MOLECULAR CHARACTERIZATION OF THE CANCER SUSCEPTIBILITY PROTEIN WRAP53β IN CAJAL BODY FORMATION AND DNA REPAIR

Sofia Henriksson

Stockholm 2013
All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.
Printed by Larserics Digital Print AB.

© Sofia Henriksson, 2013
MOLECULAR CHARACTERIZATION OF THE CANCER SUSCEPTIBILITY PROTEIN WRAP53β IN CAJAL BODY FORMATION AND DNA REPAIR

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Radiumhemmets föreläsningssal, P1:01, Karolinska Universitetssjukhuset, Solna

Fredagen den 29 november, 2013, klockan 09:30

av

Sofia Henriksson

Huvudhandledare:
Docent Marianne Farnebo
Karolinska Institutet
Institutionen för Onkologi-Patologi

Bihandledare:
Professor Klas Wiman
Karolinska Institutet
Institutionen för Onkologi-Patologi

Professor Per Kogner
Karolinska Institutet
Institutionen för kvinnors och barns hälsa

Dr. Salah Mahmoudi
Standford University
Department of Genetics

Fakultetsopponent:
Associate Professor Simon Bekker-Jensen
University of Copenhagen
The Novo Nordisk Foundation Center for Protein Research

Betygsämnd:
Docent Teresa Frisan
Karolinska Institutet
Institutionen för Cell- och Molekylärbiologi

Professor Nico Dantuma
Karolinska Institutet
Institutionen för Cell- och Molekylärbiologi

Professor Neus Visa
Stockholms Universitet
Institutionen för Molekylär Biovetenskap

Stockholm 2013
ABSTRACT

WRAP53β is a multifaceted protein involved in several biological processes including Cajal body maintenance, cancer cell survival and DNA damage repair. By directing factors to Cajal bodies and DNA double-strand breaks (DSBs), WRAP53β facilitates site-specific interactions necessary for proper biological responses.

The Cajal body is a subnuclear organelle implicated in cellular processes such as splicing machinery maturation and telomere maintenance. In Paper I, we reveal that WRAP53β is an essential structural component of Cajal bodies. Furthermore, WRAP53β is required for the intracellular targeting of factors to this site. WRAP53β associates with the survival of motor neuron (SMN) complex in the cytoplasm, mediates its nuclear import and subsequent Cajal body accumulation. In addition, we find that the interaction between WRAP53β and SMN is disrupted in the severe neurodegenerative disorder spinal muscular atrophy, suggesting clinical relevance of WRAP53β-mediated SMN transport.

In Paper II, we study the relationship between WRAP53β expression and cancer cell survival. We demonstrate that WRAP53β is overexpressed in a panel of different cancer cell lines in comparison to primary cells. WRAP53β depletion results in massive induction of cancer cell death, whereas normal human fibroblasts are largely insensitive to WRAP53β knockdown. The cell death associated with WRAP53β silencing occurs via the intrinsic mitochondrial pathway as demonstrated by Bax/Bak activation, loss of mitochondrial membrane potential and release of cytochrome c. Finally, we show that high WRAP53β expression levels correlate with poor prognosis and radioresistance of head and neck cancer patients.

In Paper III, we establish WRAP53β as a novel player in the DNA damage response. We show that WRAP53β rapidly and transiently localizes to DNA DSBs in an ATM- and PARP-dependent manner. WRAP53β binds the E3 ligase RNF8 and facilitates its interaction with MDC1, which is essential for the downstream recruitment of repair proteins 53BP1, BRCA1 and RAD51 to damaged sites. Knockdown of WRAP53β results in deficient DNA DSB repair, whereas WRAP53β overexpression enhances repair efficiency and provides resistance to DNA damaging agents. Furthermore, reduced expression of WRAP53β is related to decreased ovarian cancer patient survival.

In summary, our data identify WRAP53β as a novel structural and regulatory component of Cajal bodies as well as an important factor in carcinogenesis and DNA repair.
LIST OF PUBLICATIONS


*Authors contributed equally

RELATED PUBLICATIONS NOT INCLUDED IN THESIS

# TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS .................................................................................. 7

2. INTRODUCTION ..................................................................................................... 10
   2.1 The Cajal body .................................................................................................. 10
      2.1.1 Coilin ........................................................................................................... 11
      2.1.2 The SMN complex ....................................................................................... 12
      2.1.3 Spinal muscular atrophy .............................................................................. 14
      2.1.4 snRNP biogenesis and splicing .................................................................... 15
      2.1.5 scaRNAs and snoRNAs .............................................................................. 17
      2.1.6 Telomerase .................................................................................................. 19
      2.1.7 Dyskeratosis congenita ............................................................................... 20
   2.2 WRAP53β ......................................................................................................... 21
   2.3 The DNA damage response .............................................................................. 23
      2.3.1 Genomic instability and carcinogenesis ...................................................... 23
      2.3.2 Cancer therapies and DNA damage ............................................................. 24
      2.3.3 DNA double-strand breaks ......................................................................... 25
      2.3.4 DNA damage signaling ................................................................................ 26
      2.3.5 γH2AX ......................................................................................................... 27
      2.3.6 MDC1 .......................................................................................................... 28
      2.3.7 RNF8 and RNF168 ..................................................................................... 28
      2.3.8 Ubiquitin-mediated protein recruitment: 53BP1, BRCA1 and RAD51 ....... 30
   2.4 Cellular responses to DNA damage: survival or death ..................................... 32
      2.4.1 Homologous recombination ....................................................................... 33
      2.4.2 Non-homologous end joining ..................................................................... 35
      2.4.3 Regulation of repair pathway choice ........................................................... 37
      2.4.4 Cell cycle checkpoints ................................................................................ 39
      2.4.5 Apoptosis ................................................................................................... 40

