, RNase H-dependent R-loops can be visualised.
* MG132 treatment triggered R-loop accumulation at nuclear ING3 and nucleolar 28S loci, suggesting that proteasomal degradation and ubiquitination levels might be involved in R-loop homeostasis.
* S9.6 IF DUB siRNA screen revealed a few potential candidates for novel R-loop regulators.
* Ubiquitin Specific Protease 11, USP11, was shown to protect MRC-5 and U-2 0S (immunofluorescence data) and HEK-293T cell lines (DRIP-qPCR data) from R-loop accumulation in our experimental settings. USP11-depletion triggered R-loop accumulation at both nucleolar and nuclear loci.
* USP11 was shown to localise to R-loop forming loci according to our cross-linking ChIP data.

## Discussion

## MG132 treatment leads to R-loop accumulation phenotype

Results presented here suggest that proteasomal inhibition might disturb normal R-loop homeostasis. As proteasomal degradation is tightly linked to ubiquitination events, gathered data suggest that E3 ligases could regulate R-loop homeostasis. However, not all functions of E3 ligases are related to the proteasome, as they also alter protein activation and localization. In fact, that was the case for the only E3 ligase, Bre1, which was associated with R-loops (Chernikova et al., 2012). Chernikova and her colleagues suggested that Bre1 E3 complex monoubiquitinates histone H2B, which in turn orchestrates correct histone mRNA processing. In absence of Bre1, reduced H2B monoubiquitination results in aberrant mRNA processing and R-loop accumulation. In order to investigate the role of proteasome in R-loop homeostasis, potential experiments are suggested in the Chapter 7: General discussion.

## Evaluation of the DUB siRNA screen

The results of the DUB siRNA screen were dependent on the efficiency of siRNA sequences (Dharmacon, ON-TARGETplus siRNA library) and the reliability of the S9.6 immunofluorescence method.

It is plausible to say that some R-loop regulators, among the DUB family, were missed due to inefficient siRNA pools. USP11 siRNA pool was shown to be highly efficient by western blotting, and hence the robust USP11-depletion triggered a significant R-loop accumulation. As only USP11 was validated after the screen, false negatives or positives were not discovered. A likelihood of a presence of false negatives was higher than that of false positives due to Dharmacon’s patented ON-TARGETplus technology, which assures high specificity and low off-target effects. Any false negatives could be discovered by repeating the screen using 100nM (instead of 50nM) final concentration of published siRNA pools to ensure substantial knockdowns. However, the disadvantages of this approach are increased cell death, due to the toxicity of transfection reagents, and increased probability of off-target effects (Jiang et al., 2011).

Growing concerns regarding S9.6 antibody in the context of immunofluorescence (Vanoosthuyse, 2018) are described below. A thorough investigation of S9.6 IF experiments from different studies suggested that S9.6 antibody does not produce a consistent signal pattern (for instance: Bhatia et al., 2014, Sollier et al., 2014, Koo et al., 2015, Yasuhara et al., 2018). Even though this could be attributed to different cell types, it was shown by the Proudfoot lab that it might be the fixation method that matters for preserving R-loops (Skourti-Stathaki et al., 2014). To date, no consensus has been reached on how to carry out S9.6 IF protocol (Vanoosthuyse, 2018). Most of the published S9.6 IF data demonstrate strong cytoplasmic and nucleolar signals. The nucleolar signal can be attributed to high transcription rates of RNA polymerase I, which promotes R-loop formation (Shen et al., 2017). The cytoplasmic signal was suggested to originate from mitochondria (Ginno et al., 2012; Silva et al., 2018). However, the Ishii lab showed that the majority of the cytoplasmic S9.6 signal might come from the outside of mitochondria – cytosol (Koo et al., 2015). Interestingly, Koo et al. demonstrated that the cytosolic S9.6 signal was dependent on RNA polymerase III activity, at least in lung carcinoma cells, suggesting a presence of RNAPIII-dependent R-loops.

There is a possibility that a part of cytosolic and mitochondrial signal could be assigned to double-stranded RNA, as J2 (anti-dsRNA) staining depicted a strong presence of dsRNA in the cytosol and mitochondria of HeLa cells (Skourti-Stathaki et al., 2014; Dhir et al., 2018). The Aguilera lab demonstrated that the majority of S9.6 staining coming from cytosol and mitochondria can be ablated by RNase III, which is specific to dsRNA (Silva et al., 2018). Therefore, it is crucial for the scientific community to reach consensus on how to preserve genuine cytosolic, mitochondrial, nuclear, and nucleolar R-loops and how to quantify them.

Given the nature of S9.6 immunofluorescence signal in this study, it was decided to focus on the S9.6 signal found in nuclei. The obvious limitation of this approach was a failure to quantify the cytosolic and mitochondrial signals. The number of S9.6 foci per nucleus was counted and total intensity of nucleolar and pan-nuclear S9.6 staining was measured. The intensity measurements were always corrected for background and conducted with a help of nucleolar masks to ensure reliable results.

S9.6 foci/cell and nucleolar intensity data both demonstrated R-loop accumulation phenotype in USP11-depleted cells. However, pan-nuclear intensity data did not show significant changes after the depletion (data not shown). As DRIP-qPCR data suggested that USP11-depletion triggered R-loop accumulation at nuclear loci, it was hypothesised that the bright S9.6 immunofluorescence foci in USP11-depleted cells might have masked any remaining pan-nuclear signal, leading to negative results. Therefore, the limitation of our S9.6 immunofluorescence protocol was the detection of nuclear S9.6 staining.

It is worth noting that methanol fixation preserves dsRNA in nuclei of HeLa cells (Skourti-Stathaki et al., 2014). As methanol:acetone fixation was used in this study, it is plausible that some dsRNA structures were preserved and recognized by the S9.6 antibody, when genuine R-loops were scarce.

Finally, S9.6 was shown to bind to at least 6bp-long R-loops *in vitro* (Phillips et al., 2013). Therefore, accumulation of R-loops shorter than 6bp was probably not tested for in our experimental setting. However, as R-loops were recorded to be ~100-600bp-long (Yu et al., 2003), this limitation may not be a significant one.

## USP11 might be a novel R-loop regulator

The ultimate aim of this chapter was to carry out a high-throughput DUB siRNA screen in order to find novel proteins involved in R-loop homeostasis. USP11 was identified as a potential R-loop regulator. USP11-depletion triggered a significant R-loop accumulation in our screening conditions. Unchallenged USP11-depleted cells displayed 2-fold R-loop enrichment, which was exacerbated by 10min CPT treatment to 4-fold.

We decided to test the R-loop accumulation phenotype in unchallenged USP11-depleted cells by immunofluorescence and DRIP-qPCR protocols. *In vivo* hs-RNase H1 over-expression was used to validate the S9.6 signal obtained by immunofluorescence. *In vitro* ec-RNase H treatment was used to assay the specificity of the signal from DRIP-qPCR assays. As expected, the signal was reduced in all cases, demonstrating its specificity. Alternatively, the signal could have been examined by a combination of ec-RNase H1 and RNase III treatments.

Interestingly, USP11 depletion was found to trigger R-loop accumulation at nucleolar and nuclear loci, according to our DRIP-qPCR data. This suggested that USP11 might control a protein that is involved in maintaining global R-loop homeostasis. Therefore, USP11 could regulate, for instance, a global transcription suppressor, a master regulator of RNA export, or a pan-R-loop resolving protein.

At the time of the discovery, USP11 was only described in twelve manuscripts and mentioned in further eight. There was no information published on how USP11 could regulate R-loops. As a result, it was decided to investigate the role of USP11 in R-loop homeostasis.

# Chapter 4: USP11 might regulate the stability of senataxin



## Introduction

After suggesting that USP11 might be a novel R-loop regulator, literature analysis was conducted on USP11 and R-loop-resolving proteins. This led to a series of experiments that shed light on a connection between USP11 and an R-loop-specific helicase – senataxin. Therefore, overviews of both proteins are presented below.

## Overview of USP11

*USP11* gene is located on X-chromosome (Swanson et al., 1996) and encodes for 105kDa USP11 protein. USP11 is mostly found in the nucleus, where it interacts with a number of substrates that control microtubule nucleation, proliferation, apoptosis, and replication of human papillomavirus (Ideguchi et al., 2002, Sun et al., 2010, Deng et al., 2018; Lin et al., 2008).

USP11 was also linked to DNA damage repair. The Cortez lab gathered data showing that USP11-depleted cells are sensitive to a PARP inhibitor - olaparib (Wiltshire et al., 2010). Olaparib blocks Single-Strand Break (SSB) repair pathway so that treated cells can only repair lesions with Double-Strand Break (DSB) repair pathways. Since USP11-depleted cells are sensitive to olaparib, they most likely have a defective DSB repair.

Accordingly, it was demonstrated that Homologous Recombination (HR) factors RAD51 and 53BP1 were misregulated in the USP11-depletion background (Wiltshire et al., 2010). In line with this, Zhao lab showed that USP11-depletion triggered high levels of ubiquitinated γH2A.X, which in turn led to prolonged, misregulated retention of 53BP1 at the sites of DSBs (Yu et al., 2016).

To corroborate the link between USP11 and HR pathway, USP11-depleted cells were shown to be sensitive to bleomycin (Wiltshire et al., 2010). Bleomycin is an anti-cancer drug, which was shown to trigger DNA breaks (Suzuki et al.1969). Therefore, all gathered data suggested that USP11 is likely to be an important HR factor. The Brody lab demonstrated that cancer cells, which possess undamaged BRCA2 (HR factor) and over-express USP11, are sensitive to a small molecule inhibitor Mitoxantrone (USP11 inhibitor) (Burkhart et al., 2013). The Aaranson lab presented evidence of physical interaction between BRCA2 and USP11, but they did not find BRCA2 to be directly deubiquitinated by USP11. Therefore, they suggested that BRCA2-USP11 interaction might only aim to recruit USP11 to the sites of DNA lesions (Schoenfeld et al., 2004). If this is true, it would resemble the mechanism by which the R-loop helicase senataxin is localized to R-loops by BRCA1 (another HR factor) (Hatchi et al., 2015).

The link between BRCA2 and USP11 was further studied by the Durocher lab. The authors found, by a series of immunoprecipitation experiments, that assembly of BRCA2-PALB2-BRCA1 complex, required for homologous recombination (HR), is inhibited by an E3 ligase KEAP1, but promoted by USP11. Deletion of KEAP1 in USP11-knockout cell line resulted in a rescue to olaparib sensitivity, suggesting restoration of HR (Orthwein et al., 2015).

Additionally, USP11 was shown to be degraded upon accumulation of DNA damage in G1 phase; probably to prevent repair by homologous recombination in the absence of sister chromatids (Orthwein et al., 2015).

It was found *in vitro* using a global proteomic screen that USP11 interacts physically with a number of RNA-associated proteins, including transcription-elongation factors and a poorly characterized mRNA export protein - RAE1 (Sowa et al., 2009; Stockum et al., 2018). RAE1 is an interesting candidate for R-loop studies, as mRNA-export factors were previously suggested to regulate R-loop homeostasis (Gonzalez-Aguilera et al., 2008). Therefore, the USP11-RAE1 interaction could be another link to R-loop accumulation phenotype found in USP11-depleted cells.

USP11 promotes translation. It was shown that phosphorylated USP11 interacts with translation initiation complex, where it deubiquitinates Eukaryotic Initiation Factor 4B (eIF4B). Deubiquitination of eIF4B results in its stabilization, which in turn promotes mRNA translation (Kapadia et al., 2018). Following this logic, depletion of USP11 could inhibit translation of an R-loop associated protein and lead to the R-loop accumulation phenotype.

USP11 was shown to act both as a tumour suppressor and an oncogene by various groups. Both low and high expression levels of USP11 were linked to poor cancer prognosis (Wu et al., 2014; Zhang et al., 2018). It was demonstrated that elevated levels of USP11 stabilise TGFβ receptor type II, which leads to an epithelial-mesenchymal transition in breast cancer cells, and subsequently to metastasis (Garcia et al., 2018).

On the other hand, low expression levels of USP11 were shown to reduce the capacity of Global Genome Nucleotide Excision Repair pathway (GG-NER) (Shah et al., 2017). GG-NER is involved in repair of bulky lesions located at transcription-free sites; and can be activated by damage caused by ultraviolet B radiation (UVB). It was shown that upon UVB exposure, USP11 becomes chromatin-bound and deubiquitinates, NER factor, XPC. Deubiquitination of XPC promotes its retention at sites of damage, where it subsequently facilitates the repair mechanism (Shah et al., 2017).

We demonstrated that USP11 depletion triggered the accumulation of nucleolar and nuclear R-loops. Therefore, it was of interest whether there were any R-loop-associated proteins that were previously shown to resolve nuclear and nucleolar R-loops. Two laboratory groups independently demonstrated that senataxin, R-loop-specific helicase, co-localizes to nucleus and nucleoli (Yüce & West, 2013, Chen et al., 2006). Even though the role of senataxin is not described in nucleoli, its R-loop-resolving activities are well described in nuclear genes (Kim et al.1999, Skourti-Stathaki et al., 2011; Hatchi et al., 2015; Martin-Tumasz et al., 2015; Cogen et al., 2018). Therefore, USP11-SETX connection, if true, could be another explanation for R-loop accumulation phenotype in USP11-depleted cells.

## Overview of senataxin

*SETX* gene encodes for 302.8 kDa senataxin protein (SETX) (Suraweera et al., 2007; Skourti-Stathaki et al., 2011), which was found to be mutated in two neurodegenerative disorders: recessively inherited Ataxia-ocular Apraxia (AOA2), and dominantly inherited Juvenile Amyotrophic Lateral Sclerosis Type 4 (ALS4) (Moreira et al., 2004; Chen et al., 2004). ALS4 and AOA2 are characterised by progressive apoptosis of motor neurons, muscle weakness, and atrophy.

ALS4 and AOA2 patients possess mutations in the C- or N-terminal domains of SETX (Chen et al., 2004; Moreira et al., 2005, Duquette et al., 2005). The N-terminal domain is the protein-binding domain, while the C-terminal domain contains DEAD-Box helicase, which resolves R-loops (Skourti-Stathaki et al., 2011).

As senataxin contains nuclear localization signal (NLS), it is predominantly found inside nuclei, although minimal amounts are also present in the cytoplasm (Skourti-Stathaki et al., 2011; Suraweera et al., 2007). The eGFP-tagged N-terminal domain of SETX was demonstrated to localize to nucleoli (Chen et al., 2006, Yüce & West, 2013), while full-length senataxin was co-immunoprecipitated with nucleolin, the key component of nucleoli (Suraweera et al., 2009).

The Gromak lab demonstrated that senataxin aids RNA Polymerase II transcription termination (Skourti-Stathaki et al., 2011). Senataxin depletion triggered a formation of abnormally long mRNA transcripts of β-globin and MAZ4 genes. They demonstrated that as RNA Polymerase II stalls at G-rich pause sites, R-loops are formed and subsequently unwound by senataxin. The unwinding event triggers degradation of the RNA moiety of an R-loop by a 5’-3’ exoribonuclease XRN2, and consequently, promotes transcription termination.

SETX was demonstrated to interact with five ubiquitin-SUMO (small ubiquitin-like modifier) modifiers by a yeast two-hybrid assay (Bennett et al., 2013). A band shift of N-terminal domain of SETX was recorded upon an NEM treatment of HEK-293T cells. NEM is an inhibitor of ubiquitin/SUMO cleavage. Therefore, the N-terminal domain of SETX was suggested to be a subject of ubiquitin-SUMO modification. However, putative E3 ligases and DUBs that could process these modifications were not suggested (Bennett et al., 2013).

SUMOylated SETX was shown to interact with a component of exosome complex, RRP45, which facilitates RNA turnover and quality control. This interaction was hs-RNase H1 dependent, as over-expression of RNase H1 reduced co-localization of SETX and RRP45. Therefore, it was suggested that exosome and SETX cooperate to resolve R-loops and degrade RNA moieties (Richard et al., 2013).

SETX was shown to form nuclear foci, which co-localized with markers of DNA damage. Interestingly, SETX foci increased in number following transcription and replication stress; but decreased following transcription inhibition or hs-RNase H1 over-expression (Yüce & West, 2013). The authors suggested that senataxin might play a key DNA damage repair role at sites of transcription-replication collisions, where it resolves R-loops (Yüce & West, 2013).

A series of ChIP and DRIP experiments revealed that senataxin is recruited to R-loops at transcription pause sites by BRCA1 (Hatchi et al., 2015). Depletion of BRCA1 reduced senataxin binding to R-loops and triggered an accumulation of single-stranded DNA breaks and γH2A.X foci. γH2A.X is a marker of DNA damage. Both γH2A.X foci and ssDNA breaks were reduced by over-expression of hs-RNase H1. ssDNA breaks were found within the unannealed DNA strands, which were displaced by R-loops. γH2A.X ChIP revealed that γH2A.X foci formed in close proximity to R-loops but not in R-loop-free loci. Therefore, Hatchi et al. provided evidence that BRCA1/SETX interaction protects human cells from genome instability by removing R-loops from transcriptionally active sites.

The Greenblatt lab corroborated the link between RNA Polymerase II and senataxin (Zhao et al., 2015). They found that symmetrically dimethylated arginine 1810 (R1810me2) of the C-terminal domain of POLR2A, which is a subunit of RNA Polymerase II, recruits SMN protein (Survival of motor neuron), which in turn recruits senataxin. The interaction between RNA Polymerase II and senataxin was reduced in the absence of SMN, suggesting that SMN stabilised the complex. Exoribonuclease XRN2, which degrades RNA moiety of R-loops in a senataxin-dependent manner (Skourti-Stathaki et al., 2011), was also found to be in the complex with SMN and RNA Polymerase II, suggesting the role of SMN in facilitating transcription termination through senataxin. As both SETX and SMN are mutated in neurological disorders, the authors suggested that defective RNA Polymerase II/SMN/senataxin pathway might contribute to neuropathies (Zhao et al., 2015).

The Legube lab investigated the link between senataxin and double-strand breaks. They found that senataxin is recruited to sites of DSBs within transcriptionally active genes (Cohen et al., 2018). Interestingly, a series of ChIP-seq and DRIP-seq experiments allowed Cohen et al. to observe an increase in R-loop signal in a span of 10kb around DSBs. However, upon SETX binding, the R-loop signal was reduced, but only in 1-2kb proximity to the binding site. Their data suggest that senataxin removes R-loops in the direct vicinity of double-strand breaks (Cohen et al., 2018).

Moreover, Cohen et al. demonstrated that senataxin depletion reduced the formation of RAD51 foci (HR factor) and triggered chromosomal translocations, which were partially rescued by RNase H1 over-expression. Therefore, the authors suggested that senataxin removes R-loops at sites of DSBs and promotes homologous recombination via RAD51 recruitment (Cohen et al., 2018).

The current literature strongly links senataxin to R-loops and RNA Polymerase II. Even though it was demonstrated that senataxin co-localizes with nucleoli and binds to nucleolin protein (Chen et al., 2006, Yüce & West, 2013; Suraweera et al., 2009), future studies are required to describe its role in ribosomal DNA and in conjunction with RNA Polymerases I & III in human cells. If true, that will be in line with current *S.cerevisiae* literature where senataxin homologue, Sen1p, was demonstrated to aid termination of all three RNA polymerases (Kawauchi et al., 2008; Mischo et al., 2011; Rivosecchi et al., 2019).

## Aims and objectives

The aim of this chapter was to elucidate the link between USP11 and R-loops. The objectives were to:

* Based on the knowledge at the time, suggest a substrate for USP11 that regulates R-loop homeostasis directly or indirectly.
* Test the epistasis between USP11 and the substrate in context of R-loop homeostasis.
* Investigate whether USP11 controls stability of the substrate post-translationally.
* Investigate whether USP11 depletion triggers accumulation of ubiquitinated species of the substrate.

## Results

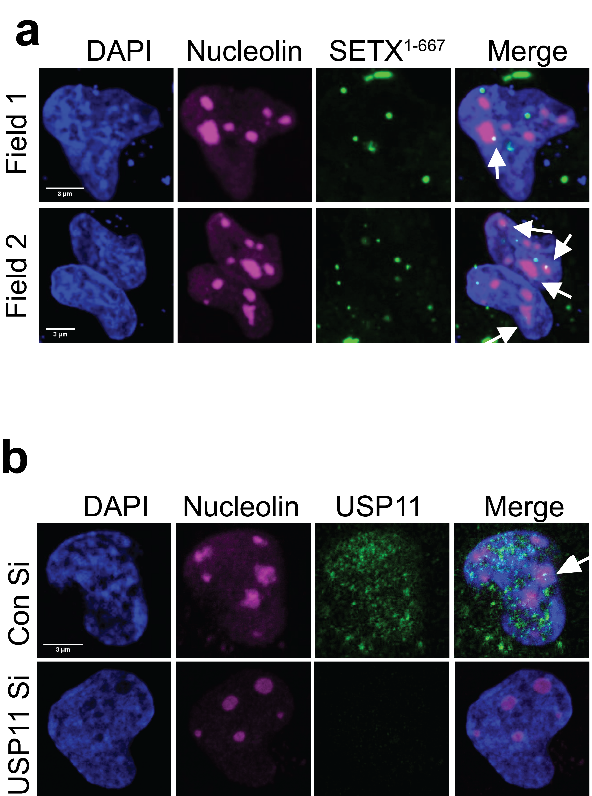
At the beginning of this study, three possible links between USP11 and R-loop homeostasis were examined. First, the impact of USP11 depletion on global mRNA levels was assessed. The rationale behind this experiment was to investigate if USP11 depletion triggers high transcription rates and therefore elevated R-loop levels (the preliminary data from this experiment was inconclusive; data not shown).

Next, the impact of USP11 depletion on the stability of mRNA export factor, RAE1, was assessed in presence of cycloheximide (Sowa et al., 2009; Stockum et al., 2018). Cycloheximide is a protein biosynthesis inhibitor (Obrig et al., 1971). The rationale was to investigate if RAE1 half-life was shortened when USP11 was depleted. Unscheduled degradation of RAE1 should block mRNA export to some extent, and consequently trigger R-loop accumulation (Gonzalez-Aguilera et al., 2008) (the data from this experiment was negative and did not support the rationale; data not shown).

Finally, senataxin was examined as a possible substrate for USP11. This led to a successful chain of experiments described below, suggesting that senataxin and USP11 act in the same pathway. Gathered data suggest that USP11 interacts with senataxin and regulates senataxin ubiquitination levels.

## USP11 and SETX are epistatic

Data described in the previous chapter suggest that USP11 feeds into nuclear and nucleolar R-loop homeostasis. Senataxin is a well-known R-loop helicase that was shown to co-localize to nucleoli (Yüce and West, 2013; Chen et al,. 2006). The N-terminal domain of SETX1-650 was demonstrated to be a subject of Ubiquitin/SUMO modification (Bennett et al., 2013). Therefore, it was hypothesised that USP11 and SETX might be a part of the same R-loop resolution pathway. MRC-5 cells were transfected with the eGFP-tagged N-terminal domain of SETX1-667 and stained for nucleoli. In line with the current literature, eGFP-SETX1-667 was found to co-localize with nucleoli (Fig.4.1a). Next, MRC-5 cells were co-stained for USP11 and nucleoli. Confocal images revealed that USP11 was not excluded from nucleoli (Fig. 4.1b), meaning that USP11 and SETX could interact in nucleoli or nucleoplasm.



**Figure 4.1 eGFP-SETX1-667 and USP11 might interact in nucleoli.**

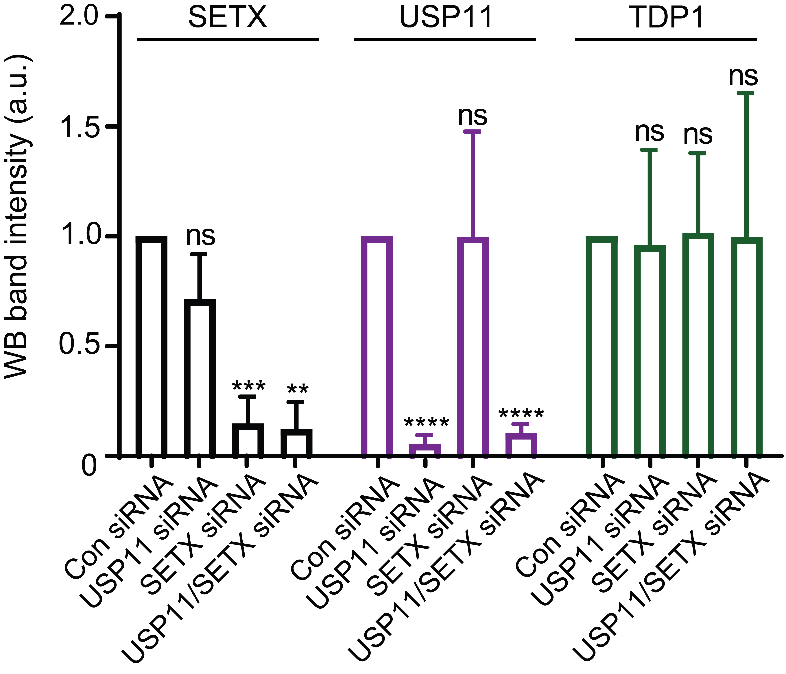
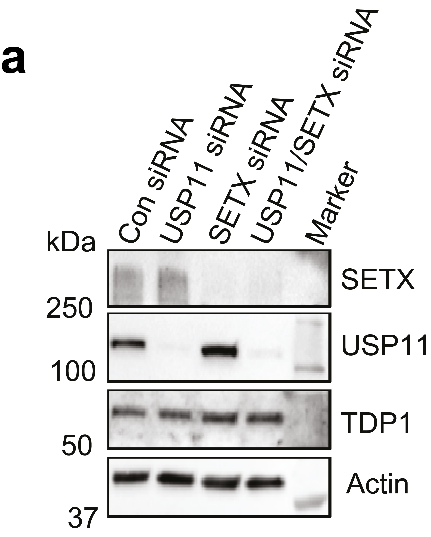
**a)** MRC-5 cells were transfected with eGFP-SETX1-667 and examined for nucleolar localisation using anti-nucleolin immunofluorescence. Representative confocal images are shown; scale bars are equal to 3 µm. Arrows depict sites of co-localization between eGFP-SETX1-667 and nucleolin.