3. AIMS OF THIS THESIS .......................................................................................... 42

4. RESULTS AND DISCUSSION .............................................................................. 43
   4.1 Paper I ............................................................................................................... 43
   4.2 Paper II ............................................................................................................. 45
   4.3 Paper III ............................................................................................................ 47
   4.4 General discussion and concluding remarks ................................................... 50

5. ACKNOWLEDGEMENTS ...................................................................................... 56

6. REFERENCES ......................................................................................................... 58
1. LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>53BP1</td>
<td>Tumor suppressor p53-binding protein 1</td>
</tr>
<tr>
<td>ALC1</td>
<td>Amplified in liver cancer 1</td>
</tr>
<tr>
<td>ALT</td>
<td>Alternative lengthening of telomeres</td>
</tr>
<tr>
<td>Alu</td>
<td>Arthrobacter luteus</td>
</tr>
<tr>
<td>Apaf-1</td>
<td>Apoptotic protease-activating factor 1</td>
</tr>
<tr>
<td>APB</td>
<td>ALT-associated PML bodies</td>
</tr>
<tr>
<td>ASF/SF2</td>
<td>Alternative splicing factor/splicing factor 2</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia telangiectasia mutated</td>
</tr>
<tr>
<td>ATR</td>
<td>Ataxia telangiectasia and Rad3-related</td>
</tr>
<tr>
<td>ATRIP</td>
<td>ATR-interacting protein</td>
</tr>
<tr>
<td>BACH1</td>
<td>BTB and CNC homology 1</td>
</tr>
<tr>
<td>Bak</td>
<td>Bcl-2 homologous antagonist/killer 1</td>
</tr>
<tr>
<td>BAL1</td>
<td>B-aggressive lymphoma 1</td>
</tr>
<tr>
<td>BARD1</td>
<td>BRCA1 associated RING domain 1</td>
</tr>
<tr>
<td>Bax</td>
<td>Bcl-2-associated X protein</td>
</tr>
<tr>
<td>BBAP</td>
<td>B-lymphoma- and BAL-associated protein</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>B-cell lymphoma 2</td>
</tr>
<tr>
<td>BIR</td>
<td>Break-induced replication</td>
</tr>
<tr>
<td>BLM</td>
<td>Bloom syndrome protein</td>
</tr>
<tr>
<td>BMI1</td>
<td>B Lymphoma Mo-MLV Insertion Region 1</td>
</tr>
<tr>
<td>BP</td>
<td>Base pair</td>
</tr>
<tr>
<td>BRCA</td>
<td>Breast cancer susceptibility gene</td>
</tr>
<tr>
<td>BRCC</td>
<td>BRCA1/BRCA2-containing complex subunit</td>
</tr>
<tr>
<td>BRCT</td>
<td>BRCA1 C terminus</td>
</tr>
<tr>
<td>CAB box</td>
<td>Cajal body box</td>
</tr>
<tr>
<td>Caspase</td>
<td>Cysteine aspartic acid specific protease</td>
</tr>
<tr>
<td>CDC</td>
<td>Cell division cycle</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin-dependent kinase</td>
</tr>
<tr>
<td>CHD4</td>
<td>Chromodomain helicase DNA-binding protein 4</td>
</tr>
<tr>
<td>Chk</td>
<td>Checkpoint kinase</td>
</tr>
<tr>
<td>CK2</td>
<td>Casein kinase 2</td>
</tr>
<tr>
<td>CSR</td>
<td>Class switch recombination</td>
</tr>
<tr>
<td>CtIP</td>
<td>CtBP-interacting protein</td>
</tr>
<tr>
<td>DC</td>
<td>Dyskeratosis congenita</td>
</tr>
<tr>
<td>DDR</td>
<td>DNA damage response</td>
</tr>
<tr>
<td>dHJ</td>
<td>double Holliday junction</td>
</tr>
<tr>
<td>DISC</td>
<td>Death-inducing signaling complex</td>
</tr>
<tr>
<td>D-loop</td>
<td>Displacement loop</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNA2</td>
<td>DNA replication ATP-dependent helicase/nuclease 2</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>DNA-dependent protein kinase, catalytic subunit</td>
</tr>
<tr>
<td>DNA pol</td>
<td>DNA polymerase</td>
</tr>
<tr>
<td>DSB</td>
<td>Double-strand break</td>
</tr>
<tr>
<td>DSBR</td>
<td>Double-strand break repair</td>
</tr>
<tr>
<td>DUB</td>
<td>Deubiquitinating enzyme</td>
</tr>
<tr>
<td>EME1</td>
<td>Essential meiotic endonuclease 1</td>
</tr>
<tr>
<td>ERCC1</td>
<td>Excision repair cross-complementing rodent repair deficiency, complementation group 1</td>
</tr>
<tr>
<td>EXO1</td>
<td>Exonuclease 1</td>
</tr>
<tr>
<td>FasL</td>
<td>Fas ligand</td>
</tr>
<tr>
<td>FHA</td>
<td>Forkhead-associated</td>
</tr>
</tbody>
</table>
GAR1  Glycine arginine rich protein 1
Gems  Gemini of Cajal bodies
GEN1  Gen endonuclease homolog 1
GFP  Green fluorescent protein
Gy  Gray
HERC2  HECT domain and RCC1-like domain-containing protein 2
HNSCC  Head and neck squamous cell carcinoma
HR  Homologous recombination
IRIF  Ionizing radiation-induced foci
IR  Ionizing radiation
JMJD2A  Jumonji domain-containing protein 2A
JNK  c-Jun N-terminal kinase
kDa  Kilodalton
L3MBTL1  Lethal (3) malignant brain tumor-like protein 1
LOH  Loss of heterozygosity
LSm  Like Sm
m⁷G cap  7-monomethylguanosine cap
m⁷G cap  2,2,7-trimethylguanosine cap
MDC1  Mediator of DNA damage checkpoint protein 1
MEF  Mouse embryonic fibroblast
MERIT40  Mediator of RAP80 interactions and targeting subunit of 40 kDa
MMC  Mitomycin C
MMSET  Multiple myeloma SET domain-containing protein
MOF  MOZ, YBF2/SAS3, SAS2 and TIP60 protein
MRE11  Meiotic recombination 11
mRNA  Messenger RNA
NBS1  Nijmegen breakage syndrome 1
NHEJ  Non-homologous end joining
NHP  Non-histone chromosome protein
NOP  Nucleolar protein
OTUB1  OTU domain, ubiquitin aldehyde binding 1
PALB2  Partner and localizer of BRCA2
PAR  Poly (ADP-ribose)
PARG  Poly (ADP-ribose) glycohydrolase
PARP  Poly (ADP-ribose) polymerase
PI3K  Phosphatidylinositol 3-kinase
PLA  Proximity ligation assay
PLK1  Polo-like kinase 1
PML  Promyelocytic leukemia
POH1  Pd1 homologue
PPT  Polypyrimidine tract
PSMD4  Proteasome (Prosome, Macropain) 26S subunit, non-ATPase 4
PTEN  Phosphatase and tensin homolog
RAD  Radiation
RanGTP  Ras-related nuclear protein guanosine triphosphate
RAP80  Receptor-associated protein 80
RDS  Radiosensitive DNA synthesis
RG  Arginine-glycine
RIDDLE  Radiosensitivity, immunodeficiency, dysmorphic features and learning difficulties
RIF1  Rap1-interacting factor 1
RING  Really interesting new gene
RMI  RecQ mediated genome instability
RNA  Ribonucleic acid
RNF  Ring finger protein
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>RPA</td>
<td>Replication protein A</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>scaRNA</td>
<td>Small Cajal body-specific RNA</td>
</tr>
<tr>
<td>SCF</td>
<td>Skp, Cullin, F-box</td>
</tr>
<tr>
<td>SDSA</td>
<td>Synthesis-dependent strand annealing</td>
</tr>
<tr>
<td>Sm</td>
<td>Smith</td>
</tr>
<tr>
<td>SMA</td>
<td>Spinal muscular atrophy</td>
</tr>
<tr>
<td>SMARCA5</td>
<td>SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5</td>
</tr>
<tr>
<td>SMN</td>
<td>Survival of motor neuron</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>snoRNA</td>
<td>Small nucleolar RNA</td>
</tr>
<tr>
<td>snRNA</td>
<td>Small nuclear RNA</td>
</tr>
<tr>
<td>SSA</td>
<td>Single-strand annealing</td>
</tr>
<tr>
<td>SSB</td>
<td>Single-strand break</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>SUMO</td>
<td>Small ubiquitin-like modifier</td>
</tr>
<tr>
<td>SWI/SNF</td>
<td>Switch/sucrose nonfermentable</td>
</tr>
<tr>
<td>TCAB1</td>
<td>Telomerase Cajal body protein 1</td>
</tr>
<tr>
<td>hTERT</td>
<td>Human telomerase reverse transcriptase</td>
</tr>
<tr>
<td>TGS1</td>
<td>Trimethylguanosine synthase 1</td>
</tr>
<tr>
<td>TIF</td>
<td>Telomere dysfunction-induced foci</td>
</tr>
<tr>
<td>TIN2</td>
<td>TRF1-interacting nuclear factor 2</td>
</tr>
<tr>
<td>TIP60</td>
<td>60 kDa Tat-interactive protein</td>
</tr>
<tr>
<td>TMG cap</td>
<td>2,2,7-trimethylguanosine cap</td>
</tr>
<tr>
<td>TopBP1</td>
<td>DNA topoisomerase 2-binding protein 1</td>
</tr>
<tr>
<td>TOPO</td>
<td>Topoisomerase</td>
</tr>
<tr>
<td>hTR</td>
<td>Human telomerase RNA</td>
</tr>
<tr>
<td>TRAIL</td>
<td>TNF-related apoptosis-inducing ligand</td>
</tr>
<tr>
<td>β-TrCP</td>
<td>Beta-transducin repeat containing protein</td>
</tr>
<tr>
<td>TRF</td>
<td>Telomere repeat binding factor</td>
</tr>
<tr>
<td>TRIP12</td>
<td>Thyroid receptor-interacting protein 12</td>
</tr>
<tr>
<td>UBC</td>
<td>E2 ubiquitin-conjugating protein</td>
</tr>
<tr>
<td>UBR5</td>
<td>Ubiquitin protein ligase E3 component n-recognition 5</td>
</tr>
<tr>
<td>Unrip</td>
<td>Unr-interacting protein</td>
</tr>
<tr>
<td>USP</td>
<td>Ubiquitin-specific protease</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VCP</td>
<td>Valosin-containing protein</td>
</tr>
<tr>
<td>V(D)J</td>
<td>Variable diversity joining</td>
</tr>
<tr>
<td>WDR79</td>
<td>WD repeat-containing protein 79</td>
</tr>
<tr>
<td>WRAP53</td>
<td>WD40-encoding RNA antisense to p53</td>
</tr>
<tr>
<td>XLF</td>
<td>XRCC4-like factor</td>
</tr>
<tr>
<td>XPF</td>
<td>Xeroderma pigmentosum group F-complementing protein</td>
</tr>
<tr>
<td>XRCC</td>
<td>X-ray repair cross-complementing protein</td>
</tr>
<tr>
<td>ZPR1</td>
<td>Zinc-finger protein 1</td>
</tr>
</tbody>
</table>
2. INTRODUCTION