**b)** MRC-5 cells were transfected with scrambled (Con Si) or USP11 siRNA and examined for USP11 nucleolar localisation using anti-USP11/nucleolin immunofluorescence. Representative confocal images are shown; the scale bar is equal to 3 µm. The arrow depicts a site of co-localization between USP11 and nucleolin.

Next, USP11 and SETX siRNAs were examined by immunoblotting (Fig. 4.2a). TDP1 was used as a positive control. As siRNA depletions were significant, it was decided to proceed and test whether USP11 and SETX were epistatic.

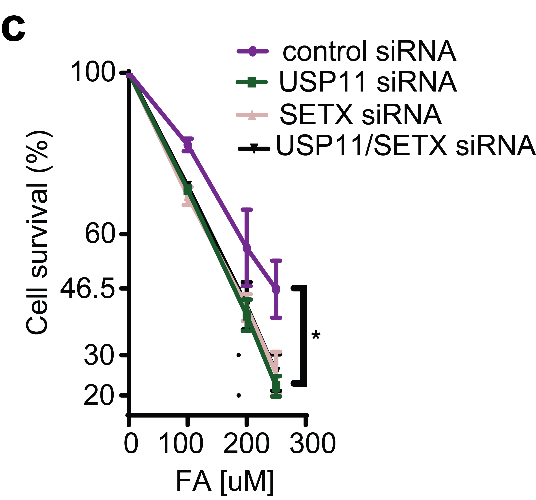
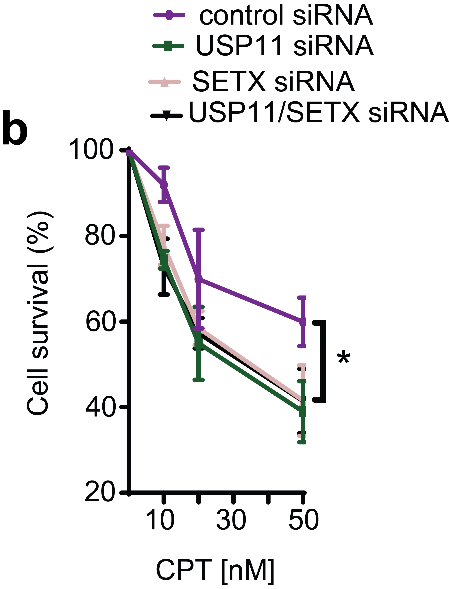
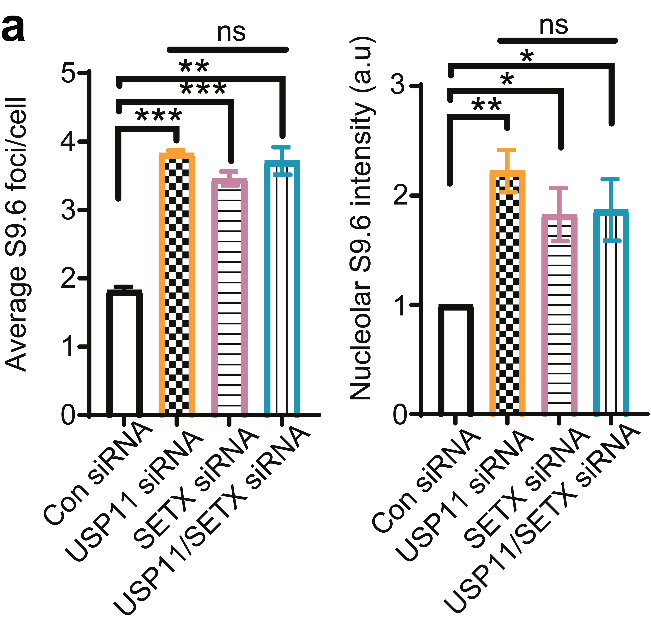
S9.6 immunofluorescence revealed that single knockdowns of USP11 and SETX triggered 2-fold R-loop accumulation in MRC-5 cells (Fig.4.3a). Double depletion of USP11 and SETX did not exacerbate the R-loop enrichment; and was not significantly different from the single knockdowns of USP11 or SETX, suggesting epistasis.

To further investigate USP11 and SETX epistasis, clonogenic cell survival assays were conducted (Fig.4.3bc). MRC-5 cells were treated with camptothecin (CPT) or formaldehyde (FA). Both CPT and FA were previously described to induce R-loop formation (Schwab et al., 2015; Tan et al., 2017). It was found that SETX-depleted, USP11-depleted, and SETX/USP11-depleted cells were significantly hypersensitive to the R-loop inducing genotoxins. The double depletions did not differ from single depletions alone. Therefore, it was concluded that USP11 and SETX are epistatic in our experimental conditions.



**Figure 4.2 Depletion of USP11 and SETX.**

**a)** Lysates from MRC-5 cells treated with scrambled (Con), USP11, SETX or USP11 and SETX siRNAs were fractionated by SDS-PAGE and analysed by immunoblotting (*top panel*). USP11, SETX and TDP1 band intensities were normalised to Actin and presented as fold reduction compared to levels in control cells (*bottom panel*). Data are the average of 4 biological repeats and presented as mean ± SD. \*\* p<0.01, \*\*\* p<0.001, \*\*\*\* p<0.0001, ns; p>0.05, Student’s t-test.



**Figure 4.3 USP11 and SETX are epistatic.**

**a)** MRC-5 cells treated with scrambled (Con), USP11, SETX or USP11 and SETX siRNA were harvested for S9.6 immunofluorescence. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated (*left panel*) and total nucleolar S9.6 fluorescence normalised to scrambled (*right panel*). ns; p>0.05, \* p <0.05, \*\* p<0.01, \*\*\* p<0.001, Student’s t-test.

**b)c)** MRC-5 cells treated with scrambled (Con), USP11, SETX or USP11 and SETX siRNA were incubated with the indicated doses of CPT or formaldehyde (FA) for 1.5 and 3 hours respectively, and left to grow for 7 days. The surviving colonies were counted and % survival calculated relative to scrambled-treated cells. Data are the mean ± SD from 3 biological repeats. \*p<0.05, Student’s t-test.

## USP11 might regulate SETX stability post-translationally

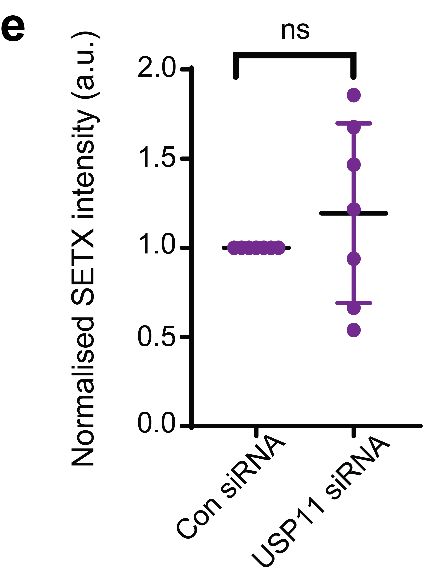
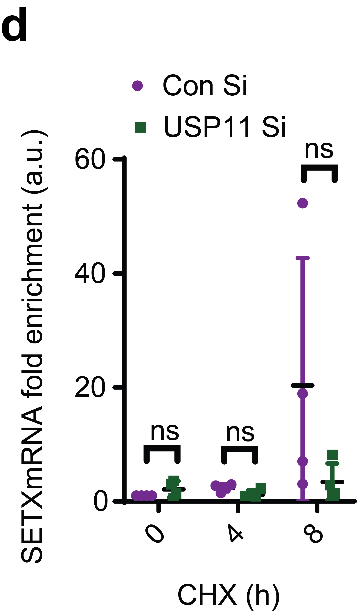
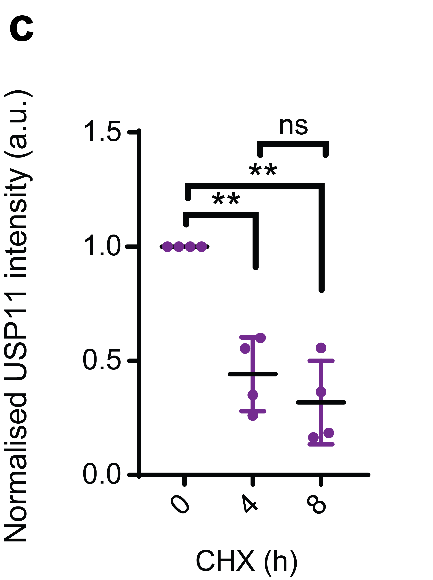
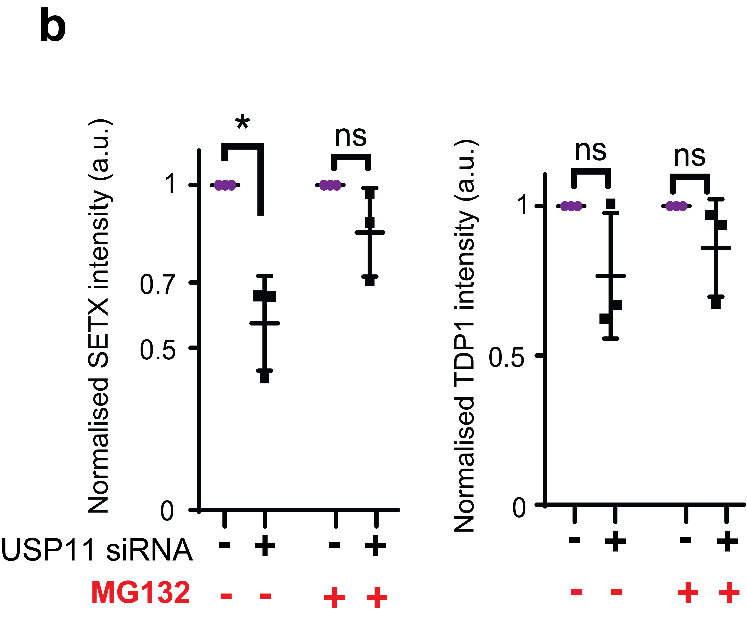
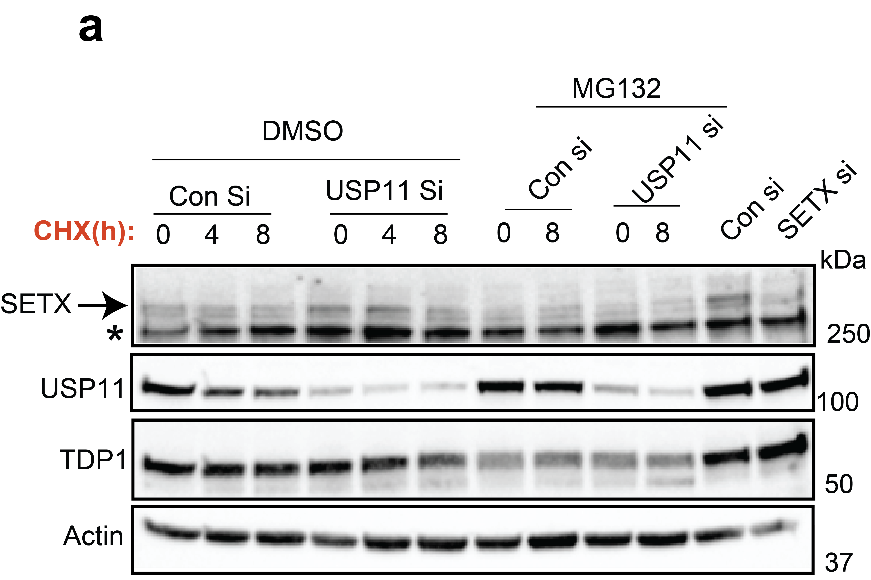
Bennett et al., (2013) previously described a band shift of the N-terminal SETX1-650, after an NEM treatment. NEM is a cysteine protease inhibitor, meaning it inhibits the majority of human DUBs including USP11. As the authors did not suggest a putative E3 ligase or DUB for senataxin, it was decided to investigate if USP11 is involved in SETX deubiquitination.

The most well described effect of ubiquitination is protein turnover. Therefore, in the absence of USP11, SETX might be highly ubiquitinated and eventually degraded. In order to test this hypothesis, a protein translation inhibitor cycloheximide (CHX) was employed. In presence of CHX, new proteins are not synthesised and the protein turnover rates can be measured.

USP11-depleted HEK-293T cells were subjected to CHX treatment for up to 8h in presence or absence of a proteasomal inhibitor MG132 (Fig. 4.4a). MG132 treatment served as a negative control for CHX activity, since it inhibits protein turnover. SETX siRNA served as a control for anti-SETX immunoblotting. Turnover rates of SETX, USP11, and TDP1 were measured (Fig. 4.4b). It was found that upon depletion of USP11, senataxin, but not TDP1, was significantly degraded after 8h of CHX treatment, which was counteracted by MG132.

The stability of USP11 was also measured, as an internal control for CHX activity (Fig. 4.4c). Together, the data acquired from the CHX chase experiment suggest that USP11 regulates the stability of SETX through a post-translational modification. Additionally, it was decided to investigate if USP11 controls SETX levels through altering its transcription rates. If USP11 promoted transcription of SETX mRNA, then SETX mRNA should be downregulated in USP11-depleted cells. Therefore, RT-qPCR approach was employed to measure SETX mRNA levels upon USP11-depletion in HEK-293T cells (Fig. 4.4d). The data suggest that 1) there was no significant misregulation of SETX mRNA upon USP11 depletion, implying that USP11 does not regulate transcription rates of SETX mRNA and that 2) the faster turnover rate of senataxin observed in Fig 4.4ab was not due to misregulated senataxin mRNA levels.

Finally, it was observed that transient siRNA depletion of USP11 without CHX treatment did not affect senataxin steady-state levels (Fig. 4.4e), suggesting that a prolonged depletion of USP11 would be required to observe reduced senataxin levels.



**Figure 4.4 USP11 regulates SETX at a post-translational level.**

**a)** HEK-293T cells were transfected with scrambled siRNA (Con), USP11 or SETX siRNA and incubated with 100 μg/ml cycloheximide (CHX) alone or additionally with 25 μM MG132 for 0, 4 and 8 hours. The expression level of SETX, USP11, TDP1 and actin was analysed by immunoblotting.

**b)** The band intensity of SETX (left panel) and TDP1 (right panel) following 8 h incubation with CHX alone or additionally with MG132 was quantified from 3 biological repeats, normalised to actin and then presented as fold reduction compared to un-treated samples ± SD.

**c)** The band intensity of USP11 following incubation with CHX from 4 biological repeats, normalised to actin and then presented as fold reduction compared to un-treated samples ± SD. \*\* p<0.01, ns; p>0.05, Student’s t-test.

**d)** Control (Con Si) or USP11 siRNA-transfected HEK-293T cells were treated with CHX for indicated times. Total RNA was extracted, reverse-transcribed to cDNA and used for qPCR. SETX mRNA levels were normalised to actin. Data are the mean ± SD from 4 biological repeats. ns; p>0.05, Student’s t-test.

**e)** Lysates from Control (Con siRNA) and USP11 siRNA-transfected HEK-293T cells were fractionated by SDS-PAGE and analysed by immunoblotting using SETX, USP11 and actin antibodies (data not shown). SETX band intensities were normalised to actin and presented as fold change compared control cells. Data are the average of 7 biological repeats and presented as mean ± SD. ns; p>0.05, Student’s t-test.

Taken together, our results suggest that USP11 might regulate degradation of senataxin through a post-translational modification, which is in line with *S.cerevisiae* data showing that the turnover of senataxin homologue, Sen1p, is conducted by the Ubiquitin-Protease System (Mischo et al., 2018).

## USP11 depletion leads to elevated ubiquitination of the N-terminal domain of SETX

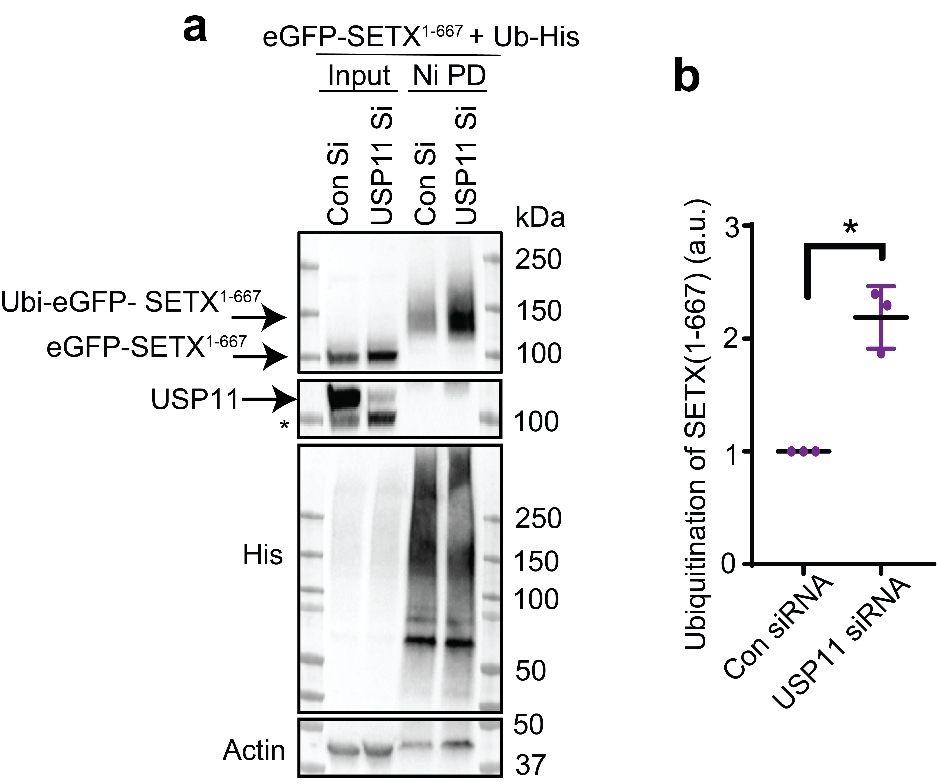
If USP11 deubiquitinates SETX in order to regulate its stability, then SETX should be hyper-ubiquitinated in the absence of USP11. Therefore, it was decided to design an experiment to purify ubiquitinated SETX. Ub-His construct and Nickel beads were employed. The premise of the experiment was to over-express Ub-His in HEK-293T cells and pull down SETX-Ub-His species using Nickel beads.

Initially, this experiment was optimised using endogenous or over-expressed full-length SETX, various concentrations of lysates, SDS-PAGE gels, and wash buffers. Unfortunately, it was found to be extremely difficult to trap ubiquitinated SETX, presumably, due to its large molecular weight of >302.8 kDa.

However, SETX1-650 was previously described to be modified by ubiquitin/SUMO pathway (Bennett et al., 2013) and SETX1-667 was shown to localize to nucleoli. As eGFP-SETX1-667 construct was in our possession, it was decided to use it instead of the full-length SETX. Fortunately, ubiquitinated eGFP-SETX1-667 was trapped by Nickel beads and could be used in the context of USP11-depletion.

Ub-His and eGFP-SETX1-667 constructs were over-expressed in USP11- or scrambled siRNA-treated HEK-293T cells (Fig. 4.5a). Nickel beads were used to pull down ubiquitinated species under denaturing conditions. The data suggest that eGFP-SETX1-667 species were significantly enriched in USP11-depleted cells as compared to the control (Fig. 4.5b), implying that USP11 specifically deubiquitinates eGFP-SETX1-667.

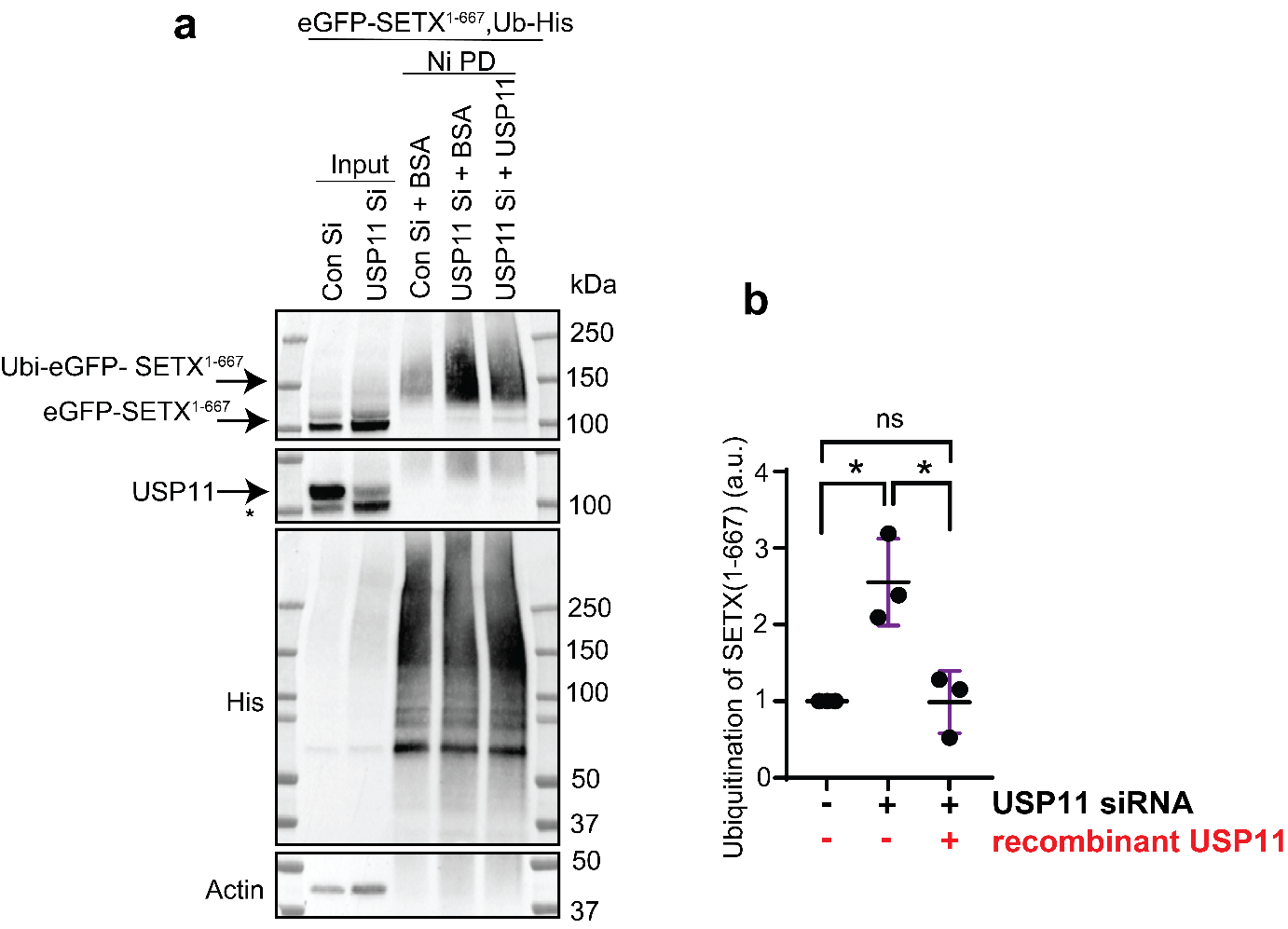
The above result was further tested by an *in vitro* deubiquitination reaction (Fig. 4.6a). Purified ubiquitinated eGFP-SETX1-667 was incubated with recombinant USP11. Deubiquitinating activity of USP11 triggered a release of eGFP-SETX1-667from the eGFP-SETX1-667 / ubiquitin-His-nickel beads complex. As a result, the amount of bead-bound His-Ub-eGFP-SETX1-667 was significantly reduced (Fig. 4.6b).



**Figure 4.5 USP11 depletion leads to enrichment of ubiquitinated eGFP-SETX1-667.**

**a)** Control (Con) and USP11 depleted HEK-293T cells (USP11si) were transfected with plasmids encoding eGFP-SETX1-667 and Ub-His, and lysates were subjected to nickel pull-down under denaturing conditions to purify ubiquitinated proteins. Samples were fractionated by SDS-PAGE and analysed by immunoblotting using anti-GFP, USP11, His and actin antibodies.