2.1 The Cajal body

The cell nucleus is a highly organized organelle that contains several sub-compartments serving to concentrate specific factors and biological processes into restricted spaces thereby optimizing their performance. Numerous distinct non-membrane-bound nuclear bodies have been identified, including structures such as nucleoli, nuclear speckles, histone locus bodies, PML bodies, Gems and Cajal bodies. The size and number of nuclear bodies depends on cell type, cell cycle stage and cellular demands. Although existing as separate nuclear entities, there are overlapping components and organizational properties between different nuclear bodies. The formation of these sub-organelles usually depends on dynamic processes such as protein modifications, protein self-association, different RNA and protein interactions and association with specific gene loci leading to tethering of central factors (1, 2).

In 1903, Santiago Ramón y Cajal identified Cajal bodies in the nuclei of vertebrate neurons and described them as “nucleolar accessory bodies” due to their close proximity to the nucleolus in these cells (3). The Cajal body is a 0.2-2 µM spherical organelle ranging in numbers of 1-10 per cell nucleus and is characterized by the presence of the marker protein coilin (Figure 1A). Cajal bodies are found across many species, including plants, yeast, insects and mammals (4). Cajal body numbers vary over the cell cycle, reaching a maximum at G1/S followed by their disassembly during mitosis (5, 6). Early in G1, coilin is diffusely distributed throughout the nucleoplasm and there is a lag period before mature Cajal bodies reappear. This phenomenon might be regulated through hyperphosphorylation of coilin in mitosis or indirectly via mitotic repression of transcription, a process intimately linked to the Cajal body (7-9). Cajal body numbers also correlate with metabolic and proliferative activity and are most abundant in cells with high transcription and splicing rates, as observed in neuronal and cancer cells (2, 9-12). Certain cell types of adult tissues such as spleen and blood vessels are devoid of Cajal bodies, whereas they are observed in all fetal tissues (13, 14).

Besides coilin, Cajal bodies contain a multitude of other molecules. Notably, several different proteins and RNA species in the form of ribonucleoprotein (RNP) complexes are enriched in Cajal bodies. These include the spliceosomal small nuclear RNPs (snRNPs), the small nucleolar RNPs (snoRNPs) and components of the telomerase RNP complex. Cajal bodies also contain the Cajal body-specific RNPs (scaRNPs) that guide the post-transcriptional modifications of snRNAs, required for their complete maturation and incorporation into the spliceosome. Accordingly, Cajal bodies have been described to play essential roles in snRNP and snoRNP maturation and telomere maintenance (Table 1). Furthermore, the snRNP-assembling survival of motor neuron (SMN) complex and factors involved in histone mRNA 3’-end processing accumulate in the Cajal body (2, 5). These factors and processes will be described in greater detail in upcoming sections.

Cajal bodies are highly dynamic structures that move within the nucleoplasm, shuttle to and from the nucleoli, fuse and undergo fission events leading to variations in size and numbers (15). Since Cajal bodies lack surrounding membranes, components
readily exchange and diffuse to the surrounding nucleoplasm. Studies analyzing the kinetics of Cajal body components revealed turnover rates of seconds to a few minutes, with coilin and members of the SMN complex demonstrating the longest residence times (16). However, a subset of Cajal bodies can also be immobilized by interacting with specific gene loci such as snRNA, snoRNA and histone gene clusters in a transcription-dependent manner (17-19). Therefore, Cajal bodies seem to alternate between a mobile state of passive diffusion within the nucleoplasm and an immobile state of transient chromatin association. Cajal bodies may associate with specific genes in order to facilitate or regulate gene expression by providing pre-assembled RNP complexes to sites of transcription (20). Additionally, Cajal bodies may facilitate nuclear export of newly transcribed snRNAs, which is a vital step in their maturation pathway (21).

Tethering experiments have revealed that immobilization of a variety of Cajal body components leads to the de novo formation of Cajal bodies. These results demonstrate that Cajal body formation does not absolutely depend on a specific locus and there is no specific factor that initiates the assembly process (22). However, others have suggested that RNA initiates Cajal body formation. In this model, RNA functions as a scaffold for the immobilization and retention of Cajal body-associated factors resulting in nuclear body assembly (23). This is in line with data showing that transcriptional inhibition results in Cajal body disassembly and furthermore prevents Cajal body re-formation following completion of mitosis (6, 24). Nevertheless, Cajal bodies seem to assemble without internal hierarchy of individual components, in a random and self-organized manner, either in the presence or absence of a RNA-mediated seeding event.

<table>
<thead>
<tr>
<th>RNA component</th>
<th>Protein components</th>
<th>RNP complex</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>snRNA</td>
<td>Sm proteins</td>
<td>snRNP</td>
<td>Splicing</td>
</tr>
<tr>
<td>scaRNA</td>
<td>Fibrillarin/Dyskerin</td>
<td>scaRNP</td>
<td>snRNA modification</td>
</tr>
<tr>
<td>snoRNA</td>
<td>Fibrillarin/Dyskerin</td>
<td>snoRNP</td>
<td>rRNA modification</td>
</tr>
<tr>
<td>hTR</td>
<td>hTERT</td>
<td>Telomerase RNP</td>
<td>Telomere elongation</td>
</tr>
</tbody>
</table>

Table 1: A simplified overview of Cajal body-associated RNP complexes. The indicated protein components represent core RNP proteins, however additional proteins (not included in the table) are also present in the RNP complexes.

2.1.1 Coilin

Cajal bodies had been recognized for almost a century when coilin was identified as a constitutive component of this nuclear organelle. By using human autoimmune sera, distinct Cajal bodies could be visualized and immunoblotting analysis revealed the presence of a 80 kDa protein that was named p80-coilin (25-27). Although coilin is an essential factor for Cajal body maintenance, coilin levels are not rate-limiting for Cajal body numbers or size. Transient overexpression of coilin does not lead to the formation of additional Cajal bodies. Instead, exogenous coilin mainly localizes throughout the nucleoplasm and appears to disrupt Cajal bodies (28). The majority of endogenous coilin resides in the nucleoplasm as well (29). Instead of altered protein levels, post-translational modifications such as methylation and phosphorylation of coilin can influence Cajal body formation through alterations in self-association properties and protein binding. The self-interaction domain of coilin resides in its N-
terminus and is required for targeting coilin to Cajal bodies (28). Coilin hyper-phosphorylation reduces its self-interaction potential, which is observed in mitosis when Cajal bodies are disrupted and in primary cells lacking Cajal bodies (7, 12, 28). In addition, the C-terminal region of coilin contains an arginine- and glycine-rich region (RG box) and an atypical Tudor domain. The classical Tudor domain binds methylated arginine and lysine residues on target proteins, however the atypical Tudor domain of coilin is not believed to retain this function due to structural aberrations (30). The arginines of the coilin RG box can be symmetrically dimethylated, which promotes its interaction with SMN (31). Coilin interacts both with SMN and snRNPs through its C-terminus, although via separate regions (32).

Knockout studies of coilin in several different species have been performed to uncover the functional importance of Cajal bodies. Since the Cajal body has been indicated as the site of snRNP maturation, deficiencies in this pathway is expected to result in splicing defects (33-36). Coilin knockout in Arabidopsis and Drosophila revealed no apparent phenotype. The levels of splicing appeared to be unaffected even though Cajal bodies were absent, suggesting that splicing machinery maturation also can occur by Cajal body-independent means in the nucleoplasm (37, 38). In mice, homozygous loss of coilin results in partial embryonic lethality. The surviving fraction of coilin−/− mice display defective fertility and fecundity (39). Cells derived from the knockout mice lack canonical Cajal bodies and instead form so-called residual Cajal bodies with distinct constituents. Although unable to accumulate proteins of the SMN complex, one subtype of residual Cajal bodies are enriched in snRNAs and the guide RNAs (scaRNAs) responsible for their modification, suggesting that snRNP maturation can occur in the absence of canonical Cajal bodies. Accordingly, coilin knockout MEFs showed no defect in the post-transcriptional modification of snRNAs. Nevertheless, these cells were later identified to display impaired artificial reporter splicing (40-42). Knockdown of coilin in zebrafish is lethal and associated with defective snRNP assembly and pre-mRNA splicing. Addition of human snRNPs can partially rescue the phenotype of coilin-deficient zebrafish, suggesting that the developing zebrafish embryo is particularly sensitive to alterations in snRNP biogenesis (43). Human HeLa cells devoid of coilin are viable but display reduced reporter splicing and decreased proliferation rates (42, 44).