**b)** The band intensities of Ubi-eGFP-SETX1-667 were normalised to His-Ub and presented as fold increase of SETX ubiquitination in USP11 knockdown cells compared to controls (Con). Data are the average ± SD from 3 biological repeats. \*p<0.05, Student’s t-test.



**Figure 4.6 Recombinant USP11 deconjugates ubiquitin from ubiquitinated eGFP-SETX1-667**

**a)** Purification of Ubi-eGFP-SETX1-667 was conducted as in Figure 4.5 and followed by incubation of the nickel beads with 2.88µg/ml BSA or recombinant USP11 in a deubiquitination buffer for 4 hours. During the *in vitro* deubiquitination reaction, recombinant USP11 liberated SETX1-667 from the SETX1-667 / ubiquitin-His-nickel beads, resulting in a reduction of the remaining bead-bound ubiquitinated SETX1-667. Samples were fractionated by SDS-PAGE and analysed by immunoblotting using anti-GFP, USP11, His and actin antibodies.

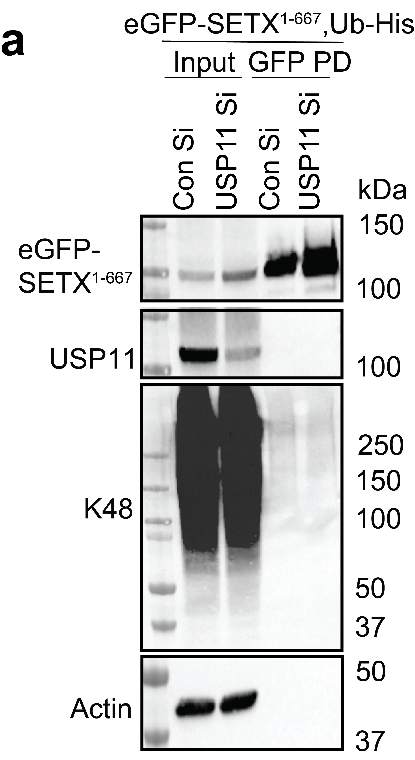
**b)** The band intensities of Ubi-eGFP-SETX1-667 were normalised to His-Ub and presented as fold increase of SETX ubiquitination compared to controls. Data are the average ± SD from 3 biological repeats. \*p<0.05, ns; p>0.05, Student’s t-test.

K11 and K48 ubiquitin chains are post-translational modifications that act as markers for protein degradation (Chau et al., 1989; Jin et al., 2008; Meyer and Rape, 2014). The Mao lab suggested that USP11 cleaves K48 ubiquitin chains off p21 protein (Deng et al., 2018). Therefore, it was hypothesised that USP11 depletion might trigger enrichment of K48-ubiquitinated eGFP-SETX1-667 species. eGFP-SETX1-667 and Ub-His constructs were over-expressed again in USP11-depleted HEK-293T cells. eGFP-SETX1-667 was pulled-down with eGFP-beads under denaturing conditions and the K48 signal was measured using anti-K48 antibody (Fig. 4.7a). The acquired data does not support the notion that USP11 deconjugates K48-linked ubiquitin chains from SETX1-667, as USP11 depletion did not trigger enrichment of K48-ubiquitinated species of SETX1-667. Therefore, it was concluded that a different ubiquitin chain, perhaps K11, is involved in senataxin proteostasis.

Based on the pull-down assays (Fig. 4.5 and 4.6), it was hypothesised that SETX and USP11 might interact together. In order to test that, protein-protein interaction domain of SETX (eGFP-SETX1-667) was over-expressed in HEK-293T cells and GFP-beads were used to pull-down eGFP-SETX1-667 under native conditions (Fig. 4.8a). The immunoblotting analysis suggested that USP11 interacts with SETX1-667 in our experimental settings.

To scrutinise this result, it was decided to assay if SETX1-667 interacts with its known partner RNA Polymerase II in our experimental conditions (Suraweera et al., 2007). As expected, RNA Polymerase II was pulled-down, which demonstrated the specificity of our experimental design (Fig. 4.8a).

Next, the connection between SETX and RNA Polymerase I was investigated. RNA Polymerase I-dependent transcription occurs in nucleoli and SETX was linked to these compartments by this study and others (this study, Yüce and West, 2013; Suraweera et al., 2007). Additionally, Senp1, *S.cerevisiae* SETX homolog, was previously shown to interact with RNA Polymerase I (Kawauchi et al., 2008), suggesting that human SETX might co-immunoprecipitate with RNA Polymerase I too. Therefore, the largest subunit of RNA Polymerase I, RPA194, was probed for and found to be in a complex with SETX1-667 in HEK-293T cells (Fig. 4.8a). Interestingly, this is the first evidence of SETX interacting with RNA Polymerase I to my knowledge, which opens up a new avenue for research studying the involvement of SETX in nucleolar transcription.



**Figure 4.7 USP11 does not remove K48-linked ubiquitin chains from eGFP-SETX1-667.**

**a)** Control (Con Si) and USP11 depleted HEK-293T cells (USP11 Si) were transfected with plasmids encoding eGFP-SETX1-667 and Ub-His, and lysates were subjected to GFP pull-down under denaturing conditions to purify eGFP-SETX1-667. Samples were fractionated by SDS-PAGE and analysed by immunoblotting using anti-GFP, USP11, K48 and actin antibodies.



**Figure 4.8 USP11 and eGFP-SETX1-667 co-immunoprecipitate.**

**a)** HEK-293T cell lysates expressing an empty vector (eGFP-EV) or a vector encoding eGFP-SETX1-667 were subjected to GFP pull-down and analysed by immunoblotting using antibodies against GFP, USP11, RNA polymerase II (RNAP II) and RNA polymerase I (RPA 194).

To summarise data from this chapter:

* USP11 and SETX were shown to be epistatic. siRNA knockdowns of USP11 & SETX triggered 2-fold R-loop enrichment, CPT & FA hypersensitivity, which were not further exacerbated by the double knock-down.
* USP11 might control SETX turnover through post-translational modifications, rather than through altering SETX transcription rates.
* USP11 was shown to deubiquitinate eGFP-SETX1-667. Depletion of USP11 triggered an enrichment of ubiquitinated eGFP-SETX1-667 species.
* USP11 was shown to co-immunoprecipitate with eGFP-SETX1-667.
* eGFP-SETX1-667 was shown to interact with RNA Polymerase II and with the largest subunit of RNA Polymerase I – RPA 194.
* USP11 depletion did not trigger enrichment of K48-ubiquitinated species of eGFP-SETX1-667, suggesting that a different ubiquitin chain might be involved in senataxin degradation.

## Discussion

## USP11 and SETX are found in nucleoli

The Harper lab carried out a global proteomic screen for 75 human DUBs, including USP11 (Sowa et al., 2009). Their data suggested that USP11 interacted with 45 substrates including RAE1 and KEAP1. USP11-RAE1 and USP11-KEAP1 interactions were later demonstrated by Stockum et al., (2018) and Orthwein et al., (2015) respectively. Sowa’s dataset of USP11 protein-protein interactions did not contain BRCA2 (Schoenfeld et al., 2004) or SETX, suggesting that the dataset was not exhaustive and could miss some >300 kDa molecular weight proteins.

Sowa et al., also suggested that USP11 interacts with POLR1A (also known as RPA 194), POLR1B (RPA 135), POLR1C (RPA 39) and POLR1E, which are all subunits of RNA Polymerase I. This is in line with confocal, X-ChIP and DRIP data from this study. Here, it was shown that USP11 immunoprecipitates with ribosomal DNA and its depletion triggers R-loop accumulation in two ribosomal DNA loci. Therefore, USP11 association with RNA Polymerase I, perhaps via senataxin, seems plausible.

It is worth to mention that eGFP-SETX1-667 was used here for immunostaining rather than the full-length SETX. It is possible that eGFP-SETX1-667 co-localization to nucleoli could have been an artefact of using a truncated version of the protein. However, full-length SETX was shown to be present in a complex with nucleolin (Suraweera et al., 2009), the key component of human nucleoli, suggesting that the SETX1-667 co-localization was not an artefact. Therefore, USP11 and SETX may not only interact in the nucleoplasm, but also in the nucleoli.

## USP11 controls stability of SETX through post-translational modifications

Two bands of SETX are visible in the cycloheximide experiment above 250kDa ladder band (Fig. 4.4a). A possibility of non-specific binding of the SETX antibody was disregarded because the bands were sensitive to SETX siRNA and thus likely represented two distinct populations of SETX. Both bands were sensitive to combined USP11-depletion & CHX treatment, which corroborated our hypothesis. As SETX does not have experimentally confirmed isoforms, it is likely that the two bands represent post-translationally modified and unmodified species of SETX. The post-translational modifications of SETX could include ubiquitination/SUMOylation (Bennett et al., 2013) or phosphorylation (Zhou et al., 2013).

Pull-down data, showing that upon USP11 depletion ubiquitinated eGFP-SETX1-667 species were enriched, was provided. It was concluded that it was the SETX1-667 fragment that was ubiquitinated rather than eGFP, as GFP was previously shown not to be ubiquitinated by others (Brown et al., 2015). The Marynen lab demonstrated that eGFP can hinder the ubiquitination process of eGFP fusion proteins (Baens et al., 2006), suggesting that eGFP-SETX1-667 fragment could potentially be more ubiquitinated than our data suggest.

Unfortunately, evidence of specific ubiquitin-chain modification of SETX1-667 was not provided. Our data suggest that K48 chains should be excluded from the future search. As a result, it is hypothesised that SETX1-667 could be degraded via K11-linked ubiquitin chains instead (Jin et al., 2008).

## Insights from USP11 and SETX clonogenic cell survival assays

SETX-depleted cells were sensitive to CPT treatment in our experimental conditions. This is in line with the research from The Lavin lab where they showed that senataxin-defective AOA2 patient cells were sensitive to CPT (Suraweera et al., 2009). USP11-depleted cells were not previously shown to be CPT sensitive. However, there is a strong evidence supporting their sensitivity to olaparib (Wiltshire et al., 2010; Orthwein et al., 2015). Olaparib is a PARP inhibitor that triggers the accumulation of single-strand breaks (Wiltshire et al., 2010). CPT is a topoisomerase poison that triggers protein-linked single-stranded DNA breaks (Desai et al., 2003). Therefore, CPT and olaparib treatments both induce single-stranded breaks. The accumulation of SSBs leads to stalling of the RNA polymerases and formation of DSBs, which were both shown to promote R-loop formation in actively transcribed genes (Marinello et al., 2013; Cohen et al., 2018). In the absence of USP11 or SETX, these lesions are not correctly repaired and result in cell death.

It was also demonstrated that USP11 and SETX-depleted cells are hypersensitive to formaldehyde (FA). FA was shown to trigger R-loop formation through depletion of BRCA2 (Tan et al., 2017). Since USP11- and SETX- depleted cells are more sensitive to FA than the control cells, these proteins might act separately from BRCA2-dependent R-loop resolution pathway. Even though USP11 and BRCA2 were shown to interact (Schoenfeld et al., 2004), data presented here suggest that this interaction might not be a part of USP11/SETX R-loop resolution mechanism, but be a part of e.g. homologous recombination (Orthwein et al., 2015). Our interpretation is in line with Hatchi et al. (2015) where they suggested that the SETX/BRCA1 R-loop resolution pathway might be independent of BRCA2.

Finally, as the experiments describing USP11/SETX epistasis were based on RNA intereference, there was a possibility of artefact results. Therefore, it was decided to corroborate the siRNA data through a CRISPR-Cas9 system described in the next chapter.

# Chapter 5: USP11 knockout CRISPR cell lines display R-loop accumulation phenotype and increased senataxin turnover



## Introduction

## CRISPR-Cas9

CRISPR-Cas9 is an adaptive immune system that evolved to protect prokaryotic cells from viral and plasmid infections. Cas9 is an endonuclease, while CRISPR is a DNA sequence. CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats (Jansen et al., 2002). There are three types of CRISPR-Cas9 systems (I, II, III). Type II has been adapted for use in eukaryotic cells (Ran et al., 2013) and is described below.

CRISPR contains repetitive regions interspaced by variable stretches of DNA called spacers (Moijica et al., 2005). The spacers are derived from viral or plasmid sequences, from previous infections, that code for precursor CRISPR RNA (pre-crRNA) (Garneau et al., 2010).

The repetitive regions code for tracrRNA, which stands for trans-activating crRNA. Both pre-crRNA and tracrRNA are non-coding RNAs. tracrRNA together with ribonuclease RNase III and Cas9 guide maturation of pre-crRNA into active crRNA (Deltcheva et al., 2011).

crRNA and tracrRNA hybridise through Watson-Crick base-pairing to form a guide RNA (gRNA). gRNA complexes with Cas9; and leads Cas9 to foreign viral or plasmid DNA (Deltcheva et al., 2011). This is possible due to the spacer sequence present in the gRNA. The spacer is complementary to viral/plasmid sequences, which are called protospacers (Moijica et al., 2005). Upon successful base-pairing of the spacer (gRNA) to the protospacer (foreign DNA), Cas9 recognises ‘protospacer-adjacent motif’ (PAM), and subsequently, cleaves both strands of the foreign DNA (Garneau et al., 2010; Jinek et al., 2012, Nishimasu et al., 2014). Cleavage of the foreign DNA renders it silent and protects the prokaryote from the invasion.

Jinek et al., demonstrated that synthetically generated single guide RNA (sgRNA; consisting of fused tracrRNA and crRNA) is able to recruit Cas9 to a target DNA; and suggested that the sgRNA could be modified to match any target DNA (Jinek et al., 2012).

This led to two studies where type II CRISPR system was adapted for use in mammalian cells to generate targeted DSBs (Cong et al., 2013, Mali et al., 2013). Upon DSB formation in human cells, non-homologous end-joining (NHEJ) or homologous recombination (HR) pathways are employed to repair the break.

HR is an error-free pathway that requires a template DNA to facilitate the repair. NHEJ does not require the template and might randomly induce insertion or deletion mutations (indels) in order to repair the DSB (Bibikova et al., 2002, Bibikova et al., 2003).

Therefore, a targeted DSB in a gene of interest X can result in the disruption (via NHEJ-associated indels) or alteration (via an HR-associated template) of the gene X.

The alteration of X is accomplished by providing a DNA template containing a desired insert Y, leading to a formation of X-Y gene fusion. To date, the CRISPR-Cas9 was proven to be a successful gene editing technology as demonstrated by numerous studies (for instance: Batista et al., 2014; Orthwein et al., 2015; Mandegar et al., 2016; Wang et al., 2017; and Zhang et al., 2017).

## Aims and objectives

The aim of this chapter was to corroborate the link between USP11, SETX, and R-loops using USP11 knockout system. The objectives were to:

* Generate USP11 knockout cell line using the CRISPR-Cas9 system.
* Assay R-loop phenotype and SETX stability in USP11 knockout cells.
* Complement USP11 knockout cells with wild type and active site mutant versions of USP11.
* Assay R-loop phenotype and SETX stability in the complemented USP11 knockout cells.

## Results

## Generation of USP11 knockout and USP11-complemented cell lines

My first try of generating USP11 knockout cell line employed MRC-5 cell line and the first generation CRISPR plasmid pSpCas9(BB)-2A-Puro, which expressed USP11 sgRNA and Cas9 protein (Ran et al., 2013). Unfortunately, all the transfected cells died after applying puromycin selection for weeks. The protocol was not optimised to yield a viable USP11 knockout cell line.

The next attempt to generate USP11 knockout cell line involved HEK-293T cell line (Orthwein et al., 2015) and the second generation CRISPR plasmid pSpCas9 (BB)-2A-GFP, which co-expressed USP11 sgRNA, Cas9, and GFP (Ran et al., 2013). As transfected cells were tracked by transient GFP expression, an antibiotic selection was not required, which alleviated the stress applied to HEK-293T cells. Eventually, transfected clones were screened by western blotting to yield two USP11 knockout clones named USP11 sgRNA Cl-1 and USP11 sgRNA Cl-2. Immediately, the USP11 sgRNA Cl-1 & 2 were stably complemented with either wild-type (WT) USP11- or active site mutant (C318S) USP11-expressing vectors, which generated additional four cell lines (Fig. 5.1a).

The indel mutations of USP11 sgRNA Cl-1 & 2 were investigated by sequencing (Fig. 5.1b). USP11 sgRNA Cl-1 was found to possess a 23nt deletion in exon 1, which introduced an UGA stop codon 45 nucleotides downstream of the deletion. USP11 sgRNA Cl-2 possessed a 2nt deletion in exon 1, which introduced an UGA stop codon 63 nucleotides downstream of the mutation.

The acquired data suggest that CRISPR-Cas9 system triggered deletion mutations and formation of premature stop codons in exon 1 of both USP11 sgRNA clones. The premature stop codons might trigger degradation of USP11 mRNAs by a non-sense mediated mRNA decay (NMD, reviewed by Baker and Parker, 2004). Alternatively, if the short mRNA transcripts are translated, the synthesised proteins would miss ~95% of amino acids of the full length USP11.

C:\Users\mbb\Desktop\sequences\clone 8 exon 1, 23nt deletion.tifC:\Users\mbb\Desktop\sequences\clone 9 exon 1, 2nt deletion.tif

**USP11 sgRNA Cl-1**

**USP11 sgRNA Cl-2**

**Figure 5.1 USP11 CRISPR knockout clones.**

**a)** HEK-293T cells were transfected with a vector expressing Cas9 and sgRNA targeting exon 1 of USP11. Single clones were obtained following a serial dilution method and USP11 knockout in two clones was demonstrated by immunoblotting using anti-USP11 and anti-tubulin antibodies. USP11 knockout clones were then stably complemented by vectors encoding full-length USP11 (WT) or catalytically inactive USP11C318S mutant (C318S).

**b)** Deletion mutations in USP11 sgRNA Cl-1 and Cl-2 were demonstrated by sequencing.

## Characterization of early passages of CRISPR cell lines

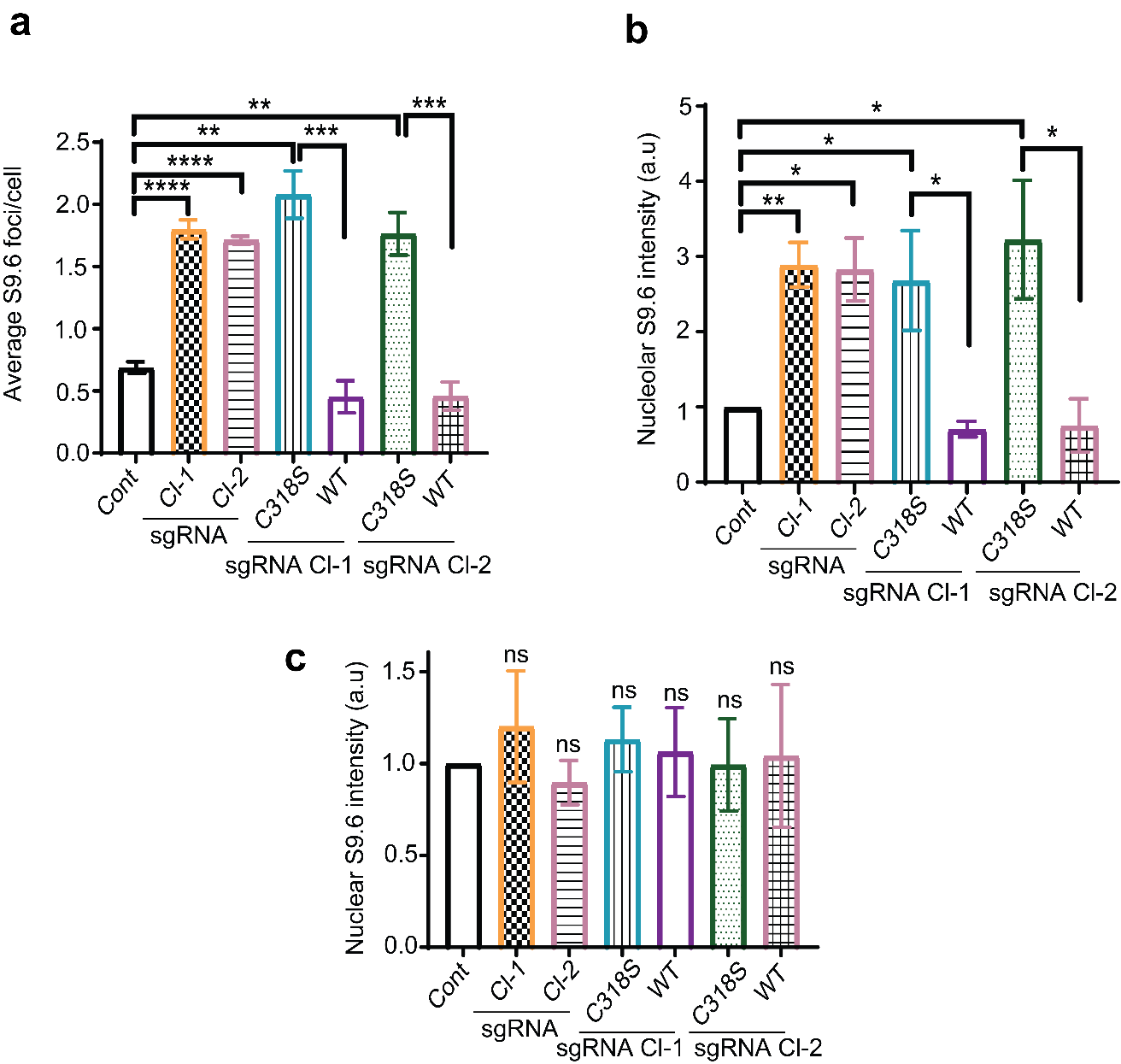
## USP11 catalytic activity is required for normal R-loop homeostasis

All USP11 sgRNA clones were examined by S9.6 immunofluorescence. Parental HEK-293T cells were used as a control. USP11 sgRNA Cl-1 and 2 displayed significant ~3-fold increase in nucleolar R-loop levels, which was ablated by over-expression of USP11WT, but not USP11C318S (Fig 5.2ab). As expected, significant changes in nuclear R-loop levels were not detected by immunofluorescence (Fig 5.2c).

Next, HEK-293T, USP11 sgRNA Cl-2, WT- and USP11C318S-complemented USP11 sgRNA Cl-2 cell lines were subjected to DRIP-qPCR (Fig 5.3). R-loop accumulation was examined at published nucleolar (28S and R7) and nuclear (Actin and ING3) loci (Karahan et al., 2015; Shen et al., 2017; Groh et al., 2014; Halász et al., 2017). Representative % input data are shown in Fig. 5.4. The S9.6 signal was validated by on-beadec-RNase H treatment. Taken together, USP11 sgRNA Cl-2 displays a significant R-loop accumulation phenotype at nucleolar and nuclear loci. Moreover, the phenotype was ablated by complementation with USP11WT, but not USP11C318S.

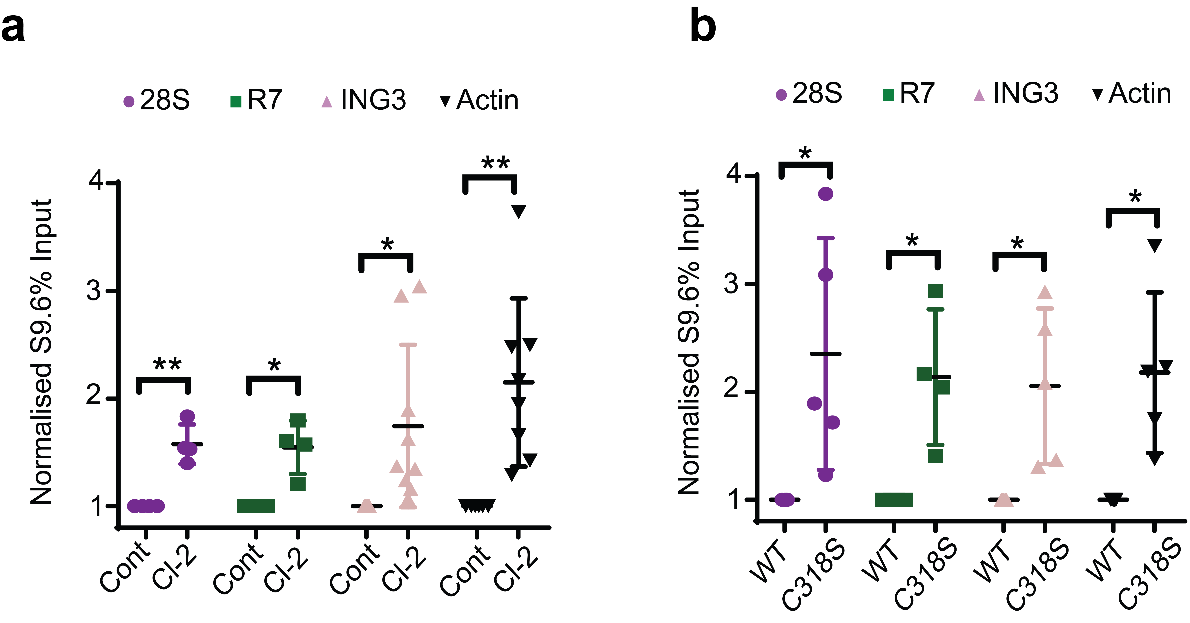
The R-loop accumulation phenotype from the complemented sgRNA Cl-2 cell lines was further assayed by a slot blot analysis (also known as a dot blot) (Stork et al., 2016). DNA extracts from all four complemented sgRNA cell lines were fixed onto a nitrocellulose membrane (Fig. 5.5a). Single-stranded DNA (ssDNA) served as a loading control. The slot blot data implies that the USP11C318S clones accumulate R-loops as compared to USP11WT clones (Fig. 5.5b).