The relatively mild phenotypes observed in most species upon coilin knockout indicate that the Cajal body is not an essential organelle per se. Given the fundamental biological importance of pre-mRNA splicing, any gross splicing defect would lead to a severe phenotype. Accordingly, disruption of the snRNP biogenesis genes SMN1, Gemin2 and ZPR1 in mice all results in embryonic lethality (45-47). Moreover, many adult cell types lack Cajal bodies (14). Instead of being essential for the splicing process, the Cajal body has been suggested to increase the efficiency of snRNP biogenesis by bringing the diverse components of the Cajal body together. Indeed, it has been estimated that snRNP assembly occurs 10-fold faster in the Cajal body compared to the surrounding nucleoplasm. Cajal bodies might therefore allow cells to efficiently adjust to fluctuating splicing demands (48).

2.1.2 The SMN complex

The SMN protein is produced from the disease-causing gene of the neurodegenerative disorder spinal muscular atrophy (SMA) (49). SMN is ubiquitously expressed and
essential for efficient assembly of splicing RNP complexes. The SMN protein contains two self-association domains and a Tudor domain that can interact directly with RG box-containing partners such as coilin and Sm proteins. Furthermore, exon 7 of SMN contains a region that confers protein stability (50, 51). Deletion of the self-association or Tudor domains disrupts SMN targeting to Cajal bodies (52). Knockdown of SMN disrupts canonical Cajal bodies, suggesting that snRNP assembly and import is essential for the integrity of this organelle (44).

SMN functions in a large macromolecular complex that, besides SMN, is composed of seven proteins termed Gemin2-8 and Unrip (53, 54). SMN, Gemin7 and Gemin8 appear to act as an interaction platform that brings the SMN complex together via multiple interactions. There have been conflicting reports regarding the contribution of specific Gemin proteins to the function of the SMN complex, although SMN protein carrying a SMA-related mutation displayed severe impairment of SMN complex formation (55). Silencing of SMN leads to reduced expression of several Gemin5, whereas depletion of Gemin2 and Gemin6 results in disrupted Sm core assembly (56). Others have shown that Gemin3 and Gemin4 are important for the maintenance of snRNP assembly-activity (57). Moreover, the WD40-domain of Gemin5 has been pinpointed to be responsible for recognizing and binding snRNAs (58). The most well studied function of the SMN complex is the cytoplasmic assembly of snRNPs. The SMN complex promotes the binding between snRNAs and Sm proteins, thereby forming the snRNP core particle that mediates pre-mRNA splicing. The SMN complex also provides sequence specificity towards snRNAs, thus preventing non-specific binding of Sm proteins to other RNAs (59). Cells lacking SMN display defective snRNP core assembly and splicing (59-62).

Figure 1: (A) Cajal bodies visualized by coilin immunostaining. Dashed white circle marks the nucleus and white arrows indicate Cajal bodies. (B) Cajal bodies and Gems visualized by SMN immunostaining. White arrows indicate Cajal bodies, whereas yellow arrows highlight two Gems. All small nuclear bodies represent Gems. Cajal bodies were distinguished from Gems by coilin co-immunostaining.

Besides being enriched in Cajal bodies, SMN localizes to the cytoplasm and to an additional nuclear body termed Gemini of Cajal bodies (Gems) (Figure 1B). As their name implies, Gems are closely associated with Cajal bodies and concentrate SMN and Gemin proteins but are not enriched in snRNPs (63). Cajal bodies and Gems exist as separate structures in fetal tissues but coincide in adult tissues (13, 64). However, in some cultured cells, the separate organizational pattern of Gems and Cajal bodies is
maintained and appears to be regulated via coilin methylation. Hypomethylated coilin promotes the accumulation of SMN in Gems, which coincides with reduced coilin-SMN interaction (31, 63). The precise function of Gems during development remains unknown, although it has been suggested that cells with both Cajal bodies and Gems have the highest snRNP assembly rates and splicing capacities (42).

2.1.3 Spinal muscular atrophy

Spinal muscular atrophy is an autosomal recessive disorder characterized by the progressive degeneration of spinal cord anterior horn α-motor neurons. SMA is the leading genetic cause of infant mortality and affects around 1:6000 live births. The gradual neurodegeneration results in muscular atrophy and in the severe cases respiratory complications lead to death. SMA is classified into four clinical forms, SMA type I-IV, based on age of onset and disease severity (51). The severity of disease correlates with SMN protein levels, where SMA type I (or Werdnig-Hoffmann disease) is the most severe form with a life expectancy of less than 2 years (66, 67).