All together immunofluorescence, slot blot, and DRIP-qPCR data (Fig. 5.2-5) suggest that USP11 knockout clones complemented with USP11WT do not accumulate R-loops, while USP11C318S clones do. Therefore, it was concluded that the R-loop accumulation phenotype observed in our experimental settings is dependent on the catalytic activity of USP11.



**Figure 5.2 USP11 sgRNA clones display R-loop accumulation phenotype according to immunofluorescence data.**

**a)b)c)** USP11 sgRNA clones and those complemented with USP11WT or the USP11C318S were examined by S9.6 immunofluorescence. The average number of S9.6 foci/cell and total nucleolar & nuclear fluorescence normalised to control cells (Cont) were calculated from 3 biological repeats, each containing at least 100 cells and presented as average ± SD. ns; p > 0.05, \*p<0.05, \*\* p<0.01, \*\*\* p<0.001, \*\*\*\* p<0.0001, Student’s t-test.



**Figure 5.3 USP11 knockout clones display R-loop accumulation phenotype according to DRIP-qPCR data**

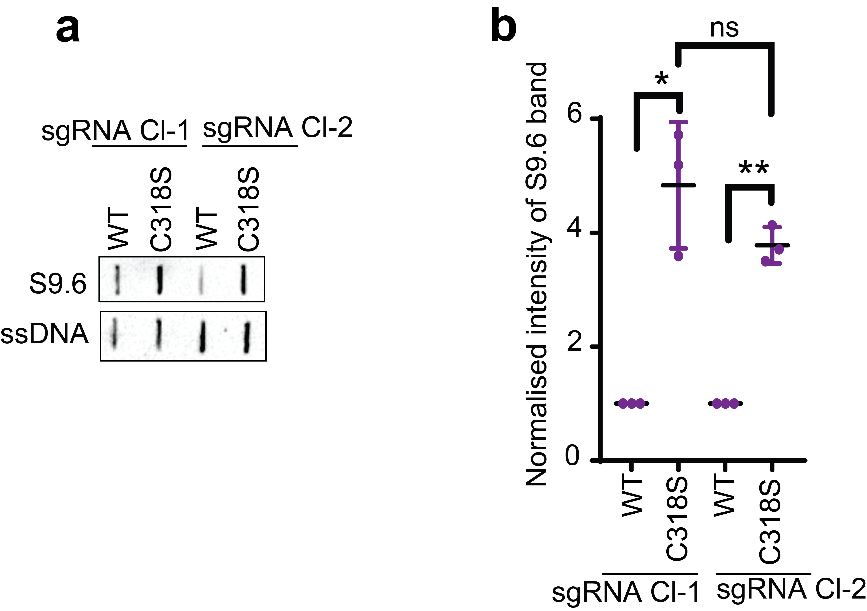
**a)b)** Lysates from Control (Cont; HEK-293T), USP11 sgRNA Cl-2 (Cl-2) & USP11 sgRNAs Cl-2 complemented with USP11WT (WT) or catalytically inactive USP11C318S mutant (C318S) cells were subjected to DNA/RNA immunoprecipitation using S9.6 antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3 and actin) loci. The raw % input values were normalised to control or USP11WT cells and presented as average ± SD from at least 3 biological repeats. \*p<0.05, \*\* p<0.01, Student’s t-test.



**Figure 5.4 USP11 knockout clones display R-loop accumulation phenotype – raw DRIP-qPCR data.**

**a) b)** Raw data from a single biological repeat of DRIP-qPCR for Figure 5.3a. Lysates from USP11 knockout Cl-2 (sgRNA2) and control HEK-293T (HEK) cells were subjected to DNA/RNA immunoprecipitation (DRIP) using S9.6 antibodies. Quantitative PCR was conducted as described above and *in vitro*, on-bead ec-RNase H (ec-RH) treatment served as a signal validation control.

**c) d)** Raw data from a single biological repeat of DRIP-qPCR for Figure 5.3b. Lysates from USP11 knockout cells complemented with either USP11WT (WT) or catalytically inactive USP11C318S (C318S) were subjected to DNA/RNA immunoprecipitation (DRIP) using S9.6 antibodies. Quantitative PCR was conducted as described above and *in vitro*, on-bead ec-RNase H (ec-RH) treatment served as a signal validation control.



**Figure 5.5 USP11C318S knockout clones display R-loop accumulation phenotype.**

**a)** Lysates from USP11 knockout cells complemented with either USP11WT (WT) or catalytically inactive USP11C318S (C318S) were subjected to a slot blot analysis using S9.6 and α-ssDNA antibodies. A representative experiment is shown. ssDNA signal was used as a loading control.

**b)** S9.6 band intensities from (a) were normalised to ssDNA and presented as fold enrichment compared to USP11WT cells. Data are the average of 3 biological repeats and presented as mean ± SD. ns; p>0.05, \* p<0.05, \*\* p<0.01, Student’s t-test.

## USP11 knockout clones display SETX knockdown

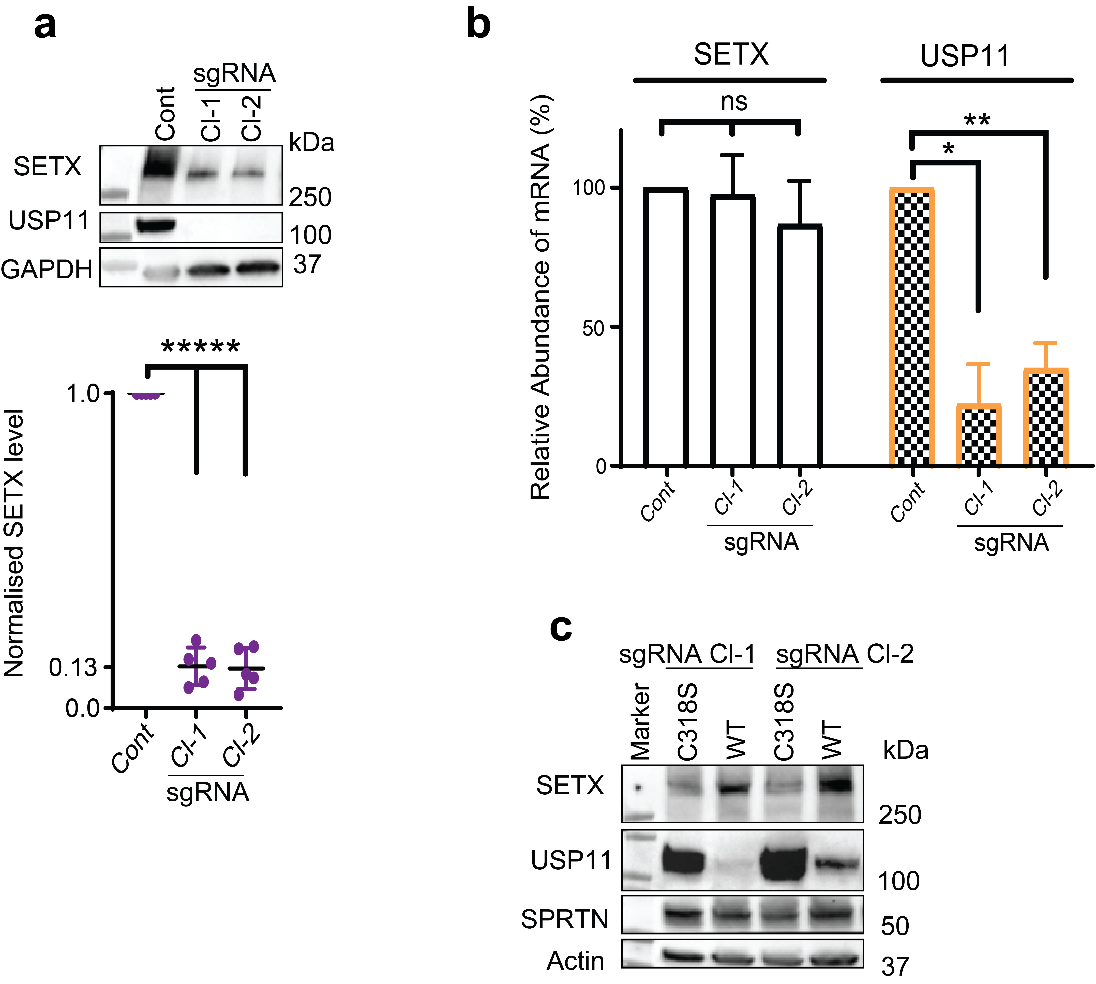
As USP11-depleted cells exhibit increased SETX ubiquitination & turnover rates (Chapter 4), it was hypothesised that USP11 knockout cells might have reduced SETX steady-state levels. Prolonged knockout of USP11 could result in an excessive ubiquitination of SETX and its subsequent degradation over time. This scenario is different from USP11 depletion by siRNAs, as the siRNA treatment triggers only a temporary effect.

Therefore, it was decided to investigate the protein level of SETX in USP11 knockout cells, using immunoblotting (Fig. 5.6a, *top panel*). GAPDH was used as a loading control. The data implies that USP11 sgRNA Cl-1 & 2 display a significant SETX knockdown (Fig. 5.6a, *bottom panel*). This is in line with our CHX and Nickel pull-down data (Chapter 4); and corroborates the hypothesis of USP11 regulating the stability of SETX.

In order to test if this observation is due to changes at the transcriptional level, SETX mRNA was quantified by RT-qPCR (Fig. 5.6b). USP11 mRNA served as a positive control. As expected, the SETX mRNA was not misregulated whereas the USP11 mRNA was significantly downregulated in USP11 knockout cells. Therefore, it was concluded that SETX is not regulated by USP11 at the transcriptional level in our experimental settings. This is in line with siRNA data described in Chapter 4.

Next, SETX protein levels were examined in USP11C318S and USP11WT sgRNA clones. As USP11C318S clones accumulate R-loops, it was hypothesised that these clones might have reduced steady-state levels of SETX (Fig. 5.6c). As expected, acquired data suggest that highly expressed USP11C318S did not restore SETX protein levels to the same extent as the USP11WT did (Fig. 5.6c). Therefore, the catalytic activity of USP11 might be essential to inhibit proteasomal degradation of senataxin.

Following this result, USP11WT and USP11C318S sgRNA Cl-2 cells were transfected with eGFP-SETX1-667 and Ub-His constructs in order to pull down ubiquitinated eGFP-SETX1-667 species using nickel beads (Fig 5.7a, *left panel*). Ubiquitinated eGFP-SETX1-667 species were found to be enriched two-fold in USP11C318S as compared to USP11WT (Fig 5.7a, *right panel*), suggesting that catalytic activity of USP11 is required to deubiquitinate the N-terminal domain of senataxin.

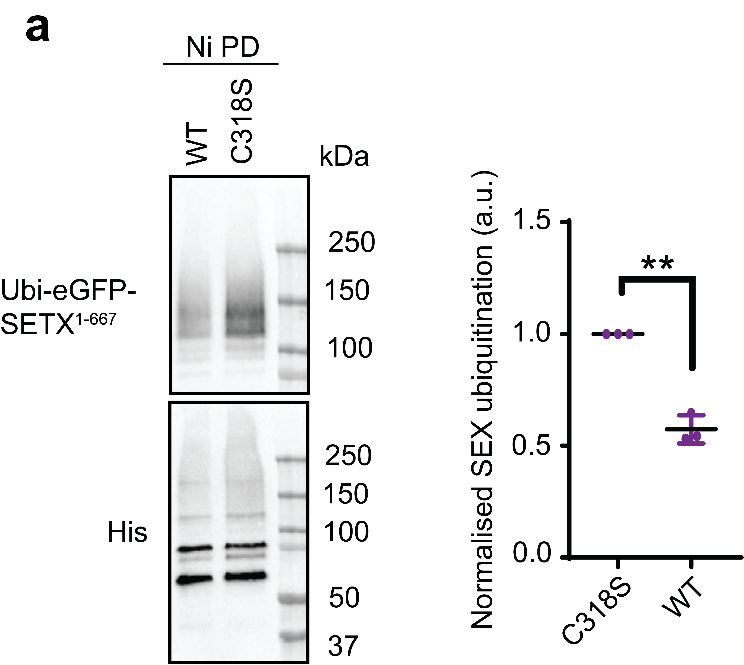


**Figure 5.6 The catalytic activity of USP11 is required to maintain SETX proteostasis**

**a)** Lysates from Control (Cont) and USP11 knockout HEK-293T cells (sgRNA Cl-1 and Cl-2) were fractionated by SDS-PAGE and analysed by immunoblotting using SETX, USP11 and GAPDH antibodies (*top*). SETX band intensities were normalised to GAPDH and presented as fold reduction compared to levels in control parental cells (*bottom*). Data are the average of 5 biological repeats and presented as mean ± SD. \*\*\*\*\* p<0.0001, Student’s t-test.

**b)** Total RNA was extracted from USP11 knockout clones at passage 2 – 7, reverse transcribed to cDNA, and USP11 and SETX transcripts quantified by qPCR. Data are the average of 3 biological repeats, normalised to actin transcript levels and presented as % reduction compared to the control parental cells**.** ns; p>0.05, \* p<0.05, \*\* p<0.01, Student’s t-test.

**c)** Lysates from USP11 knockout cells complemented with USP11WT (WT) or catalytically inactive USP11C318S mutant (C318S) were fractionated by SDS-PAGE and analysed by immunoblotting. SPRTN was used as a negative control.



**Figure 5.7 USP11C318S mutants accumulate ubiquitinated eGFP-SETX1-667.**

**a)** USP11 knockout Cl-2 cells expressing USP11WT (WT) or the catalytically inactive USP11C318S mutant (C318S) were transfected with plasmids encoding eGFP-SETX1-667 and Ub-His. Ubiquitinated eGFP-SETX1-667 was purified using Nickel beads under denaturing conditions as described in Figure 4.5a. The band intensities of Ubi-eGFP-SETX1-667 were normalised to His-Ub and presented as fold increase of SETX ubiquitination in USP11C318S sgRNA Cl-2 as compared to USP11WT sgRNA Cl-2 (right). Data are the average ± SD from 3 biological repeats. \*\* p<0.01, Student’s t-test.

## Characterization of late passages of USP11 knockout cells

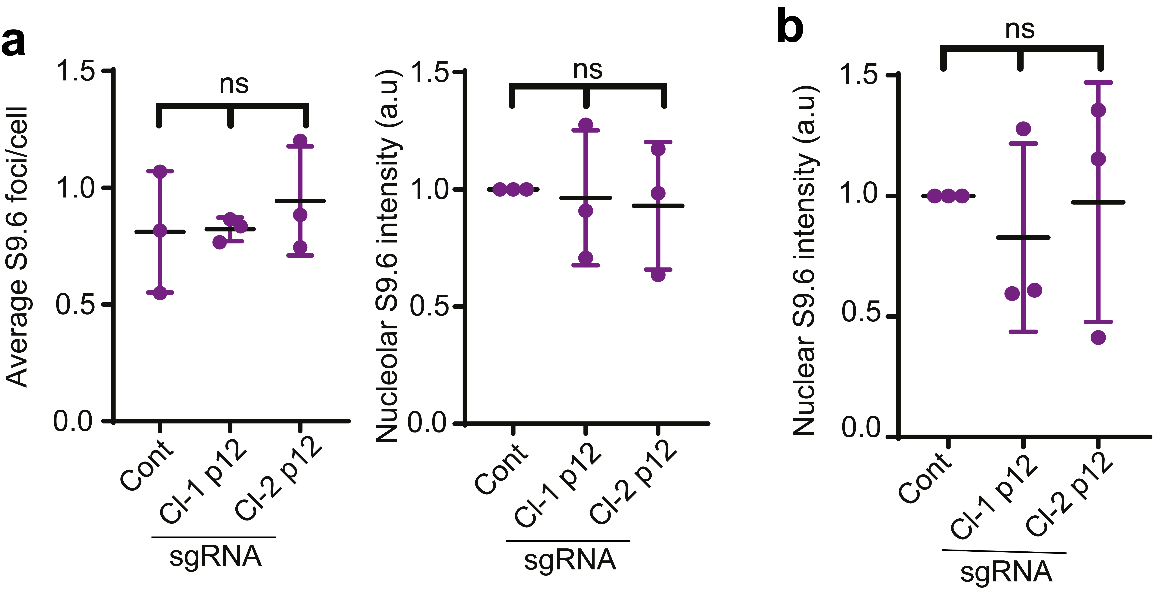
USP11 sgRNA Cl-1 & 2 were passaged and assayed for R-loops for two months (twelve passages) when it was realised that the R-loop accumulation phenotype was gone. Fig. 5.8a demonstrates S9.6 immunofluorescence data, which implies that aged USP11 sgRNA clones, passaged at least 12 times, do not accumulate nucleolar R-loops anymore. Significant changes in nuclear R-loops were not recorded either (Fig. 5.8b).

S9.6 phenotype of aged USP11 sgRNA Cl-2 cells was also investigated by DRIP-qPCR (Fig. 5.9a). R-loops were assessed at nucleolar (28S and R7) and nuclear (ING3 and actin) loci. Representative raw % input data are presented in Fig. 5.9b. The acquired data implies that aged USP11 sgRNA Cl-2 cells do not accumulate R-loops at these loci; and do not differ significantly from parental HEK-293T cell line.

As R-loop phenotype is directly linked to SETX in our experimental conditions, SETX steady-state levels were assayed in aged USP11 sgRNA Cl-1 & 2 cells (Fig. 5.10, *left panel*). SPRTN was used as a negative control. The acquired data suggest that aged USP11 sgRNA clones over-express SETX (Fig. 5.10, *right panel*), without restoring USP11 levels.

It was hypothesised that aged USP11 sgRNA clones reverted to normal R-loop phenotype by restoring SETX stability via transcriptional or post-translational modifications. In order to discern between the two, SETX mRNA levels were assayed in both non-aged and aged USP11 knockout cell lines (Fig. 5.11a). USP11 mRNA served as a negative control (Fig. 5.11b).

Interestingly, SETX mRNA in USP11 knockout cells did not differ significantly from the parental HEK-293T cells regardless of the age of USP11 knockout cells (Fig. 5.11a), implying that former downregulation and latter upregulation of SETX protein was due to post-translational modifications. Therefore, it was concluded that aged USP11 knockout cells reverted the R-loop accumulation phenotype to wild type via stabilising SETX steady-state levels through, presumably, a post-translational modification.



**Figure 5.8 Aged USP11 sgRNA Clones revert to wild-type R-loop homeostasis.**

**a)** Parental HEK-293T cells (Cont) and USP11 knockout clones 1 and 2 (Cl-1 and Cl-2) at passages 12-15 (p12) were examined for R-loop levels using S9.6 immunostaining. The average number of S9.6 foci/cell (*left panel*) and total nucleolar fluorescence normalised to control cells (*right panel*) were calculated from 3 biological repeats, each containing at least 100 cells and presented as average ± SD. ns; p>0.05, Student’s t-test.

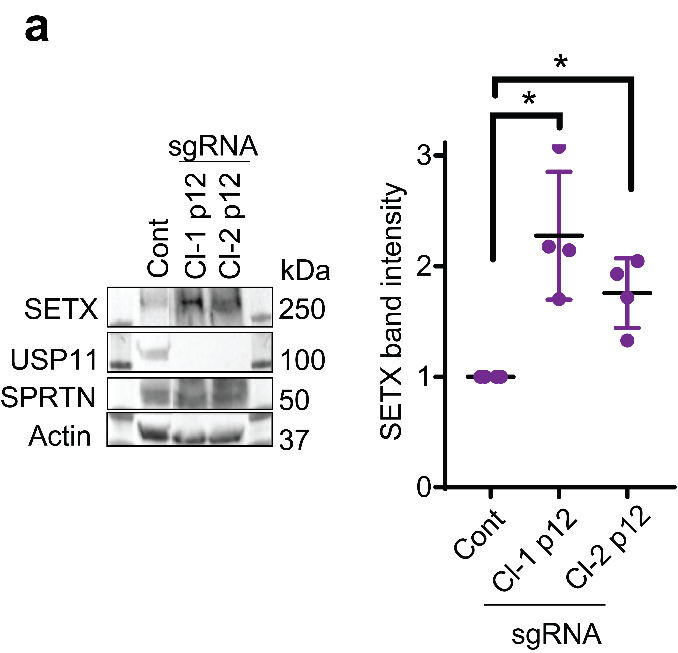
**b)** Corrected total nuclear fluorescence of aged (p-12-15) USP11 sgRNA Cl-1 and 2 (Cl-1 p12, Cl-2 p12) was measured using ImageJ and normalised to mock (Cont; HEK-293T). Data are the average ± SD from 3 biological repeats, each containing at least 100 cells. ns; p > 0.05, Student t-test.



**Figure 5.9 Aged USP11 sgRNA Clones do not accumulate R-loops according to DRIP data.**

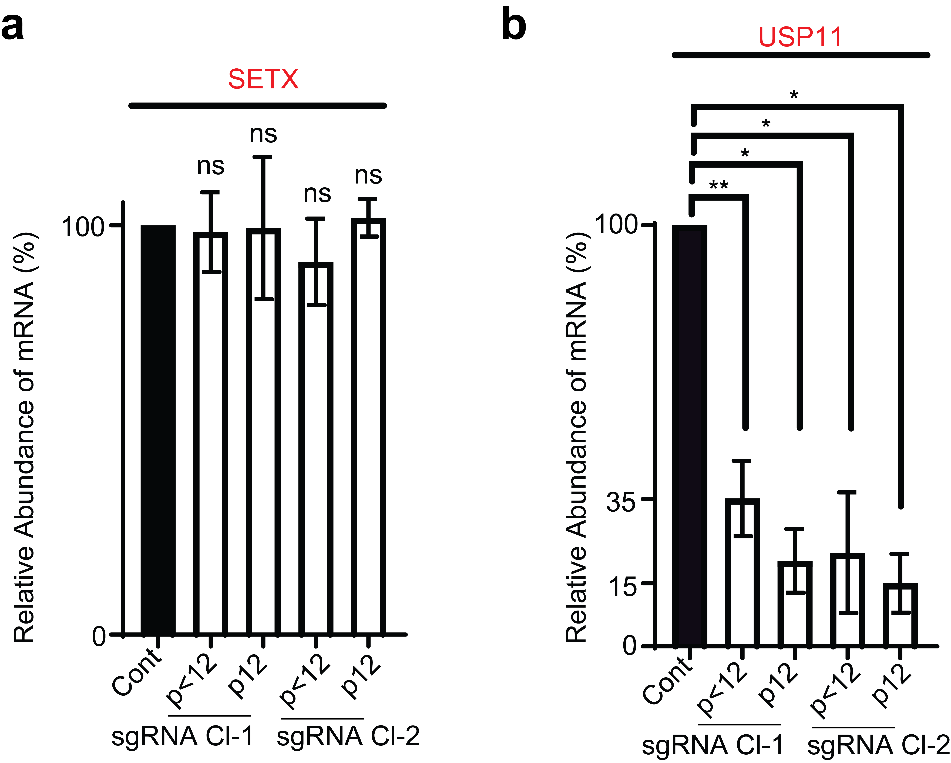
**a)** Lysates from Control HEK-293T cells (Cont) and USP11 knockout clone-2 at passages 12-15 (p12) were subjected to DNA/RNA immunoprecipitation (DRIP) using S9.6 antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3 and actin) loci. Raw % input values were normalised to control cells and presented as average ± SD from 3 biological repeats. ns; p>0.05, Student’s t-test.

**b)** Raw data from a representative experiment of DRIP-qPCR for Figure 5.9a. Lysates from Control HEK-293T (HEK) and aged USP11 sgRNA Cl-2 (p12-15; p12) were subjected to DNA/RNA immunoprecipitation (DRIP) using S9.6 antibodies. Quantitative PCR was conducted using primers targeting nucleolar (28S and R7) and nuclear (ING3 and actin) loci. In vitro, on-bead ec-RNase H (ec-RH) treatment served as a signal validation control.



**Figure 5.10 Aged USP11 sgRNA Clones over-express senataxin.**

**a)** Lysates from parental HEK-293T cells (Cont) and USP11 knockout clones 1 and 2 (Cl-1 and Cl-2) at passages 12-15 (p12) were fractionated by SDS-PAGE and analysed by immunoblotting (*left panel*). SETX band intensities were normalised to actin and presented as fold-increase compared to control cells (*right panel*). Data are the mean of 3 biological repeats ± SD. \*p<0.05, Student’s t-test.



**Figure 5.11 SETX mRNA levels do not** **change over time as USP11 sgRNA clones are passaged.**

**a)b)** Total RNA was extracted from young (p2-9; p<12) and aged (p12-15; p12) USP11 sgRNA clones, reverse transcribed to cDNA and transcript levels of USP11 and SETX were quantified by qPCR. USP11 and SETX mRNA levels were first normalised to actin and then presented as % change compared to levels in control cells. Data are the mean ± SD from 3 biological repeats. \*p<0.05, \*\*p<0.01, ns; p>0.05, Student’s t-test.

To summarise data from this chapter:

* Young, early passage, USP11 knockout cell lines accumulate R-loops and display SETX knockdown in our experimental conditions.
* The catalytic activity of USP11 is required for maintenance of normal R-loop homeostasis in our experimental conditions.
* The catalytically inactive USP11C318S does not restore SETX steady-state levels and does not deubiquitinate Ubi-eGFP-SETX1-667, as compared to USP11WT.
* Aged, late passage, USP11 knockout cell lines revert to wild type R-loop homeostasis presumably via stabilising SETX steady-state levels.

## Discussion

Here, presented USP11 knockout data corroborated USP11 knockdown data from Chapter 4. Upon temporary USP11-depletion, the N-terminal domain of senataxin was observed to be significantly ubiquitinated, which was presumably the reason behind increased turnover rates of full-length senataxin. USP11 knockout cell lines allowed to study this effect on senataxin over a prolonged period. In line with our previous findings, USP11 knockout cell lines displayed significant senataxin depletion, implying that extended absence of USP11 allowed for proteasomal degradation of ~90% of senataxin. This led to the R-loop accumulation phenotype, which was rescued by complementing cells with wild type, but not active site mutant, USP11.

Notably, these findings are in line with recent research by Mischo et al., where the authors demonstrated that *S.cerevisiae* SETX homologue, Sen1p, is ubiquitinated and subsequently degraded (Mischo et al., 2018). Interestingly, both Sen1p and hs-USP11 were shown to be specifically degraded in G1 stage (Mischo et al., 2018; Durocher et al., 2015), suggesting a conserved mechanism that might regulate SETX/Sen1p levels via (de)ubiquitination.

USP11C318S expression

USP11C318S was expressed to a higher extent than the USP11WT (Fig. 5.6c). USP11C318S was generated by introducing a single substitution into the USP11WT plasmid. Therefore, the strikingly different expression pattern of USP11C318S and USP11WT was not due to different gene promoters or any other features of the backbone. As a result, it is hypothesised that the differential expression could have been triggered by an uncharacterised positive feedback loop. As the initial expression of USP11WT could have restored normal R-loop phenotype, it presumably led to a downregulation of USP11WT to levels sufficient to maintain normal R-loop homeostasis. On the other hand, persistent accumulation of R-loops in the USP11C318S clones could have triggered a constant over-expression of USP11C318S by a regulatory mechanism. Therefore, future research might be needed to confirm a positive feedback loop linking R-loops and USP11 expression.

Lessons from RT-qPCR data

Surprisingly, USP11 mRNA levels in USP11 knockout cells were found to be between 15% and 35% of the wild-type HEK-293T expression level. As the used qPCR primers anneal to a sequence within the exon 3 of USP11 mRNA, and the introduced stop codon is located in exon 1, the polymerase transcribed for at least 6000nt past the stop codon. This suggests that either 1) the premature stop codon triggers degradation of exon1 mRNA of USP11, but the transcription resumes due to abundant methionine codons in the intron 1 of *USP11*, or 2) full-length USP11 mRNA is produced. However, generated USP11mRNA, whether truncated or not, is not translated, as USP11 sgRNA clones display a clear USP11 knockout and R-loop accumulation phenotype, which is rescued by the USP11WT complementation.

Finally, qPCR data revealed that SETX mRNA level was neither upregulated or downregulated throughout the consecutive passages of non-aged and aged USP11 sgRNA clones, suggesting that SETX is not regulated at the transcriptional level in our experimental settings. Therefore, SETX could be regulated at translational or post-translational levels. As it was demonstrated that the N-terminal domain of SETX is post-translationally modified in chapter 4, it was decided to continue the research into post-translational modifications of SETX. Since USP11-depleted cells accumulate ubiquitinated species of SETX, SETX must be a substrate for an E3 ligase. Therefore, it was hypothesised that a downregulation of SETX-specific E3 ligase could result in reduced levels of ubiquitinated SETX, and therefore, prolong its half-life. This led to a series of experiments described in the next chapter, which suggest a possible E3 ligase for senataxin and a potential explanation for this age-dependent adaptation phenomenon.

# Chapter 6: Possible complementation mechanism in late passages of USP11 knockout cells



## Introduction

Even though senataxin was shown to be ubiquitinated by this study and others (Bennett et al., 2013), a potential E3 ligase was not suggested. Due to time constraints, it was decided not to carry out a SETX E3 ligase screen but focuse on published literature to deduce a likely candidate.

Harper lab showed that USP11 and an E3 ligase, KEAP1, interact *in vitro* (Sowa et al., 2009). KEAP1-deletion rescued olaparib-sensitivity of USP11 knockout cells, which led to a series of experiments, implying that the two proteins might have antagonistic roles in regulating homologous recombination (Orthwein et al. 2015). Given the evidence of USP11-KEAP1 link, it was decided to investigate the role of KEAP1 in the context of R-loops and senataxin.

## Overview of KEAP1

KEAP1 (Kelch Like ECH Associated Protein 1) is a 70kDa substrate adaptor protein that is involved in the DNA damage response. KEAP1 binds to CUL3 and RBX1 to form an E3 ligase complex, which was shown to target a transcription factor, NRF2, for degradation (Zhang et al., 2004; Zhang et al., 2005, Suzuki and Yamamoto, 2018).

KEAP1 role changes from being a substrate-adaptor to a substrate, depending on the oxidative and electrophilic stress. Under unchallenged conditions, KEAP1 recruits the CUL3-RBX1 E3 complex to ubiquitinate NRF2, which is subsequently degraded. However, under stressed conditions, KEAP1 acts as a substrate for CUL3-RBX1 E3 complex and is subsequently degraded. This stabilises NRF2, which promotes transcription of stress response genes (Itoh et al., 1997; Zhang et al., 2005).

KEAP1 was shown to interact with NRF2 in the cytoplasm and with PALB2 (an HR factor) in the nucleus (Sun et al., 2007; Ma et al., 2012). This led the Durocher lab to elucidate the role of KEAP1 in inhibiting homologous recombination and counteracting USP11 (Orthwein et al., 2015).

Research on lung cancer revealed accumulation of methylated CpG islands over transcription start site of *KEAP1*, which in turn downregulated the transcription of the gene. Normal KEAP1 mRNA levels were restored by treating the cells with 5-Aza, which is a DNA methylation inhibitor (Wang et al., 2008), implying a transcriptional regulation.

All gathered data suggest that KEAP1 is a DNA damage response protein that is transcriptionally and post-translationally regulated (Wang et al., 2008; Zhang et al., 2005); it is not excluded from the nucleus (Sun et al., 2007); and it counteracts USP11 (Orthwein et al., 2015). Therefore, it was proposed that KEAP1 might be the E3 ligase that could ubiquitinate senataxin and be downregulated in aged USP11 knockout clones.

## Aims and objectives

The aim of this chapter was to find a putative E3 ligase that ubiquitinates senataxin. The objectives were to:

* Investigate the role of KEAP1 in R-loop homeostasis and its link to USP11.
* Investigate the role of KEAP1 in restoring R-loop phenotype and SETX steady-state levels in aged USP11 knockout cells.

## Results

## Olaparib sensitivity of USP11-depleted cells is rescued by RNase H1

USP11-depleted cells were reported to be sensitive to olaparib (Wiltshire et al., 2010). Olaparib is a PARP inhibitor that inhibits single-strand break (SSB) repair. This led to a series of experiments, which demonstrated that inhibition of SSB repair was lethal to USP11-depleted cells, because of their inability to carry out double-strand break repair via HR (Wiltshire et al., 2010; Orthwein et al., 2015). Combination of olaparib and USP11 siRNA led to a synthetic lethality, as cells were not able to repair single and double-strand breaks.

Durocher lab demonstrated that USP11-sensitivity to olaparib is rescued by a double KEAP1/USP11 knockout. This led to a discovery of antagonistic roles of KEAP1 & USP11 in controlling double-strand break repair by homologous recombination (Orthwein et al., 2015).

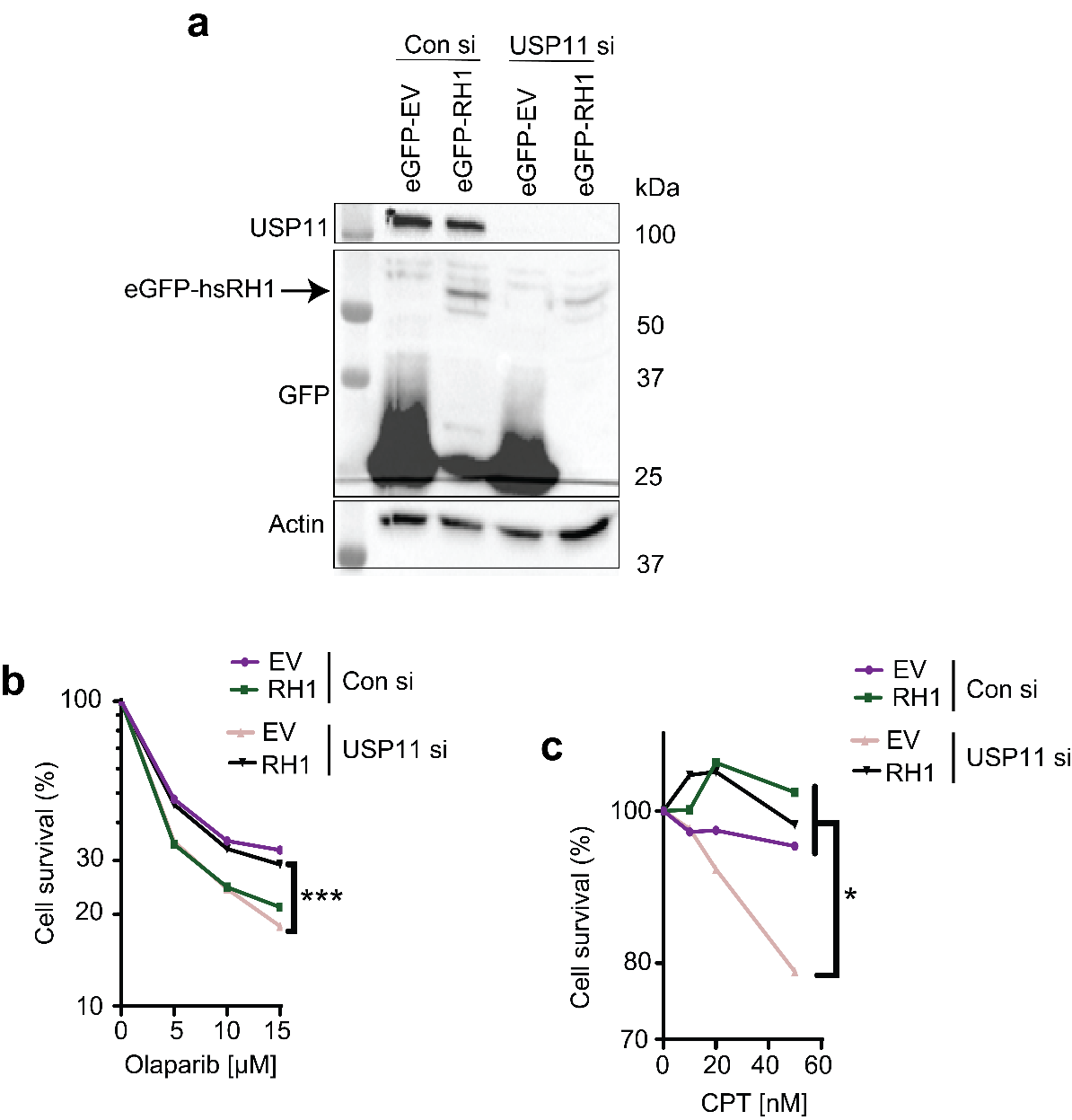
Interestingly, olaparib inhibits PARP1 (SSB repair factor) which was shown to interact with R-loop helicase, DHX9, in order to resolve R-loops (Cristini et al., 2018). Therefore, olaparib sensitivity of USP11-depleted cells could have been not only due to the loss of HR (Orthwein et al., 2015), but also due to olaparib-induced R-loop accumulation (Cristini et al., 2018).

In order to test that, simultaneous USP11-depletion and hs-RNase H1 over-expression were first validated in MRC-5 cells by immunoblotting (Fig. 6.1a). Next, a cell survival clonogenic assay was employed to investigate the role of R-loops in olaparib sensitivity of USP11-depleted cells (Fig. 6.1b). As expected, hs-RNase H1 over-expression rescued USP11-depleted cells from olaparib, suggesting that olaparib triggers R-loop formation in our experimental settings.

Surprisingly, MRC-5 cells transfected with scrambled siRNA and hs-RNase H1 were found to be sensitive to olaparib as well (Fig. 6.1b). It was hypothesised that olaparib did not trigger formation of sufficient amount of R-loops to fully engage over-expressed RNase H1 in these cells. This presumably led to off-target effects within the replication machinery, as *S.cerevisiae* RNase H and bovine RNase H1 were both demonstrated to resolve Okazaki fragments (Qiu et al., 1999; Murante et al., 1998).

Therefore, stressing MRC-5 cells with a more effective R-loop-inducing toxin, e.g. CPT (Cristini et al., 2018) should prevent RNase H1 from disassembling Okazaki fragments and interfering with replication (Fig. 6.1c). As expected, MRC-5 cells transfected with scrambled siRNA and hs-RNase H1 were found not to be sensitive to CPT at concentrations highly toxic to USP11-depleted cells.

To summarise, olaparib sensitivity of USP11-depleted cells can be rescued by either KEAP1 depletion or RNase H1 over-expression. Therefore, it was hypothesised that KEAP1 might not only be involved in HR, but also in R-loop homeostasis. If RNase H1 over-expression led to R-loop resolution, then KEAP1-depletion could have the same result. In order to test that, it was decided to investigate if KEAP1 depletion would also rescue R-loop accumulation phenotype in USP11-depleted cells (section 6.3.2).



**Figure 6.1 RNase H1 rescues sensitivity to olaparib of USP11-depleted cells.**

**a)** Lysates from MRC-5 cells transfected with scrambled (Con) or USP11 siRNAs and eGFP-RNaseH1 (eGFP-RH1) or empty vector (eGFP-EV) plasmids were fractionated by SDS-PAGE and analysed by immunoblotting.

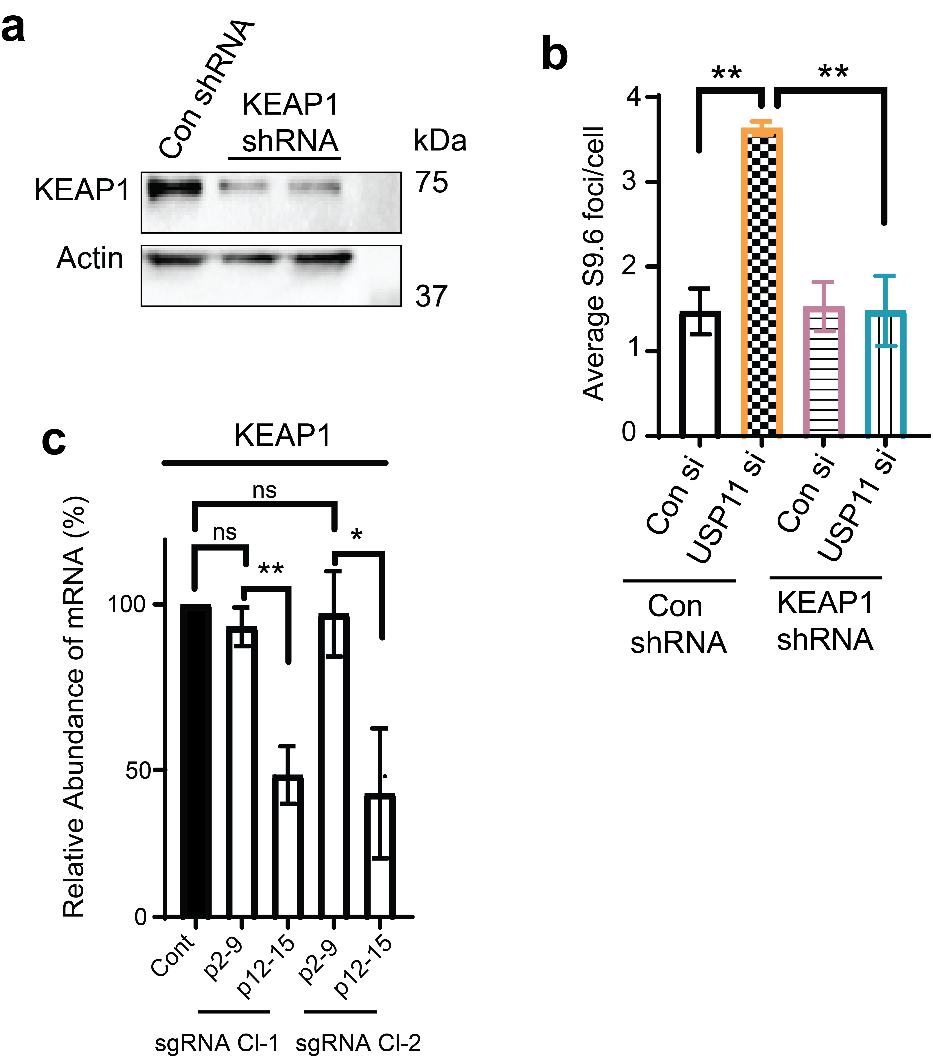
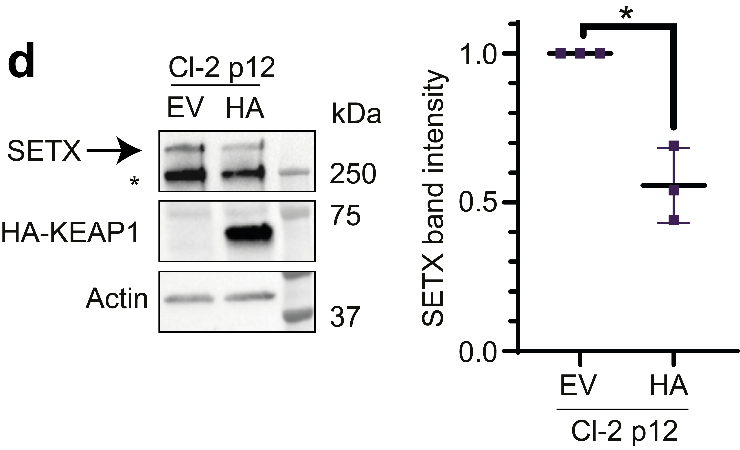
**b)c)** MRC-5 cells transfected with scrambled (Con) or USP11 siRNAs and eGFP-RNaseH1 (RH1) or empty vector (eGFP-EV) plasmids were incubated with the indicated doses of olaparib or camptothecin (CPT) for 24 and 1.5 hours respectively, and left to grow for 7 days. The surviving colonies were counted and % survival calculated relative to scrambled-treated cells. Data are the mean from 3 biological repeats. \*p<0.05, \*\*\*p<0.001, Student’s t-test.

## KEAP1 depletion rescues the R-loop accumulation phenotype in USP11-depleted cells

A heterogeneous KEAP1-depleted MRC-5 cell line was generated using a KEAP1 shRNA lentivirus targeting KEAP1 mRNA (Fig. 6.2a). Lenti KEAP1 cell line was later transfected with USP11 siRNA and used for S9.6 immunofluorescence experiment (Fig. 6.2b). Interestingly, USP11 depletion triggered significant R-loop accumulation, which was rescued by a double USP11/KEAP1 depletion. This might be the first up-to-date evidence suggesting that KEAP1 might have a role in R-loop homeostasis.

As double KEAP1/USP11 depletion rescued the R-loop accumulation phenotype in MRC-5 cells, it was hypothesised that aged USP11 knockout clones could have downregulated KEAP1 in order to stabilise SETX and revert R-loop accumulation phenotype back to wild type. Consequently, it was decided to assay KEAP1 mRNA in young and aged USP11 knockout clones (Fig. 6.2c). The acquired data suggest that KEAP1 mRNA levels were significantly reduced from ~100% to ~45% as the sgRNA Cl-1 and 2 aged, suggesting a correlation between KEAP1 downregulation and SETX upregulation. KEAP1 protein levels were not investigated as a reliable KEAP1 antibody was not available.

Next, it was decided to over-express KEAP1 in aged sgRNA Cl-1 cells in order to investigate its effect on the steady-state level of SETX (Fig. 6.2d). Notably, it was observed that upon KEAP1 overexpression, SETX steady-state level was reduced, suggesting that KEAP1 might promote SETX degradation, presumably through ubiquitination. Therefore, a link between KEAP1 and SETX was identified.



**Figure 6.2 KEAP1 knockdown rescues R-loop accumulation phenotype of USP11-depleted cells.**

**a)** Lysates from MRC-5 cells transducted with scrambled (Con shRNA) or KEAP1 shRNA lentiviruses were fractionated by SDS-PAGE and analysed by immunoblotting.

**b)** MRC-5 cells transduced with scrambled (Con shRNA) or KEAP1 shRNA lentiviruses and USP11 or scrambled (Con) siRNAs were harvested for S9.6 immunofluorescence. Data are the average ±SD from 3 biological repeats, each consisting of at least 100 cells. The average number of S9.6 foci/cell was calculated \*\* p<0.01, Student’s t-test.

**c)** Total RNA was extracted from young (p2-9) and aged (p12-15) USP11 sgRNA clones, reverse transcribed to cDNA and transcript levels of KEAP1 were quantified by qPCR. KEAP1 mRNA levels were first normalised to actin and then presented as % change compared to levels in control cells. Data are the mean ± SD from 3 biological repeats. ns; p>0.05, \*p<0.05, \*\*p<0.01, Student’s t-test.

**d)** Lysates from USP11 knockout clone 1 (Cl-1) at passages 12-15 (p12) transfected with HA-KEAP1 (HA) or empty vector (EV) were fractionated by SDS-PAGE and analysed by immunoblotting (*left panel*). SETX band intensities were normalised to actin and presented as fold-change compared to EV cells (*right panel*). Data are the mean of 3 biological repeats ± SD. \*p<0.05, Student’s t-test.

To summarise data from this chapter:

* Olaparib-sensitivity of USP11-depleted cells can be rescued by RNase H1 over-expression.
* KEAP1 depletion can rescue the R-loop accumulation phenotype of USP11-depleted cells according to the immunofluorescence data.
* KEAP1 mRNA expression was downregulated as the USP11 sgRNA Cl-1 and 2 cells were passaged according to qPCR data.
* Over-expression of endogenous KEAP1 in USP11 sgRNA Cl-2 p12 reduced SETX steady-state levels, suggesting that KEAP1 might regulate SETX stability.

## Discussion

Here, data suggesting that aged CRISPR-Cas9 cell lines undergo transcriptional adaptations in order to restore physiological functions was presented. Therefore, any data acquired from aged CRISPR-Cas9 cellular models should be interpreted with caution and separated from the data from the young models. According to my knowledge, this is the first description of culture adaptation of CRISPR-Cas9 HEK-293T cell line, which opens up a new avenue of research. Comparative analysis of young and aged CRISPR models might help to elucidate new feedback loops and genetic interactions activated by gene deletions. Additionally, CRISPR models might require a corroboration with other genetic approaches, like RNA interference, in order to validate the findings.

A connection between USP11, R-loops & KEAP1 was identified through an S9.6 IF read-out. The acquired data suggest that KEAP1 depletion rescues R-loop accumulation phenotype in USP11-depleted cells. As KEAP1 and USP11 were previously identified to ubiquitinate and deubiquitinate PALB2 respectively (Orthwein et al., 2015), the antagonistic roles of KEAP1 and USP11 in maintaining R-loop homeostasis are plausible. Nevertheless, this finding will need to be corroborated with DRIP-qPCR and S9.6 immunoblotting (slot blot) assays.

Next, a connection between SETX stability and KEAP1 expression was found. Although speculative, KEAP1 fits the profile of an E3 ligase that would target SETX for degradation. It is known that KEAP1 promotes degradation of NRF2 via ubiquitination (Suzuki and Yamamoto, 2018), localises to nuclei (Ma et al., 2012), counteracts USP11 (Orthwein et al., 2015), and is presumably involved in R-loop homeostasis (this study). Therefore, future research will be required to address the role of KEAP1 in R-loop homeostasis and SETX ubiquitination (Future directions, Chapter 7).

# Chapter 7: General discussion

## Overview

R-loops are transcription by-products that are tightly regulated. Physiological R-loops promote transcription termination, class-switch recombination and transcription-associated homologous recombination, while aberrant R-loops are linked to genome instability and cell death (Yu et al., 2003; Skourti-Stathaki et al., 2011, Yasuhara et al., 2018; Wahba et al., 2011; Sollier et al., 2014; Garcia-Rubio et al., 2015).