The genetic background of the different SMA subtypes explains the variations in SMN proteins levels. In humans, the SMN gene is present in two copies as an inverted repeat on chromosome 5q13. All SMA patients carry deletions or mutations of the telomeric SMN1 gene, whereas at least one copy of the centromeric SMN2 gene remains intact (68). The SMN2 gene is nearly identical to SMN1, however a C-to-T transition in exon 7 disrupts an exonic splicing enhancer and results in the production of a truncated SMN protein termed SMNΔ7 (69, 70). Although not dominant negative, SMNΔ7 display oligomerization defects and is highly unstable. Nevertheless, SMN2 retains the capacity to produce approximately 10% of full-length SMN protein, which appears to be sufficient for most cell types except motor neurons (71, 72). SMN1-to-SMN2 gene conversion events result in increased copy numbers of SMN2, which is associated with milder forms of SMA and prolonged survival (73, 74). Carriers and control persons with homozygous deletions of SMN2 show no clinical manifestations, supporting the idea that SMN1 is the main producer of full-length SMN protein (75). Mice only have one SMN gene and homozygous SMN disruption results in embryonic lethality. Introduction of human SMN2 to SMN knockout mice rescues the lethality and instead leads to the development of SMA, demonstrating that low SMN protein levels triggers the disease (47, 76-78). SMA type I motor neurons demonstrate defective targeting of SMN and coilin to Cajal bodies and a severe decrease in Cajal body numbers (79).

Given the fundamental importance of splicing, the underlying reason why SMA affects lower motor neurons to such a high extent remains unexplained. It has been suggested that motor neurons have higher energy demands and might express SMN at the levels required whereas other cell types express an excess of SMN (80, 81). Animal and cell culture models of SMA have revealed a correlation between reduced capacity of snRNP assembly and disease severity in the form of motor neuron degeneration (46, 82-84). However, in a Drosophila SMA model, the role of SMN in snRNP biogenesis was independent from its effect on viability (85). Moreover, SMN deficiency causes aberrant splicing that is not restricted to motor neurons. Therefore, the motor neuron-specific effects of SMA cannot be exclusively explained by deficient splicing, although it has been suggested that specific transcripts required for
motor neuron activity are affected in SMA (62, 82). SMN is also proposed to function in the axonal transport of β-actin mRNA in motor neurons, important for local protein synthesis and axonal sprouting (86, 87).

2.1.4 snRNP biogenesis and splicing

The biogenesis of splicing snRNPs is a complex series of events that includes both nuclear and cytoplasmic phases. Briefly, this process includes three major steps: (1) nuclear transcription of the snRNA, (2) cytoplasmic snRNA export followed by assembly of the core snRNP particle, and (3) nuclear re-entry and final maturation of the snRNP complex. Upon completion of the maturation pathway, the snRNP particles participate in pre-mRNA splicing as elementary units of the spliceosome (88-90). A simplified scheme of snRNP biogenesis is depicted in Figure 2.

There are five spliceosomal snRNAs, named U1, U2, U4, U5 and U6 due to their high uridine content. All snRNAs except U6 are transcribed by RNA polymerase II and acquire a 7-monomethylguanosine (m7G) cap at their 5’ end. Nuclear export proteins interact with the m7G cap structure, which serves as a nuclear export signal for snRNA precursors (Step 1 and 2, Figure 2) (91). In contrast, the U6 snRNA is transcribed by RNA polymerase III and acquires a different cap structure. The U6 snRNA is not exported into the cytoplasm and acquires its associated core proteins in a different manner (92).

In the cytoplasm, the SMN complex promotes the assembly of the seven Sm proteins (SmB/B’, D1, D2, D3, E, F and G) in the form of a heptameric ring onto the snRNA, thus forming the core snRNP particle (Step 3, Figure 2) (60, 93). The Sm ring assembles on a conserved structural motif present in snRNAs called the Sm site (94). Although snRNPs can assemble spontaneously in vitro, the SMN complex appears to be required for their assembly in vivo (95). To achieve this, SMN complex components first associate with Sm proteins and then with the snRNA (59, 96). Once the snRNP is assembled, the snRNA m7G cap is hypermethylated by the methyltransferase TGS1 into a 2,2,7-trimethylguanosine (m7G or TMG) cap and the 3’ end is exonucleolytically processed (Step 4, Figure 2) (97, 98). The TMG cap and the Sm proteins form a bipartite snRNP nuclear localization signal (99, 100). The nuclear import adaptor snurportin1 interacts with the TMG cap structure and mediates binding to the nuclear import receptor importinβ (101). The Sm core-dependent snRNP import is suggested to occur via the SMN complex or an unknown adaptor that directly interacts with importinβ (Step 5 and 6, Figure 2) (102, 103).