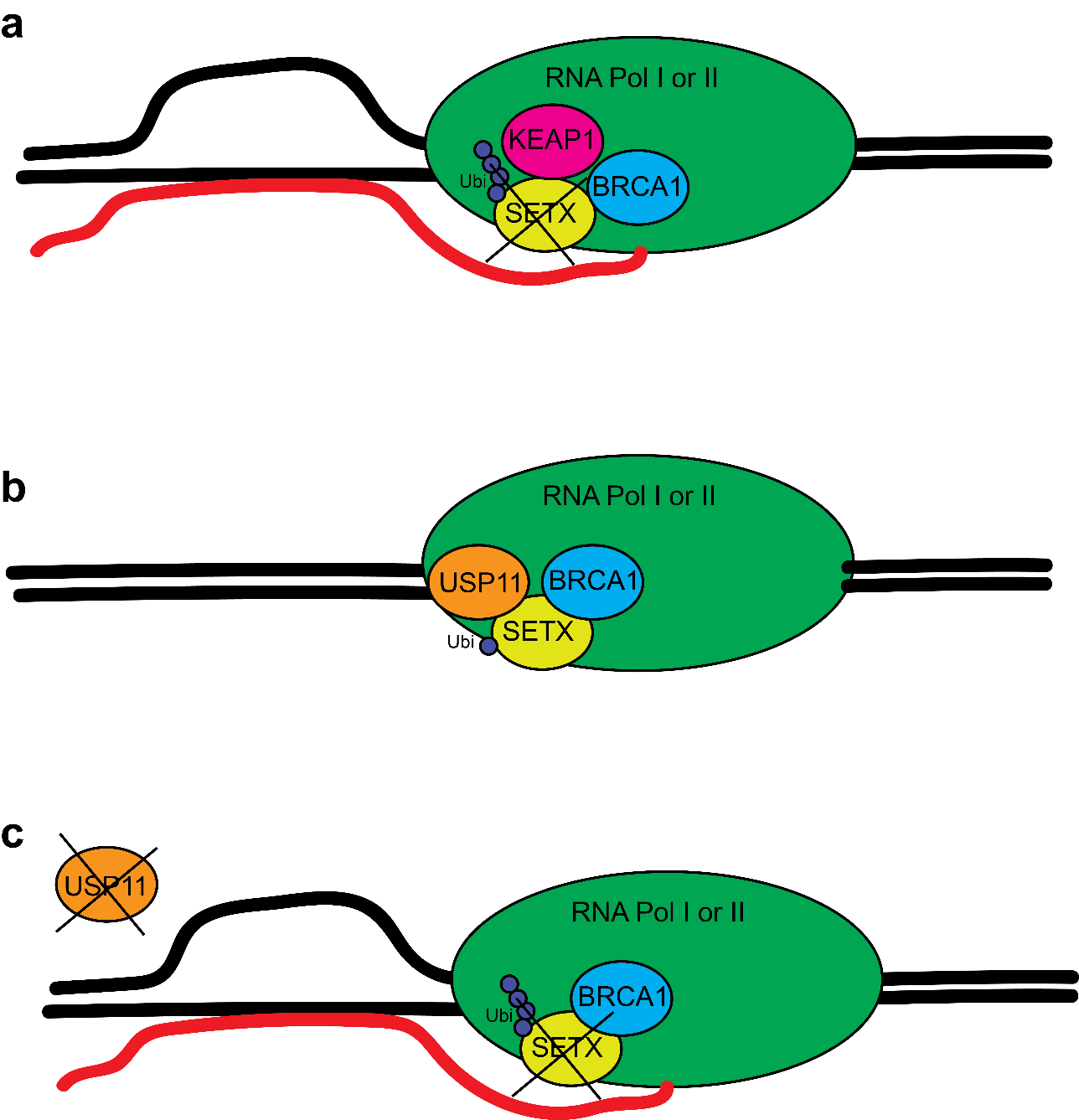
R-loop homeostasis is regulated by R-loop specific helicases (e.g. senataxin) and ribonucleases (e.g. RNase H1 and H2). Even though R-loop homeostasis requires tight regulation, the ways by which cells regulate R-loop metabolising enzymes are largely unknown. Here, an unbiased siRNA screen was carried out that led to the discovery of a novel R-loop regulating factor - ubiquitin specific protease 11 (USP11).

USP11 depletion by RNA interference triggered an accumulation of nuclear and nucleolar R-loops in human cells. Literature analysis suggested that SETX might be a putative substrate for USP11. Depletion of USP11 or SETX triggered R-loop accumulation phenotype, which was not exacerbated by a double USP11/SETX depletion, suggesting USP11-SETX epistasis. This result was further demonstrated by a cell survival assay, where USP11-, SETX-, and USP11/SETX-depleted cells were treated with R-loop inducing agents: camptothecin and formaldehyde.

The data obtained from cycloheximide-chase and qPCR experiments implied that USP11 regulates the turnover rate of SETX via post-translational modifications. A series of pull-down experiments provided evidence suggesting that: USP11 interacts with SETX1-667; depletion of USP11 triggers excessive ubiquitination of SETX1-667; and that USP11 *in vitro* deconjugates ubiquitin from Ub-SETX1-667. Moreover, a previously undescribed interaction between the N-terminal domain of SETX and RNA polymerase I was demonstrated, which resembles the interaction of Sen1p with RNA polymerase I in *Saccharomyces cerevisiae* (Kawauchi et al., 2008).

In order to corroborate USP11-SETX epistasis, USP11-deleted cell lines were generated using CRISPR-Cas9 technology. USP11-deleted cells demonstrated R-loop accumulation and SETX-depletion phenotypes, which were restored to physiological levels by a complementation with wild type, but not active site mutant, USP11.

Continued growth of USP11-deleted cells (passage≥12) triggered an adaptation, which restored physiological R-loop and SETX steady-state levels in both sgRNA Cl-1 & 2. Literature analysis led to a hypothesis that an E3 ligase KEAP1 might have an antagonistic function to USP11 in R-loop homeostasis and consequently, have a role in the adaptation of the CRISPR clones. Immunofluorescence data showed that KEAP1/USP11-depleted cells did not display an R-loop accumulation phenotype, implying that KEAP1 and USP11 antagonistically modulate R-loop homeostasis. Therefore, it was hypothesised that KEAP1 might be gradually downregulated over the age of USP11 sgRNA clones in order to stabilise SETX and restore physiological R-loop levels. qPCR data demonstrated that KEAP1 mRNA levels are halved in aged CRISPR cell lines, suggesting a possible transcriptional adaptation mechanism that alleviates deleterious effects of *USP11* deletion. Finally, KEAP1 over-expression was shown to trigger reduction in SETX steady-state levels, suggesting that KEAP1 might ubiquitinate SETX. As a result, a model was suggested where USP11 and KEAP1 work together to regulate the steady-state levels of SETX, which in turn resolves R-loops (Fig. 7.1)



**Figure 7.1 Proposed model of R-loop resolution via SETX.**

**a)** An untested, suggested KEAP1-SETX-BRCA1-RNA Pol I/II complex. BRCA1 recruits SETX to RNA Pol I or II. KEAP1 presumably ubiquitinates SETX, which leads to SETX degradation and preservation of the R-loop.

**b)** A suggested USP11-SETX-BRCA1-RNA Pol I/II complex. BRCA1 recruits SETX to RNA Pol I or II. USP11 deubiquitinates SETX, which stabilises SETX and enables R-loop resolution.

**c)** A suggested model of R-loop formation in USP11-depleted cells. In the absence of USP11, BRCA1-recruited SETX (Hatchi et al., 2015) is ubiquitinated and subsequently degraded, which results in the preservation of the R-loop.

## Future directions

MG132 leads to R-loop accumulation

MG132 treatment was shown to induce R-loop accumulation at nuclear and nucleolar foci in Chapter 3. MG132 inhibits proteasomes, which are responsible for protein degradation. Therefore, it is possible that MG132 treatment could reduce turnover rates of transcriptional activators; and consequently trigger elevated transcription rates and R-loop accumulation. This hypothesis can be assessed by conducting a 5-ethynyl uridine staining, as it yields quantitative mRNA staining, which reflects transcription rates.

Alternatively, MG132 treatment could trigger prolonged chromatin binding of unknown factors that either stabilise R-loops or promote their formation. In order to assay this hypothesis, a DRIP-Mass Spec approach could be employed to analyse R-loop proteome in the presence or absence of MG132, as described by Gromak or Cheung labs (Cristini et al., 2018; Wang et al., 2018).

Another approach to assay a potential role of the proteasome in R-loop homeostasis would be through an E3 ligase siRNA screen that would resemble described here DUB siRNA screen.

Ubiquitination of SETX

*In silico* search suggested that SETX1-667 domain contains ~40 lysines that could potentially be ubiquitinated. A mass spectrometry or extensive site-directed mutagenesis approaches could be employed in order to map the ubiquitinated lysines.

Additionally, linkage-specific K11 antibodies could be used to immunoprecipitate ubiquitinated SETX1-667 to investigate the nature of ubiquitination in the absence of USP11.

Moreover, a question was posed why USP11 depletion causes R-loop accumulation if ubiquitinated SETX is not immediately degraded. As an answer, a hypothesis was brought to light suggesting that ubiquitination might somehow deactivate or re-localise SETX prior to its degradation. Recently, it was shown that once ubiquitinated, DNA damage response factor Chk1 is sequestered to proteasomes by hHR23A protein (Tan et al., 2015).

hHR23A and B proteins contain ubiquitin-like and ubiquitin-associated domains that interact with proteasomes and ubiquitinated proteins respectively (Chen and Madura, 2002). Therefore, it is possible that once SETX is ubiquitinated, it interacts with hHR23, which leads to sequestration of SETX, and its subsequent proteasomal degradation. In order to test this, SETX immunoprecipitation could be followed by immunoblotting with anti-hHR23 antibodies. If the SETX-hHR23 interaction is true, over-expression of hHR23 in presence of a proteasomal inhibitor MG132 should not trigger SETX degradation but reduce SETX capacity to bind to R-loop forming loci, which could be tested by X-ChIP.

USP11, SETX and RNA Polymerase I link

Our co-immunoprecipitation data implies that SETX can be found in complexes with RPA194 (RNA Polymerase I subunit) and USP11, and our confocal images suggest that USP11 is not excluded from nucleoli, which overall suggest that formation of SETX-USP11-RPA194 complex is plausible. In order to assay this hypothesis, subsequent experiments are suggested.

Co-immunoprecipitation experiments using anti-USP11 or anti-RPA194 antibodies to show that these proteins form a complex with each other. USP11, SETX, and RPA194 X-ChIP experiments followed by qPCR using DRIP primers to investigate if these proteins bind to the same R-loop forming loci. A series of experiments involving radioactive RNA labelling followed by autoradiography to investigate if USP11, SETX, and USP11/SETX depletions affect transcription rates of RNA Polymerase I and ribosomal RNA processing.

USP11-R-loops feedback loop

Interestingly, immunoblotting data of complemented sgRNA cell lines suggest that catalytically inactive USP11C318S is over-expressed when compared to USP11WT. Therefore, it is hypothesised that USP11 expression levels might be tightly regulated to match R-loop levels through a positive feedback loop. In order to assay this, an RT-qPCR approach could be employed to measure USP11 mRNA levels in USP11WT- or USP11C318S-complemented CRISPR cell lines. Alternatively, if USP11 regulation occurs at the post-translational level, cycloheximide chase approach could be employed to measure the half-life of USP11WT and USP11C318S.

USP11 mRNA levels were present in USP11 knockout cell lines according to our RT-qPCR data (Fig. 5.11b). USP11 qPCR primers were designed to bind to exon 3 of *USP11*, while the introduced stop codons were found in exon 1 in both sgRNA clones. The distance between the stop codon and the primer annealing locus was ~6kb, implying that the detected mRNA was at least 6kb long. In order to assay if USP11 mRNA in USP11-deleted cell lines is 6kb long, a qPCR approach involving primers that anneal between exon 1 and intron 2 of *USP11* could be used. If the USP11 mRNA is shorter than 6kb, *USP11* gene might contain an alternative start codon that codes for a shorter, non-functional variant of USP11.

KEAP1-USP11-SETX-BRCA1 complex

Our model suggests that KEAP1 and USP11 might regulate R-loop homeostasis via SETX proteostasis (Fig. 7.1). It was shown that KEAP1 and USP11 regulate the association of BRCA1-PALB2-BRCA2 complex that initiates homologous recombination (Orthwein et al., 2015). Therefore, future research is required to assay if KEAP1 is the E3 ligase for USP11-substrates including SETX.

We demonstrated that USP11-depleted cells are sensitive to formaldehyde, which was shown by others to induce R-loops via depletion of BRCA2 (Tan et al., 2017). As USP11 depletion exacerbated cell death among formaldehyde-treated cells, BRCA2 and USP11 may not be involved in the same pathway that regulates R-loop homeostasis.

As BRCA1 and BRCA2 depletions were shown to result in distinct R-loop accumulation phenotypes (Bhatia et al., 2014; Hatchi et al., 2015) and BRCA1 was shown to be in a complex with SETX (Hatchi et al., 2015), it is plausible to hypothesise that BRCA2 may not interact with SETX.

Taken together, SETX-BRCA1 complex (Hatchi et al., 2015) might be regulated by USP11 and KEAP1. In line with this, USP11 was recently shown to interact with BRCA1 (Herold et al., 2019). Therefore, future pull-down and X-ChIP assays are required to examine the existence of USP11-SETX-BRCA1 and KEAP1-SETX-BRCA1 complexes at transcription sites of RNA Polymerase I and II.

Furthermore, Hatchi et al. (2015) demonstrated that BRCA1 is essential for SETX recruitment to transcription end sites, while Herold et al. (2019) demonstrated that BRCA1 depletion triggers SETX recruitment to promoter proximal regions. Therefore, BRCA1 might have locus-dependent antagonistic functions in SETX recruitment. To assay if the function of USP11/SETX interaction is upstream of the BRCA1/SETX function, a clonogenic cell survival assay is suggested. Olaparib-treated USP11-depleted cells could be rescued by overexpression of SETX but not BRCA1 if the BRCA1/SETX function is downstream. Moreover, USP11 and SETX X-ChIP-seq experiments might shed light on whether USP11 regulates SETX at both promoter proximal regions and transcription end sites. It would also be of importance to demonstrate SETX binding sites by X-ChIP-seq in USP11- or BRCA1-depleted cells in order to investigate the role of these proteins in SETX recruitment.

Finally, USP11-SETX-BRCA1 interaction might be cell-cycle dependent as both sc-Sen1p and hs-USP11 were shown to be degraded in G1 (Mischo et al., 2018; Orthwein et al., 2015). Therefore, it is plausible to hypothesise that SETX steady-state levels might also fluctuate, perhaps via differential regulation of USP11 or KEAP1. Consequently, SETX/USP11/KEAP1 immunoblotting of asynchronised and G1-arrested cells is suggested.

The role of KEAP1 in R-loop homeostasis

DRIP-qPCR and slot blot approaches can be used to demonstrate if KEAP1 antagonises USP11 in the context of R-loop homeostasis. If that is true, over-expression of KEAP1 in aged CRISPR cell lines should restore R-loop accumulation phenotype as it reduces SETX steady-state levels. On the contrary, KEAP1 downregulation in young CRISPR cell lines could restore wild-type phenotype and presumably upregulate steady-state levels of SETX.

Moreover, cycloheximide-chase and pull-down experiments could be employed to indirectly investigate if KEAP1 controls stability of SETX via ubiquitination. Depletion of KEAP1 should prolong the half-life of SETX and reduce the extent to which SETX is ubiquitinated. Finally, an *in vitro* ubiquitination assay could be conducted to directly investigate KEAP1- SETX interaction.

## USP11 as a drug target for novel therapies

Mitoxantrone was described to be a small-molecule inhibitor of USP11 and suggested for use in cancer therapies. It effectively kills tumours that are dependent on USP11 over-expression (Burkhart et al., 2013). Clonogenic cell survival assays described here revealed that USP11-depleted cells are hypersensitive to R-loop inducing toxins: camptothecin and formaldehyde. Consequently, a combination of an R-loop-inducing agent and mitoxantrone might be a novel method of treatment for USP11-over-expressing tumours.

Furthermore, mitoxantrone might be also effective for disorders caused by pathogenic expression or activity of SETX. For instance, gain-of-function SETX mutation is found in amyotrophic lateral sclerosis type 4 (ALS4) (Bennett et al., 2013). Down-regulation of SETX via mitoxantrone in ALS4 patients could potentially alleviate symptoms of the disorder.

Finally, patients with disorders like ALS, which are characterised by an R-loop accumulation phenotype and intact USP11-SETX-KEAP1 pathway, might benefit from treatment with USP11 activators or KEAP1 inhibitors (Walker et al., 2017; Abed et al., 2015).

## Summary

This thesis described USP11 as a novel R-loop regulator. It showed that USP11 interacts with and deubiquitinates R-loop-specific helicase, senataxin. It provided evidence of culture adaptation in CRISPR-edited cell lines and suggested that the adaptation mechanism could have occurred via transcriptional downregulation of the E3 ligase KEAP1. These findings might contribute to the development of novel methods of treatment for cancer and neurological disorders and open up new avenues to study CRISPR-edited cell lines to unravel novel regulatory feedback mechanisms.

# References

Abdul Rehman, S.A., Kristariyanto, Y.A., Choi, S.Y., Nkosi, P.J., Weidlich, S., Labib, K., Hofmann, K., and Kulathu, Y. (2016). MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. Mol. Cell *63*, 146–155.

Alagoz, M., Chiang, S.-C., Sharma, A., and El-Khamisy, S.F. (2013). ATM deficiency results in accumulation of DNA-topoisomerase I covalent intermediates in neural cells. PLoS One *8*, e58239.

Al-Hakim, A., Escribano-Diaz, C., Landry, M.C., Donnell, L., Panier, S., Szilard, R.K., and Durocher, D. (2010). The ubiquitous role of ubiquitin in the DNA damage response. DNA Repair (Amst). *9*, 1229–1240.

Al-Salihi, M. A., Herhaus, L., Macartney, T., and Sapkota, G.P. (2012). USP11 augments TGF signalling by deubiquitylating ALK5. Open Biol. *2*, 120063–120063.

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

Arab, K., Karaulanov, E., Musheev, M., Trnka, P., Schäfer, A., Grummt, I., and Niehrs, C. (2019). GADD45A binds R-loops and recruits TET1 to CpG island promoters. Nat. Genet. *51*, 217–223.

Arana, M.E., Kerns, R.T., Wharey, L., Gerrish, K.E., Bushel, P.R., and Kunkel, T. a (2012). Transcriptional responses to loss of RNase H2 in Saccharomyces cerevisiae. DNA Repair (Amst). *11*, 933–941.

Argaud, D., Boulanger, M.-C., Chignon, A., Mkannez, G., and Mathieu, P. (2019). Enhancer-mediated enrichment of interacting JMJD3–DDX21 to ENPP2 locus prevents R-loop formation and promotes transcription. Nucleic Acids Res. 1–15.

Arning, L., Epplen, J.T., Rahikkala, E., Hendrich, C., Ludolph, A.C., and Sperfeld, A.D. (2013). The SETX missense variation spectrum as evaluated in patients with ALS4-like motor neuron diseases. Neurogenetics *14*, 53–61.

Azvolinsky, A., Giresi, P.G., Lieb, J.D., and Zakian, V. a (2009). Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae. Mol. Cell *34*, 722–734.

Baens, M., Noels, H., Broeckx, V., Hagens, S., Fevery, S., Billiau, A.D., Vankelecom, H., and Marynen, P. (2006). The dark side of EGFP: Defective polyubiquitination. PLoS One *1*, 1–6.

Baker, K.E., and Parker, R. (2004). Nonsense-mediated mRNA decay: Terminating erroneous gene expression. Curr. Opin. Cell Biol. *16*, 293–299.

Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). M6A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell *15*, 707–719.

Beale, R.C.L., Petersen-Mahrt, S.K., Watt, I.N., Harris, R.S., Rada, C., and Neuberger, M.S. (2004). Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. J. Mol. Biol. *337*, 585–596.

Beletskii, A., and Bhagwat, A. (1996). Transcription-induced mutations: increase in C to T mutations in the nontranscribed strand during transcription in Escherichia coli. Proc. Natl. … *93*, 13919–13924.

Bennett, C.L., Chen, Y., Vignali, M., Lo, R.S., Mason, A.G., Unal, A., Huq Saifee, N.P., Fields, S., and La Spada, A.R. (2013). Protein Interaction Analysis of Senataxin and the ALS4 L389S Mutant Yields Insights into Senataxin Post-Translational Modification and Uncovers Mutant-Specific Binding with a Brain Cytoplasmic RNA-Encoded Peptide. PLoS One *8*, e78837.

Bermejo, R., Doksani, Y., Capra, T., Katou, Y., Tanaka, H., Shirahige, K., and Foiani, M. (2007). Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. 1921–1936.

Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F., et al. (2013). An estimation of the number of cells in the human body. Ann. Hum. Biol. *40*, 463–471.

Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. Science (80-.). *300*, 764.

Bibikova, M., Golic, M., Golic, K.G., and Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics *161*, 1169–1175.

Bochman, M.L., Paeschke, K., and Zakian, V. a (2012). DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet. *13*, 770–780.

Bochman, M.L., Sabouri, N., and Zakian, V.A. (2010). Unwinding the functions of the Pif1 family helicases. DNA Repair (Amst). *9*, 237–249.

Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., and Carrico, R.J. (1986). Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. J. Immunol. Methods *89*, 123–130.

Bonnet, A., Grosso, A.R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S.C., Janbon, G., Géli, V., de Almeida, S.F., and Palancade, B. (2017). Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. Mol. Cell *67*, 608-621.e6.

Boros-Oláh, B., Dobos, N., Hornyák, L., Szabó, Z., Karányi, Z., Halmos, G., Roszik, J., and Székvölgyi, L. (2019). Drugging the R-loop interactome: RNA-DNA hybrid binding proteins as targets for cancer therapy. DNA Repair (Amst). 102642.

Boulé, J.-B., and Zakian, V. a (2007). The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. Nucleic Acids Res. *35*, 5809–5818.

Brown, J.S., Lukashchuk, N., Sczaniecka-Clift, M., Britton, S., le Sage, C., Calsou, P., Beli, P., Galanty, Y., and Jackson, S.P. (2015). Neddylation Promotes Ubiquitylation and Release of Ku from DNA-Damage Sites. Cell Rep. *11*, 704–714.

Buchler, N.E., Gerland, U., and Hwa, T. (2005). Nonlinear protein degradation and the function of genetic circuits. Proc. Natl. Acad. Sci. *102*, 9559–9564.

Burkhart, R. a., Peng, Y., Norris, Z. a., Tholey, R.M., Talbott, V. a., Liang, Q., Ai, Y., Miller, K., Lal, S., Cozzitorto, J. a., et al. (2013). Mitoxantrone Targets Human Ubiquitin-Specific Peptidase 11 (USP11) and Is a Potent Inhibitor of Pancreatic Cancer Cell Survival. Mol. Cancer Res. *11*, 901–911.

Caldecott, K.W. (2008). Single-strand break repair and genetic disease. Nat. Rev. Genet. *9*, 619–631.

Casellas, R., Basu, U., Yewdell, W.T., Chaudhuri, J., Robbiani, D.F., and Di Noia, J.M. (2016). Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat. Rev. Immunol. *16*, 164–176.

Castel, S.E., Ren, J., Bhattacharjee, S., Chang, A.-Y., Sánchez, M., Valbuena, A., Antequera, F., and Martienssen, R.A. (2014). Dicer Promotes Transcription Termination at Sites of Replication Stress to Maintain Genome Stability. Cell *159*, 572–583.

Castellano-Pozo, M., Santos-Pereira, J., Rondón, A.G., Barroso, S., Andújar, E., Pérez-Alegre, M., García-Muse, T., and Aguilera, A. (2013). R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. Mol. Cell *52*, 583–590.

Cerritelli, S., Frolova, E., Feng, C., and Grinberg, A. (2003). Failure to Produce Mitochondrial DNA Results in Embryonic Lethality in Rnaseh1 Null Mice. Mol. Cell *11*, 807–815.

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

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

Chakraborty, P., Huang, J.T.J., and Hiom, K. (2018). DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. Nat. Commun. *9*.

Chau, V., Tobias, J.W., Bachmair, A., Marriotr, D., Ecker, D.J., Gonda, D.K., and Varshavsky, A. (1989). A Multiubiquitin Chain is COnfined to Specific Lysine in a Targeted SHort-Lived Protein. Science (80-. ). *243*, 1576–1583.

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

Chávez, S., and Aguilera, A. (1997). The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. Genes Dev. *11*, 3459–3470.

Chávez, S., Beilharz, T., Rondón, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q., Lithgow, T., and Aguilera, A. (2000). A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. EMBO J. *19*, 5824–5834.

Chen, J., Bozza, W., and Zhuang, Z. (2011). Ubiquitination of PCNA and Its Essential Role in Eukaryotic Translesion Synthesis. Cell Biochem. Biophys. *60*, 47–60.

Chen, L., and Madura, K. (2002). Rad23 Promotes the Targeting of Proteolytic Substrates to the Proteasome. Mol. Cell. Biol. *22*, 4902–4913.

Chen, L., Chen, J.Y., Huang, Y.J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., et al. (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. Mol. Cell *69*, 412-425.e6.

Chen, Y., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M.L., Rabin, B.A., et al. (2004). DNA/RNA Helicase Gene Mutations in a Form of Juvenile Amyotrophic Lateral Sclerosis (ALS4). *34*, 1128–1135.

Chen, Y.Z., Hashemi, S.H., Anderson, S.K., Huang, Y., Moreira, M.C., Lynch, D.R., Glass, I.A., Chance, P.F., Bennett, C.L., Yüce, Ö., et al. (2006). Senataxin, the yeast Sen1p orthologue: Characterization of a unique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. Neurobiol. Dis. *23*, 97–108.

Chernikova, S.B., Razorenova, O. V., Higgins, J.P., Sishc, B.J., Nicolau, M., Dorth, J.A., Chernikova, D.A., Kwok, S., Brooks, J.D., Bailey, S.M., et al. (2012). Deficiency in mammalian histone H2B ubiquitin ligase Bre1 (Rnf20/Rnf40) leads to replication stress and chromosomal instability. Cancer Res. *72*, 2111–2119.

Chiang, H.C., Zhang, X., Li, J., Zhao, X., Chen, J., Wang, H.T.H., Jatoi, I., Brenner, A., Hu, Y., and Li, R. (2019). BRCA1-associated R-loop affects transcription and differentiation in breast luminal epithelial cells. Nucleic Acids Res. *47*, 5086–5099.

Chon, H., Sparks, J.L., Rychlik, M., Nowotny, M., Burgers, P.M., Crouch, R.J., and Cerritelli, S.M. (2013). RNase H2 roles in genome integrity revealed by unlinking its activities. Nucleic Acids Res. *41*, 3130–3143.

Chou, C.C., Alexeeva, O.M., Yamada, S., Pribadi, A., Zhang, Y., Mo, B., Williams, K.R., Zarnescu, D.C., and Rossoll, W. (2015). PABPN1 suppresses TDP-43 toxicity in ALS disease models. Hum. Mol. Genet. *24*, 5154–5173.

Chuang, T.W., Lu, C.C., Su, C.H., Wu, P.Y., Easwvaran, S., Lee, C.C., Kuo, H.C., Hung, K.Y., Lee, K.M., Tsai, C.Y., et al. (2019). The RNA Processing Factor Y14 Participates in DNA Damage Response and Repair. Food Sci. Hum. Wellness *13*, 402–415.

Ciccia, A., and Elledge, S.J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. Mol. Cell *40*, 179–204.

Cohen, S., Puget, N., Lin, Y.L., Clouaire, T., Aguirrebengoa, M., Rocher, V., Pasero, P., Canitrot, Y., and Legube, G. (2018). Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. Nat. Commun. *9*.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (80-. ). *339*, 819–823.

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

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

D’Arcy, P., Wang, X., and Linder, S. (2015). Deubiquitinase inhibition as a cancer therapeutic strategy. Pharmacol. Ther. *147*, 32–54.

Dantuma, N.P., and van Attikum, H. (2016). Spatiotemporal regulation of posttranslational modifications in the DNA damage response. EMBO J. *35*, 6–23.

David, S.S., O’Shea, V.L., and Kundu, S. (2007). Base-excision repair of oxidative DNA damage. Nature *447*, 941–950.

de Kock, L., Sabbaghian, N., Druker, H., Weber, E., Hamel, N., Miller, S., Choong, C.S., Gottardo, N.G., Kees, U.R., Rednam, S.P., et al. (2014). Germ-line and somatic DICER1 mutations in pineoblastoma. Acta Neuropathol. *128*, 583–595.

De Magis, A., Manzo, S.G., Russo, M., Marinello, J., Morigi, R., Sordet, O., and Capranico, G. (2019). DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. Proc. Natl. Acad. Sci. *116*, 816–825.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., and Vogel, J. (2011). UKPMC Funders Group UKPMC Funders Group Author Manuscript factor RNase III. *471*, 602–607.

Deng, T., Yan, G., Song, X., Xie, L., Zhou, Y., Li, J., Hu, X., Li, Z., Hu, J., Zhang, Y., et al. (2018). Deubiquitylation and stabilization of p21 by USP11 is critical for cell-cycle progression and DNA damage responses. Proc. Natl. Acad. Sci. 201714938.

Desai, S.D., Zhang, H., Rodriguez-Bauman, A., Yang, J.-M., Wu, X., Gounder, M.K., Rubin, E.H., and Liu, L.F. (2003). Transcription-Dependent Degradation of Topoisomerase I-DNA Covalent Complexes. Mol. Cell. Biol. *23*, 2341–2350.

Deshpande, A., and Newlon, C. (1996). DNA replication fork pause sites dependent on transcription. Science (80-. ). *272*, 1030–1033.

Dhir, A., Dhir, S., Borowski, L.S., Jimenez, L., Teitell, M., Rötig, A., Crow, Y.J., Rice, G.I., Duffy, D., Tamby, C., et al. (2018). Mitochondrial double-stranded RNA triggers antiviral signalling in humans. Nature *560*, 238–242.

Dickerson, S.K., Market, E., Besmer, E., and Papavasiliou, F.N. (2003). AID mediates hypermutation by deaminating single stranded DNA. J. Exp. Med. *197*, 1291–1296.

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

Elizalde, M., Ávila, M.A., Berasain, C., Elizalde, M., Urtasun, R., Azkona, M., Latasa, M.U., Goñi, S., Lujambio, A., Prieto, J., et al. (2014). Splicing regulator SLU7 is essential for maintaining liver homeostasis Find the latest version : Splicing regulator SLU7 is essential for maintaining liver homeostasis. *124*, 2909–2920.

Erkizan, H.V., Schneider, J. a, Sajwan, K., Graham, G.T., Griffin, B., Chasovskikh, S., Youbi, S.E., Kallarakal, A., Chruszcz, M., Padmanabhan, R., et al. (2015). RNA helicase A activity is inhibited by oncogenic transcription factor EWS-FLI1. Nucleic Acids Res. *5*, 1–12.

Filippo, J.S., Sung, P., and Klein, H. (2008). Mechanism of Eukaryotic Homologous Recombination. Annu. Rev. Biochem. *77*, 229–257.

Garcia, D.A., Baek, C., Estrada, M.V., Tysl, T., Bennett, E.J., Yang, J., and Chang, J.T. (2018). USP11 Enhances TGFβ-Induced Epithelial–Mesenchymal Plasticity and Human Breast Cancer Metastasis. Mol. Cancer Res. *16*, 1172 LP-1184.

García-Pichardo, D., Cañas, J.C., García-Rubio, M.L., Gómez-González, B., Rondón, A.G., and Aguilera, A. (2017). Histone Mutants Separate R Loop Formation from Genome Instability Induction. Mol. Cell *66*, 597–609.e5.

García-Rubio, M.L., Pérez-Calero, C., Barroso, S.I., Tumini, E., Herrera-Moyano, E., Rosado, I. V., and Aguilera, A. (2015). The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. PLoS Genet. *11*, 1–17.

Garneau, J.E., Dupuis, M.È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Ginno, P., Lott, P., Christensen, H., Korf, I., and Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell *45*, 814–825.

Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I., and Chedin, F. (2013). 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.

Gómez-Herreros, F., Schuurs-Hoeijmakers, J.H.M., McCormack, M., Greally, M.T., Rulten, S., Romero-Granados, R., Counihan, T.J., Chaila, E., Conroy, J., Ennis, S., et al. (2014). TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. Nat. Genet. *46*, 516–521.

González-Aguilera, C., Tous, C., Gómez-González, B., Huertas, P., Luna, R., and Aguilera, A. (2008). The THP1-SAC3-SUS1-CDC31 Complex Works in Transcription Elongation-mRNA Export Preventing RNA-mediated Genome Instability. Mol. Biol. Cell *19*, 4310–4318.

Gorthi, A., Romero, J.C., Loranc, E., Cao, L., Lawrence, L.A., Goodale, E., Iniguez, A.B., Bernard, X., Masamsetti, V.P., Roston, S., et al. (2018). EWS-FLI1 increases transcription to cause R-Loops and block BRCA1 repair in Ewing sarcoma. Nature *555*, 387–391.

Gowrishankar, J., and Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? Mol. Microbiol. *54*, 598–603.

Groh, M., Lufino, M.M.P., Wade-Martins, R., and Gromak, N. (2014). R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. PLoS Genet. *10*, e1004318.

Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A., Lanman, T., Chen, K., Schindler, A.B., et al. (2018). Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. Mol. Cell *69*, 426–437.e7.

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

Halász, L., Karányi, Z., Boros-oláh, B., Kuik-rózsa, T., Sipos, É., Nagy, É., Mosolygó-l, Á., Mázló, A., Rajnavölgyi, É., Halmos, G., et al. (2017). RNA-DNA hybrid (R-loop) immunoprecipitation mapping: an analytical workflow to evaluate inherent biases. 1–12.

Hamperl, S., and Cimprich, K. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. DNA Repair (Amst). *19*, 84–94.

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

Harper, J.W., and Elledge, S.J. (2007). The DNA Damage Response: Ten Years After. Mol. Cell *28*, 739–745.

Hatchi, E., Skourti-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., Dimitrov, S., Pathania, S., McKinney, K.M., Eaton, M.L., et al. (2015). BRCA1 Recruitment to Transcriptional Pause Sites Is Required for R-Loop-Driven DNA Damage Repair. Mol. Cell *57*, 636–647.

Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol. Cell *44*, 966–977.

Herhaus, L., Al-Salihi, M., MacArtney, T., Weidlich, S., and Sapkota, G.P. (2013). OTUB1 enhances TGFβ signalling by inhibiting the ubiquitylation and degradation of active SMAD2/3. Nat. Commun. *4*, 1–13.

Hermanns, T., Pichlo, C., Woiwode, I., Klopffleisch, K., Witting, K.F., Ovaa, H., Baumann, U., and Hofmann, K. (2018). A family of unconventional deubiquitinases with modular chain specificity determinants. Nat. Commun. *9*.

Herold, S., Kalb, J., Büchel, G., Ade, C.P., Baluapuri, A., Xu, J., Koster, J., Solvie, D., Carstensen, A., Klotz, C., et al. (2019). Recruitment of BRCA1 limits MYCN-driven accumulation of stalled RNA polymerase. Nature *567*, 545–549.

Herrera-Moyano, E., Megui, X., Garcia-Rubio, M.L., Barroso, S., and Aguilera, A. (2014). The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription–replication conflicts. Genes Dev. *28*, 735–748.

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Merhi, R.A., and Maiorano, D. (2017). An ATR-dependent function for the Ddx 19 RNA helicase in nuclear R-loop metabolism. 1–17.

Householder, K.T., DiPerna, D.M., Chung, E.P., Wohlleb, G.M., Dhruv, H.D., Berens, M.E., and Sirianni, R.W. (2015). Intravenous delivery of camptothecin-loaded PLGA nanoparticles for the treatment of intracranial glioma. Int. J. Pharm. *479*, 374–380.

Huang, T.T., Nijman, S.M.B., Mirchandani, K.D., Galardy, P.J., Cohn, M. a, Haas, W., Gygi, S.P., Ploegh, H.L., Bernards, R., and D’Andrea, A.D. (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. Nat. Cell Biol. *8*, 339–347.

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

Hussong, M., Börno, S.T., Kerick, M., Wunderlich, A., Franz, A., Sültmann, H., Timmermann, B., Lehrach, H., Hirsch-Kauffmann, M., and Schweiger, M.R. (2014). The bromodomain protein BRD4 regulates the KEAP1/NRF2-dependent oxidative stress response. Cell Death Dis. *5*, 1–11.

Hutchins, A.P., Liu, S., Diez, D., and Miranda-Saavedra, D. (2013). The repertoires of ubiquitinating and deubiquitinating enzymes in eukaryotic genomes. Mol. Biol. Evol. *30*, 1172–1187.

Ideguchi, H., Ueda, A., Tanaka, M., Yang, J., Tsuji, T., Ohno, S., Hagiwara, E., Aoki, A., and Ishigatsubo, Y. (2002). Structural and functional characterization of the USP11 deubiquitinating enzyme, which interacts with the RanGTP-associated protein RanBPM. Biochem. J. *367*, 87–95.

Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., and Hatayama, Ø.I. (1997). An Nrf2 / Small Maf Heterodimer Mediates the Induction of Phase II Detoxifying Enzyme Genes through Antioxidant Response Elements. *322*, 313–322.

Jackson, S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687–696.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jackson, S.P., and Durocher, D. (2013). Regulation of DNA damage responses by ubiquitin and SUMO. Mol. Cell *49*, 795–807.

Jangi, M., Fleet, C., Cullen, P., Gupta, S. V, Mekhoubad, S., Chiao, E., Allaire, N., Bennett, C.F., Rigo, F., Krainer, A.R., et al. (2017). SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. Proc. Natl. Acad. Sci. *114*, E2347–E2356.

Jansen, R., Embden, J.D.A. Van, Gaastra, W., and Schouls, L.M. (2002). Jansen\_Identiﬁcation of genes that are associated with DNA repeats in prokaryotes.pdf. *43*, 1565–1575.

Jendrisak, J. (1980). The use of amanitin to inhibit in vivo RNA synthesis and germination in wheat embryos. J. Biol. Chem. *255*, 8529–8533.

Jiang, L.L., Xue, W., Hong, J.Y., Zhang, J.T., Li, M.J., Yu, S.N., He, J.H., and Hu, H.Y. (2017). The N-terminal dimerization is required for TDP-43 splicing activity. Sci. Rep. *7*, 1–12.

Jiang, M., Instrell, R., Saunders, B., Berven, H., and Howell, M. (2011). Tales from an academic RNAi screening facility; FAQs. Brief. Funct. Genomics *10*, 227–237.

Jiménez, M., Urtasun, R., Elizalde, M., Azkona, M., Latasa, M.U., Uriarte, I., Arechederra, M., Alignani, D., Bárcena-Varela, M., Álvarez-Sola, G., et al. (2019). Splicing events in the control of genome integrity: role of SLU7 and truncated SRSF3 proteins. Nucleic Acids Res. *47*, 3450–3466.

Jin, L., Williamson, A., Banerjee, S., Philipp, I., and Rape, M. (2008). Mechanism of Ubiquitin-Chain Formation by the Human Anaphase-Promoting Complex. Cell *133*, 653–665.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided. Science (80-. ). *337*, 816–822.

Jiricny, J. (2006). The multifaceted mismatch-repair system. Nat. Rev. Mol. Cell Biol. *7*, 335–346.

Johnston, R., D’Costa, Z., Ray, S., Gorski, J., Paul Harkin, D., Mullan, P., and Panov, K.I. (2016). The identification of a novel role for BRCA1 in regulating RNA polymerase I transcription. Oncotarget *7*, 68097–68110.

Joo, H.-Y., Zhai, L., Yang, C., Nie, S., Erdjument-Bromage, H., Tempst, P., Chang, C., and Wang, H. (2007). Regulation of cell cycle progression and gene expression by H2A deubiquitination. Nature *449*, 1068–1072.

Joshi, R.S., Piña, B., and Roca, J. (2010). Positional dependence of transcriptional inhibition by DNA torsional stress in yeast chromosomes. EMBO J. *29*, 740–748.

Joshi, R.S., Piña, B., and Roca, J. (2010). Positional dependence of transcriptional inhibition by DNA torsional stress in yeast chromosomes. EMBO J. *29*, 740–748.

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

Kapadia, B., Nanaji, N.M., Bhalla, K., Bhandary, B., Lapidus, R., Beheshti, A., Evens, A.M., and Gartenhaus, R.B. (2018). Fatty Acid Synthase induced S6Kinase facilitates USP11-eIF4B complex formation for sustained oncogenic translation in DLBCL. Nat. Commun. *9*.

Karahan, G., Sayar, N., Gozum, G., Bozkurt, B., Konu, O., and Yulug, I.G. (2015). Relative expression of rRNA transcripts and 45S rDNA promoter methylation status are dysregulated in tumors in comparison with matched-normal tissues in breast cancer. Oncol. Rep. *33*, 3131–3145.

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

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

Kee, Y., Yang, K., Cohn, M. a., Haas, W., Gygi, S.P., and D’Andrea, A.D. (2010). WDR20 regulates activity of the USP12-UAF1 deubiquitinating enzyme complex. J. Biol. Chem. *285*, 11252–11257.

Kessler, B.M., and Edelmann, M.J. (2011). PTMs in Conversation: Activity and Function of Deubiquitinating Enzymes Regulated via Post-Translational Modifications. Cell Biochem. Biophys. *60*, 21–38.

Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C.W., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.B., Murphy, M.E., et al. (2015). SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell *27*, 617–630.

Kim, H., Choe, J., and Seo, Y. (1999). The sen1 + Gene of Schizosaccharomyces pombe , a Homologue of Budding Yeast SEN1, Encodes an RNA and DNA Helicase †. 14697–14710.

Komander, D., Clague, M.J., and Urbé, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat. Rev. Mol. Cell Biol. *10*, 550–563.

Koo, C.X.G.E., Kobiyama, K., Shen, Y.J., LeBert, N., Ahmad, S., Khatoo, M., Aoshi, T., Gasser, S., and Ishii, K.J. (2015). RNA polymerase III regulates cytosolic RNA:DNA hybrids and intracellular microRNA expression. J. Biol. Chem. *290*, 7463–7473.

Labbadia, J., and Morimoto, R.I. (2015). The Biology of Proteostasis in Aging and Disease. Annu. Rev. Biochem. *84*, 435–464.

Lafuente-Barquero, J., Luke-Glaser, S., Graf, M., Silva, S., Gómez-González, B., Lockhart, A., Lisby, M., Aguilera, A., and Luke, B. (2017). The Smc5/6 complex regulates the yeast Mph1 helicase at RNA-DNA hybrid-mediated DNA damage. PLoS Genet. *13*, 1–25.

Lam, E.Y.N., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013). G-quadruplex structures are stable and detectable in human genomic DNA. Nat. Commun. *4*, 1796.

Lee, D.H. (1998). Proteasome inhibitors: Valuable new tools for cell biologists. Trends Cell Biol. *8*, 397–403.

Lee, E.-W., Seong, D., Seo, J., Jeong, M., Lee, H.-K., and Song, J. (2015). USP11-dependent selective cIAP2 deubiquitylation and stabilization determine sensitivity to Smac mimetics. Cell Death Differ. 1–14.

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

Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell *122*, 365–378.

Liang, Z., Liang, F., Teng, Y., Chen, X., Liu, J., Longerich, S., Rao, T., Green, A.M., Collins, N.B., Xiong, Y., et al. (2019). Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination. Cell Rep. *26*, 564–572.e5.

Liao, C., Beveridge, R., Hudson, J.J.R., Parker, J.D., Chiang, S.C., Ray, S., Ashour, M.E., Sudbery, I., Dickman, M.J., and El-Khamisy, S.F. (2018). UCHL3 Regulates Topoisomerase-Induced Chromosomal Break Repair by Controlling TDP1 Proteostasis. Cell Rep. *23*, 3352–3365.

Lieber, M.R. (2008). The mechanism of human nonhomologous DNA end joining. J. Biol. Chem. *283*, 1–5.

Lima, W.F., and Crooke, S.T. (1997). Binding affinity and specificity of Escherichia coli RNase H1: Impact on the kinetics of catalysis of antisense oligonucleotide-RNA hybrids. Biochemistry *36*, 390–398.

Lin, C.H., Chang, H.S., and Yu, W.C.Y. (2008). USP11 stabilizes HPV-16E7 and further modulates the E7 biological activity. J. Biol. Chem. *283*, 15681–15688.

Liu, L., and Wang, J. (1987). Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA *84*, 7024–7027.

Liu, Z., Lee, A., and Gilbert, W. (1995). Gene disruption of a G4-DNA-dependent nuclease in yeast leads to cellular senescence and telomere shortening. Proc. Natl. Acad. Sci. *92*, 6002–6006.

Ma, J., Cai, H., Wu, T., Sobhian, B., Huo, Y., Alcivar, A., Mehta, M., Cheung, K.L., Ganesan, S., Kong, A.-N.T., et al. (2012). PALB2 Interacts with KEAP1 To Promote NRF2 Nuclear Accumulation and Function. Mol. Cell. Biol. *32*, 1506–1517.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. Science (80-. ). *339*, 823–826.

Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., et al. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. Cell Stem Cell *18*, 541–553.

Marabitti, V., Lillo, G., Malacaria, E., Palermo, V., Sanchez, M., Pichierri, P., and Franchitto, A. (2019). ATM pathway activation limits R-loop-associated genomic instability in Werner syndrome cells. Nucleic Acids Res. 1–18.

Marinello, J., Chillemi, G., Bueno, S., Manzo, S.G., and Capranico, G. (2013). Antisense transcripts enhanced by camptothecin at divergent CpG-island promoters associated with bursts of topoisomerase I-DNA cleavage complex and R-loop formation. Nucleic Acids Res. *41*, 10110–10123.

Marteijn, J.A., Lans, H., Vermeulen, W., and Hoeijmakers, J.H.J. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. Nat. Rev. Mol. Cell Biol. *15*, 465–481.

Martin-Tumasz, S., and Brow, D.A. (2015). Saccharomyces cerevisiae sen1 helicase domain exhibits 5′- to 3′-helicase activity with a preference for translocation on DNA rather than RNA. J. Biol. Chem. *290*, 22880–22889.

McClurg, U.L., and Robson, C.N. (2015). Deubiquitinating enzymes as oncotargets. Oncotarget *6*, 9657–9668.

Mermerian, A.H., Case, A., Stein, R.L., and Cuny, G.D. (2007). Structure-activity relationship, kinetic mechanism, and selectivity for a new class of ubiquitin C-terminal hydrolase-L1 (UCH-L1) inhibitors. Bioorganic Med. Chem. Lett. *17*, 3729–3732.

Mersaoui, S.Y., Yu, Z., Coulombe, Y., Karam, M., Busatto, F.F., Masson, J.-Y., and Richard, S. (2019). Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. 1–20.

Meyer, H.J., and Rape, M. (2014). Enhanced protein degradation by branched ubiquitin chains. Cell *157*, 910–921.

Mischo, H.E., Chun, Y., Harlen, K.M., Smalec, B.M., Dhir, S., Churchman, L.S., and Buratowski, S. (2018). Cell-Cycle Modulation of Transcription Termination Factor Sen1. Mol. Cell *70*, 312–326.e7.

Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol. Cell *41*, 21–32.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. *60*, 174–182.

Moreira, M.-C., Klur, S., Watanabe, M., Németh, A.H., Le Ber, I., Moniz, J.-C., Tranchant, C., Aubourg, P., Tazir, M., Schöls, L., et al. (2004). Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. Nat. Genet. *36*, 225–227.

Motohashi, H., and Yamamoto, M. (2004). Nrf2-Keap1 defines a physiologically important stress response mechanism. Trends Mol. Med. *10*, 549–557.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. Cell *102*, 553–563.

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

Nguyen, H.D., Leong, W.Y., Li, W., Reddy, P.N.G., Sullivan, J.D., Walter, M.J., Zou, L., and Graubert, T.A. (2018). Spliceosome mutations induce R loop-associated sensitivity to ATR inhibition in myelodysplastic syndromes. Cancer Res. *78*, 5363–5374.

Nicassio, F., Corrado, N., Vissers, J.H. a, Areces, L.B., Bergink, S., Marteijn, J. a, Geverts, B., Houtsmuller, A.B., Vermeulen, W., Di Fiore, P.P., et al. (2007). Human USP3 is a chromatin modifier required for S phase progression and genome stability. Curr. Biol. *17*, 1972–1977.

Nijman, S.M.B., Luna-Vargas, M.P. a, Velds, A., Brummelkamp, T.R., Dirac, A.M.G., Sixma, T.K., and Bernards, R. (2005). A genomic and functional inventory of deubiquitinating enzymes. Cell *123*, 773–786.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell *156*, 935–949.

Nowotny, M., Gaidamakov, S.A., Ghirlando, R., Cerritelli, S.M., Crouch, R.J., and Yang, W. (2007). Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. Mol. Cell *28*, 264–276.

Obeng, E.A., Chappell, R.J., Seiler, M., Chen, M.C., Campagna, D.R., Schmidt, P.J., Schneider, R.K., Lord, A.M., Wang, L., Gambe, R.G., et al. (2016). Physiologic Expression of Sf3b1 K700E Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. Cancer Cell *30*, 404–417.

Oberle, C., and Blattner, C. (2010). Regulation of the DNA Damage Response to DSBs by Post-Translational Modifications. Curr. Genomics *11*, 184–198.

Obrig, T.G., Culp, W.J., McKeehan, W.L., and Hardesty, B. (1971). The Mechanism by which Cycloheximide and Related Glutarimide Antibiotics Inhibit Peptide Synthesis on Reticulocyte Ribosomes. J. Biol. Chem. *246*, 174–181.

Orthwein, A., Noordermeer, S.M., Wilson, M.D., Landry, S., Enchev, R.I., Sherker, A., Munro, M., Pinder, J., Salsman, J., Dellaire, G., et al. (2015). A mechanism for the suppression of homologous recombination in G1 cells. Nature *528*, 422–426.

Osaka, H., Wang, Y.L., Takada, K., Takizawa, S., Setsuie, R., Li, H., Sato, Y., Nishikawa, K., Sun, Y.J., Sakurai, M., et al. (2003). Ubiquitin carboxy-terminal hydrolase L1 binds to and stabilizes monoubiquitin in neuron. Hum. Mol. Genet. *12*, 1945–1958.

Panier, S., and Durocher, D. (2009). Regulatory ubiquitylation in response to DNA double-strand breaks. DNA Repair (Amst). *8*, 436–443.

Parsons, J.L., Dianova, I.I., Khoronenkova, S. V, Edelmann, M.J., Kessler, B.M., and Dianov, G.L. (2011). USP47 is a deubiquitylating enzyme that regulates base excision repair by controlling steady-state levels of DNA polymerase β. Mol. Cell *41*, 609–615.

Paulsen, R.D., Soni, D. V, Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J. a, Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., et al. (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell *35*, 228–239.

Periyasamy, M., Patel, H., Lai, C.F., Nguyen, V.T.M., Nevedomskaya, E., Harrod, A., Russell, R., Remenyi, J., Ochocka, A.M., Thomas, R.S., et al. (2015). APOBEC3B-Mediated Cytidine Deamination Is Required for Estrogen Receptor Action in Breast Cancer. Cell Rep. *13*, 108–121.

Petryk, N., Kahli, M., D’Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., Thermes, C., Chen, C.L., and Hyrien, O. (2016). Replication landscape of the human genome. Nat. Commun. *7*, 1–13.

Pezone, A., Zuchegna, C., Tramontano, A., Romano, A., Russo, G., de Rosa, M., Vinciguerra, M., Porcellini, A., Gottesman, M.E., and Avvedimento, E.V. (2019). RNA Stabilizes Transcription-Dependent Chromatin Loops Induced By Nuclear Hormones. Sci. Rep. *9*, 1–12.

Phillips, D.D., Garboczi, D.N., Singh, K., Hu, Z., Leppla, S.H., and Leysath, C.E. (2013). The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. J. Mol. Recognit. *26*, 376–381.

Posse, V., Al-Behadili, A., Uhler, J.P., Clausen, A.R., Reyes, A., Zeviani, M., Falkenberg, M., and Gustafsson, C.M. (2019). RNase H1 directs origin-specific initiation of DNA replication in human mitochondria. PLOS Genet. *15*, e1007781.

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

Prakash, A., Natarajan, A., Marky, L. a, Ouellette, M.M., and Borgstahl, G.E.O. (2011). Identification of the DNA-Binding Domains of Human Replication Protein A That Recognize G-Quadruplex DNA. J. Nucleic Acids *2011*, 896947.

Qiu, J., Qian, Y., Frank, P., Wintersberger, U., and Shen, B. (1999). Saccharomyces cerevisiae RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. Mol. Cell. Biol. *19*, 8361–8371.

Rabe, B. (2013). Aicardi-Goutières syndrome: clues from the RNase H2 knock-out mouse. J. Mol. Med. (Berl). *91*, 1235–1240.

Raha, D., Wang, Z., Moqtaderi, Z., Wu, L., Zhong, G., Gerstein, M., Struhl, K., and Snyder, M. (2010). Close association of RNA polymerase II and many transcription factors with Pol III genes. Proc. Natl. Acad. Sci. *107*, 3639–3644.

Ran, F.A., Hsu, P.D.P., Wright, J., Agarwala, V., Scott, D. a, and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S., Chaudhuri, R., Fu, W., Prabhakar, B.S., Johnson, M.E., et al. (2008). A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. Proc. Natl. Acad. Sci. U. S. A. *105*, 16119–16124.

Ray, S., Panova, T., Miller, G., Volkov, A., Porter, A.C.G., Russell, J., Panov, K.I., and Zomerdijk, J.C.B.M. (2013). Topoisomerase IIα promotes activation of RNA polymerase i transcription by facilitating pre-initiation complex formation. Nat. Commun. *4*, 1512–1598.

Reijns, M. a M., Rabe, B., Rigby, R.E., Mill, P., Astell, K.R., Lettice, L. a, Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., et al. (2012). Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. Cell *149*, 1008–1022.

Reyes, A., Melchionda, L., Nasca, A., Carrara, F., Lamantea, E., Zanolini, A., Lamperti, C., Fang, M., Zhang, J., Ronchi, D., et al. (2015). RNASEH1 Mutations Impair mtDNA Replication and Cause Adult-Onset Mitochondrial Encephalomyopathy. Am. J. Hum. Genet. *97*, 186–193.

Rice, G., Patrick, T., Parmar, R., Taylor, C.F., Aeby, A., Aicardi, J., Artuch, R., Montalto, S.A., Bacino, C. a, Barroso, B., et al. (2007). Clinical and molecular phenotype of Aicardi-Goutieres syndrome. Am. J. Hum. Genet. *81*, 713–725.

Rivosecchi, J., Larochelle, M., Teste, C., Grenier, F., Malapert, A., Ricci, E.P., Bernard, P., Bachand, F., and Vanoosthuyse, V. (2019). Sen1 is required for RNA Polymerase III transcription termination in an R-loop independent manner. BioRxiv 571737.

Roy, D., and Lieber, M.R. (2009). G clustering is important for the initiation of transcription-induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter. Mol. Cell. Biol. *29*, 3124–3133.

Roy, D., Yu, K., and Lieber, M.R. (2008). Mechanism of R-loop formation at immunoglobulin class switch sequences. Mol. Cell. Biol. *28*, 50–60.

Roy, D., Zhang, Z., Lu, Z., Hsieh, C.-L., and Lieber, M.R. (2010). Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. Mol. Cell. Biol. *30*, 146–159.

Sahtoe, D.D., and Sixma, T.K. (2015). Layers of DUB regulation. Trends Biochem. Sci. *40*, 456–467.

Santos-Pereira, J., Herrero, A.B., Garcia-Rubio, M.L., Marin, A., Moreno, S., and Aguilera, A. (2013). The Npl3 hnRNP prevents R-loop-mediated transcription–replication conflicts and genome instability. Genes Dev. *27*, 2445–2458.

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

Sarkar, K., Han, S.S., Wen, K.K., Ochs, H.D., Dupré, L., Seidman, M.M., and Vyas, Y.M. (2018). R-loops cause genomic instability in T helper lymphocytes from patients with Wiskott-Aldrich syndrome. J. Allergy Clin. Immunol. *142*, 219–234.

Schoenfeld, A.R., Apgar, S., Dolios, G., Wang, R., and Aaronson, S. (2004). BRCA2 is ubiquitinated in vivo and interacts with USP11, a deubiquitinating enzyme that exhibits prosurvival function in the cellular response to DNA damage. Mol. Cell. Biol. *24*, 7444–7455.

Schwab, R.A., Nieminuszczy, J., Shah, F., Langton, J., Martinez, D.L., Liang, C., Cohn, M.A., Gibbons, R.J., Deands, A.J., and Niedzwiedz, W. (2015). The Fanconi Anemia Pathway Maintains Genome Stability. Mol. Cell *60*, 1–11.

Shahid, T., Soroka, J., Kong, E.H., Malivert, L., McIlwraith, M.J., Pape, T., West, S.C., and Zhang, X. (2014). Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. Nat. Struct. Mol. Biol. *21*, 962–968.

Shalhout, S., Haddad, D., Sosin, A., Holland, T.C., Al-Katib, A., Martin, A., and Bhagwat, A.S. (2014). Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. Mol. Cell. Biol. *34*, 4019–4032.

Shanbhag, N., and Rafalska-Metcalf, I. (2010). An ATM-Dependent Transcriptional Silencing Program is Transmitted Through Chromatin in Cis to DNA Double Strand Breaks. Cell *141*, 970–981.

Shen, W., Sun, H., De Hoyos, C.L., Bailey, J.K., Liang, X.H., and Crooke, S.T. (2017). Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops. Nucleic Acids Res. *45*, 10672–10692.

Shirai, C.L., Ley, J.N., White, B.S., Kim, S., Tibbitts, J., Shao, J., Ndonwi, M., Wadugu, B., Duncavage, E.J., Okeyo-Owuor, T., et al. (2015). Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. Cancer Cell *27*, 631–643.

Silva, B., Pentz, R., Figueira, A.M., Arora, R., Lee, Y.W., Hodson, C., Wischnewski, H., Deans, A.J., and Azzalin, C.M. (2019). FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. Nat. Commun. *10*, 1–16.

Silva, S., Camino, L.P., and Aguilera, A. (2018). Human mitochondrial degradosome prevents harmful mitochondrial R loops and mitochondrial genome instability [Genetics]. Proc. Natl. Acad. Sci. U. S. A.

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

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

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell *42*, 794–805.

Skourti-Stathaki, K., Torlai Triglia, E., Warburton, M., Voigt, P., Bird, A., and Pombo, A. (2019). R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes. Mol. Cell *73*, 930-945.e4.

Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., and Cimprich, K.A. (2014). Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. Mol. Cell 1–9.

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

Sowa, M.E., Bennett, E.J., Gygi, S.P., Harper, J.W., and Results, I.S. (2009). Supplemental Data Defining the Human Deubiquitinating Enzyme Interaction Landscape Table of Contents I . Supplemental Results 1 . Detailed analysis of four previously studied Dub complexes 2 . Detailed analysis of several largely unstudied Dub complexes a. Cell *138*.

Spiliotopoulos, A., Blokpoel Ferreras, L., Densham, R.M., Caulton, S.G., Maddison, B.C., Morris, J.R., Dixon, J.E., Gough, K.C., and Dreveny, I. (2018). Discovery of peptide ligands targeting a specific ubiquitin-like domain-binding site in the deubiquitinase USP11. J. Biol. Chem. jbc.RA118.004469.

Stavnezer, J., Guikema, J.E.J., and Schrader, C.E. (2008). Mechanism and regulation of class switch recombination. Annu. Rev. Immunol. *26*, 261–292.

Stein, H., and Hausen, P. (1969). Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. Science (80-. ). *166*, 393–395.

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

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

Stockum, A., Snijders, A.P., and Maertens, G.N. (2018). USP11 deubiquitinates RAE1 and plays a key role in bipolar spindle formation. PLoS One *13*, 1–21.

Stolz, R., Sulthana, S., Hartono, S.R., Malig, M., Benham, C.J., and Chedin, F. (2019). Interplay between DNA sequence and negative superhelicity drives R-loop structures. Proc. Natl. Acad. Sci. *116*, 6260–6269.

Stork, C.T., Bocek, M., Crossley, M.P., Sollier, J., Sanz, L.A., Chédin, F., Swigut, T., and Cimprich, K.A. (2016). Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. Elife *5*, 1–21.

Sträßer, K., Masuda, S., Mason, P., and Pfannstiel, J. (2002). TREX is a conserved complex coupling transcription with messenger RNA export. Nature *417*, 304–308.

Sun, H., Yabuki, A., and Maizels, N. (2001). A human nuclease specific for G4 DNA. Proc. Natl. Acad. Sci. USA *98*, 12444–12449.

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., and Dean, C. (2013). R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. Science *340*, 619–621.

Sun, W., Tan, X., Shi, Y., Xu, G., Mao, R., Gu, X., Fan, Y., Yu, Y., Burlingame, S., Zhang, H., et al. (2010). USP11 Negatively Regulates TNFα-Induced NF-κB Activation by Targeting on IκBα. Cell Signal *22*, 386–394.

Sun, Z., Zhang, S., Chan, J.Y., and Zhang, D.D. (2007). Keap1 Controls Postinduction Repression of the Nrf2-Mediated Antioxidant Response by Escorting Nuclear Export of Nrf2. Mol. Cell. Biol. *27*, 6334–6349.

Suraweera, A., Becherel, O.J., Chen, P., Rundle, N., Woods, R., Nakamura, J., Gatei, M., Criscuolo, C., Filla, A., Chessa, L., et al. (2007). Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. J. Cell Biol. *177*, 969–979.

Suzuki, H., Nagai, K., Yamaki, H., and Tanaka, N. (1969). On The Mechanism of Action of Bleomycin: Scission of DNA Strands in vitro and in vivo. J. Antibiot. Sept *22*, 446–448.

Swanson, D. a., Freund, C.L., Ploder, L., McInnes, R.R., and Valle, D. (1996). A ubiquitin C-terminal hydrolase gene on the proximal short arm of the X chromosome: Implications for X-linked retinal disorders. Hum. Mol. Genet. *5*, 533–538.

Takahashi, T., Burguiere-Slezak, G., Van der Kemp, P.A., and Boiteux, S. (2011). Topoisomerase 1 provokes the formation of short deletions in repeated sequences upon high transcription in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *108*, 692–697.

Takashima, H., Boerkoel, C.F., John, J., Saifi, G.M., Salih, M. a M., Armstrong, D., Mao, Y., Quiocho, F. a, Roa, B.B., Nakagawa, M., et al. (2002). Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. Nat. Genet. *32*, 267–272.

Tan, S.L.W., Chadha, S., Liu, Y., Gabasova, E., Perera, D., Ahmed, K., Constantinou, S., Renaudin, X., Lee, M.Y., Aebersold, R., et al. (2017). A Class of Environmental and Endogenous Toxins Induces BRCA2 Haploinsufficiency and Genome Instability. Cell *169*, 1105–1118.e15.

Tan, X., Liang, R.Y., and Chuang, S.M. (2015). HHR23A is required to control the basal turnover of Chk1. Cell. Signal. *27*, 2304–2313.

Tian, M., and Alt, F.W. (2000). Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. J. Biol. Chem. *275*, 24163–24172.

Tinnikov, A.A., and Samuels, H.H. (2013). A Novel Cell Lysis Approach Reveals That Caspase-2 Rapidly Translocates from the Nucleus to the Cytoplasm in Response to Apoptotic Stimuli. PLoS One *8*.

Tran, P.L.T., Pohl, T.J., Chen, C.F., Chan, A., Pott, S., and Zakian, V.A. (2017). PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. Nat. Commun. *8*, 1–10.

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell *168*, 644–656.

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., et al. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol. *11*, 1315–1324.

Vance, J.R., and Wilson, T.E. (2002). Yeast Tdp1 and Rad1-Rad10 function as redundant pathways for repairing Top1 replicative damage. Proc. Natl. Acad. Sci. U. S. A. *99*, 13669–13674.

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

Vologodskii, A. V., Levene, S.D., Klenin, K. V., Frank-Kamenetskii, M., and Cozzarelli, N.R. (1992). Conformational and thermodynamic properties of supercoiled DNA. J. Mol. Biol. *227*, 1224–1243.

Veloso, A., Biewen, B., Paulsen, M.T., Berg, N., Carmo de Andrade Lima, L., Prasad, J., Bedi, K., Magnuson, B., Wilson, T.E., and Ljungman, M. (2013). Genome-wide transcriptional effects of the anti-cancer agent camptothecin. PLoS One *8*, e78190.

Wahba, L., Amon, J.D., Koshland, D., Vuica-ross, M., and Biology, C. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA-DNA hybrids from generating genome instability. Mol. Cell *44*, 978–988.

Wahba, L., Gore, S.K., and Koshland, D. (2013). The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. Elife *2*, e00505.

Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W., Lukashchuk, V., Chiang, S.C., Ray, S., Mulcahy, P.J., et al. (2017). C9orf72 expansion disrupts ATM-mediated chromosomal break repair. Nat. Neurosci. *20*, 1225–1235.

Wang, I.X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M., and Cheung, V.G. (2018). Human proteins that interact with RNA / DNA hybrids. Cold Spring Harb. Lab. Press *28*, 1–10.

Wang, Q., Zou, Y., Nowotschin, S., Kim, S.Y., Li, Q. V., Soh, C.-L., Su, J., Zhang, C., Shu, W., Xi, Q., et al. (2017). The p53 Family Coordinates Wnt and Nodal Inputs in Mesendodermal Differentiation of Embryonic Stem Cells. Cell Stem Cell *20*, 70–86.

Wang, R., An, J., Ji, F., Jiao, H., Sun, H., and Zhou, D. (2008). Hypermethylation of the Keap1 gene in human lung cancer cell lines and lung cancer tissues. Biochem. Biophys. Res. Commun. *373*, 151–154.

Wanrooij, P.H., Engqvist, M.K.M., Forslund, J.M.E., Navarrete, C., Nilsson, A.K., Sedman, J., Wanrooij, S., Clausen, A.R., and Chabes, A. (2017). Ribonucleotides incorporated by the yeast mitochondrial DNA polymerase are not repaired. Proc. Natl. Acad. Sci. U. S. A. *114*, 12466–12471.

Weinmann, R., and Roeder, R.G. (1974). Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. Proc. Natl. Acad. Sci. U. S. A. *71*, 1790–1794.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science *303*, 1014–1016.

Wiltshire, T.D., Lovejoy, C. a., Wang, T., Xia, F., O’Connor, M.J., and Cortez, D. (2010). Sensitivity to poly(ADP-ribose) polymerase (PARP) inhibition identifies ubiquitin-specific peptidase 11 (USP11) as a regulator of DNA double-strand break repair. J. Biol. Chem. *285*, 14565–14571.

Wolberger, C. (2014). Mechanisms for regulating deubiquitinating enzymes. Protein Sci. *23*, 344–353.

Wongsurawat, T., Jenjaroenpun, P., Kwoh, C.K., and Kuznetsov, V. (2012). Quantitative model of R-loop forming structures reveals a novel level of RNA-DNA interactome complexity. Nucleic Acids Res. *40*.

Wu, H.-C., Lin, Y.-C., Liu, C.-H., Chung, H.-C., Wang, Y.-T., Lin, Y.-W., Ma, H.-I., Tu, P.-H., Lawler, S.E., and Chen, R.-H. (2014). USP11 regulates PML stability to control Notch-induced malignancy in brain tumours. Nat. Commun. *5*, 3214.

Xu, L., and Qu, Z. (2012). Roles of protein ubiquitination and degradation kinetics in biological oscillations. PLoS One *7*.

Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., and Miyagawa, K. (2018). Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. Cell 1–13.on. Mol. Cell *42*, 794–805.

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

Yu, M., Liu, K., Mao, Z., Luo, J., Gu, W., and Zhao, W. (2016). USP11 is a Negative Regulator to γH2AX Ubiquitylation by RNF8/RNF168. J. Biol. Chem. *291*, jbc.M114.624478.

Yüce, Ö., and West, S.C. (2013). Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. Mol. Cell. Biol. *33*, 406–417.

Zhang, D., Zaugg, K., Mak, T.W., and Elledge, S.J. (2006). A Role for the Deubiquitinating Enzyme USP28 in Control of the DNA-Damage Response. Cell *126*, 529–542.

Zhang, D.D., Lo, S.-C., Cross, J. V., Templeton, D.J., and Hannink, M. (2004). Keap1 Is a Redox-Regulated Substrate Adaptor Protein for a Cul3-Dependent Ubiquitin Ligase Complex. Mol. Cell. Biol. *24*, 10941–10953.

Zhang, D.D., Lo, S.C., Sun, Z., Habib, G.M., Lieberman, M.W., and Hannink, M. (2005). Ubiquitination of Keap1, a BTB-Kelch substrate adaptor protein for Cul3, targets Keap1 for degradation by a proteasome-independent pathway. J. Biol. Chem. *280*, 30091–30099.

Zhang, S., Xie, C., Li, H., Zhang, K., Li, J., Wang, X., and Yin, Z. (2018). Ubiquitin-specific protease 11 serves as a marker of poor prognosis and promotes metastasis in hepatocellular carcinoma. Lab. Investig. *98*, 883–894.

Zhang, Z., Zhang, Y., Gao, F., Han, S., Cheah, K.S., Tse, H.F., and Lian, Q. (2017). CRISPR/Cas9 Genome-Editing System in Human Stem Cells: Current Status and Future Prospects. Mol. Ther. - Nucleic Acids *9*, 230–241.

Zhang, Z.Z., Pannunzio, N.R., Han, L., Hsieh, C.-L., Yu, K., and Lieber, M.R. (2014). The strength of an Ig switch region is determined by its ability to drive R loop formation and its number of WGCW sites. Cell Rep. *8*, 557–569.

Zhao, D.Y., Gish, G., Braunschweig, U., Li, Y., Ni, Z., Schmitges, F.W., Zhong, G., Liu, K., Li, W., Moffat, J., et al. (2015). SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. Nature *529*, 48–53.

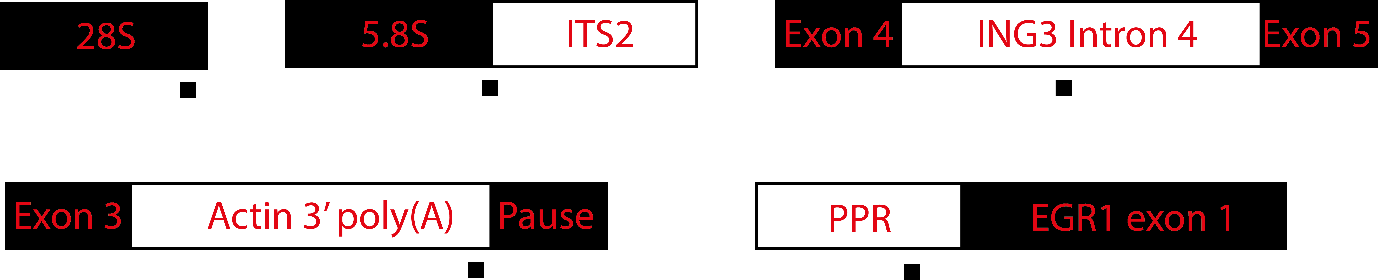
Zhou, H., Di Palma, S., Preisinger, C., Peng, M., Polat, A.N., Heck, A.J.R., and Mohammed, S. (2013). Toward a comprehensive characterization of a human cancer cell phosphoproteome. J. Proteome Res. *12*, 260–271.

Zhou, Z., Luo, A., Shrivastava, I., He, M., Huang, Y., Bahar, I., Liu, Z., and Wan, Y. (2017). Regulation of XIAP Turnover Reveals a Role for USP11 in Promotion of Tumorigenesis. EBioMedicine *15*, 48–61.

# Appendix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Gene** | **NCBI Gene ID** | **Primer** | **Amplicon** | **Location** | **Targeted splice variants** | **Source** | **Comment** |
| ACTB | 60 | Actin β | 121bp | Exon 3 | All | Li et al., 2015 | Reference gene for RT-qPCR |
| USP11 | 8237 | USP11 | 102bp | Exon 4 | All | Zhou et al., 2017 | Target gene for RT-qPCR |
| SETX | 23064 | SETX | 176bp | Exon 26 | All | Skourti-Stathaki et al., 2011 | Target gene for RT-qPCR |
| KEAP1 | 9817 | KEAP1 | 60bp | Exon 3 | All | Hussong et al., 2014 | Target gene for RT-qPCR |
| RNA28SN3 | 109910382 | 28S | 173bp | 3’ end | All | Johnston et al., 2016 | Positive target gene for DRIP-qPCR |
| RNA5-8SN3 | 109910381 | R7 | 55bp | 3’ end of 5.8S and 5’ end of internal transcribed spacer 2 (ITS2) | All | Shen et al., 2017 | Positive target gene for DRIP-qPCR |
| ING3 | 54556 | ING3 | 99bp | Center of intron 4 | All | Halász et al., 2017 | Positive target gene for DRIP-qPCR |
| ACTB | 60 | Actin | 104bp | 5' of the pause site | All | Skourti-Stathaki et al., 2011 | Positive target gene for DRIP-qPCR |
| EGR1 | 1958 | EGR1 | 90bp | Proximal promoter region (PPP). 1.3kbp upstream of the reading frame. | All | Ginno et al., 2012 | Positive target gene for DRIP-qPCR |
| SNRPN | 6638 | SNRPN-neg | 102bp | 5' untranslated region (UTR). 67.2kbp upstream of the reading frame. | All | Ginno et al., 2012 | Negative target gene for DRIP-qPCR |
| MYADM | 91663 | MYADM-neg | 146bp | 5' UTR. 4.4kbp upstream of the reading frame. | All | Ginno et al., 2012 | Negative target gene for DRIP-qPCR |

Appendix-1: Detailed primer data

Schematic of the binding of positive DRIP-qPCR primer pairs is presented below. For MYADM and SNRPN negative DRIP-qPCR primers, please see the table above.

Appendix-2: Figures obtained during my master’s project: 3.2a; 3.4a *left panel*; 3.5b; 3.6b *left panel*

Appendix-3: Processed screen data – knockdowns of the following DUBs triggered R-loop suppression or accumulation in the following conditions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **R-loop suppression** | | | **R-loop accumulation** | | |
| **DMSO** | **CPT 10min** | **CPT 30min** | **DMSO** | **CPT 10min** | **CPT 30min** |
| UEVLD |  | JOSD1 | USP19 | OTUD7 |  |
| OTUB1 |  | USP33 | USP29 | USP2 |  |
| STAMBPL1 |  | USP45 | **USP11** | **USP11** |  |
|  |  | USP3 | USP48 |  |  |