Once re-imported into the nucleus through the nuclear pore complex, the import proteins disassociate and the snRNPs are targeted to Cajal bodies together with the SMN complex (104-106). Coilin has been shown to associate with both SmB and SMN. Moreover, coilin appears to compete with SMN for binding SmB. Coilin might therefore function to release the snRNP cargo from the SMN complex (107). The Cajal body contains an additional class of RNAs termed scaRNAs. These RNAs guide post-transcriptional modifications of snRNAs, essential for the final maturation of snRNPs and their proper incorporation into the spliceosome (33-36). By associating with specific enzymes, scaRNAs direct site specific 2′-O′-methylation and pseudouridylation of snRNAs via sequence complementarity (Step 7, Figure 2) (40, 108, 109). As mentioned previously, assembly of the U6 snRNP is believed to occur in the nucleus. The U6 snRNA lacks an Sm site and instead associates with seven Sm-
like proteins (LSm2-8) \((110)\). In addition to the common core proteins there are also snRNP-specific proteins, which might be added in the Cajal body \((111)\). Finally, some of the snRNPs are assembled into higher order particles, where the U4, U5 and U6 snRNPs form a functional tri-snRNP. First, base paring between U4 and U6 snRNAs forms a di-snRNP, which is assembled into a tri-snRNP upon incorporation of the U5 snRNP. Di- and tri-snRNP formation is suggested to take place within the Cajal body \((112-114)\). Following modifications and assembly in the Cajal body, snRNP particles accumulate in splicing speckles that serve as snRNP storage sites before and/or after splicing \((\text{Step 8, Figure 2})\) \((105, 115)\).

![Figure 2: Cellular biogenesis of snRNPs. 1) Nuclear transcription of pre-snRNA, 2) Pre-snRNA nuclear export, 3) SMN-mediated snRNP assembly, 4) 5’ cap hypermethylation, 5) Import complex formation, 6) snRNP nuclear import and localization to Cajal bodies, 7) scaRNP-mediated snRNP modification, 8) Mature snRNP particles exit the Cajal body and participate in splicing.](image)

The spliceosome is a complex macromolecular machinery composed of the five spliceosomal snRNPs and a multitude of additional proteins serving to remove non-coding introns from pre-mRNAs to generate mature mRNAs that can be translated into proteins. Introns are defined by specific sequence elements, which include the 5’ and 3’ splice sites, the branch point sequence and a polypyrimidine tract (PPT). The 5’ splice site is characterized by a GT dinucleotide, whereas the 3’ intron boundary contains AG nucleotides and the PPT. The spliceosome assembles in a stepwise and sequential manner on pre-mRNA and catalyzes two successive transesterification reactions, where phosphodiester bonds are cleaved by hydrolysis for subsequent intron removal. Upon completion of splicing, the spliceosome is released and components are recycled to participate in additional splicing reactions \((116)\). Besides
the major spliceosome described above, there is also a minor spliceosome that removes a rare group of introns using distinct snRNP.s. These non-canonical introns display AT and AC terminal nucleotides at their respective 5' and 3' splice sites, instead of the conventional GT and AG. Moreover, the low-abundance snRNP.s U11, U12 and U4atac/U6atac are utilized as alternatives to U1, U2 and U4/U6. The U5 snRNP participates in splicing reactions performed by both the major and minor spliceosomes. The low-abundance snRNAs also associate with Sm cores and SMN has been implied in the assembly of the minor snRNP.s as well (117-120).

Interestingly, the Cajal body appears to be involved in all nuclear steps of the snRNP maturation pathway. As mentioned above, Cajal bodies associate with snRNA gene loci and seem to accumulate pre-snRNA (17, 121). Moreover, Cajal bodies have been demonstrated to enable nuclear export of snRNA precursors by facilitating their binding to RNA export factors (21). Following nuclear re-entry, Cajal bodies ensure proper snRNA base modification by accumulating the mediating guide scaRNAs (108). Cajal bodies have also been implicated as the site where certain snRNP-specific proteins are added to the core snRNP.s and in the formation of di- and tri-snRNP.s. Finally, the Cajal body is involved in snRNP recycling following splicing, for instance in the reassembly of the tri-snRNP (48, 90, 122). Furthermore, Cajal body integrity is closely linked to proper snRNP biogenesis. Knockdown of factors involved in snRNP biogenesis prior to their re-entry into the nucleus disrupt Cajal bodies (44). Inhibition of nuclear export and silencing of the import protein snurportin1 also results in loss of Cajal bodies (5, 57, 101, 123, 124). Whereas coilin overexpression does not enhance Cajal body formation, increased expression of snRNPs and SMN induce Cajal bodies even in cell normally lacking them (28, 123, 125).

2.1.5 scaRNAs and snoRNAs

The concept of the Cajal body as the cellular site of spliceosomal snRNA post-transcriptional modification came from the identification of the Cajal body-specific RNAs (108). The structurally and functionally related snoRNAs were already established as guiding factors for the 2'-O-methylation and pseudouridylation of ribosomal RNA, which is essential for proper ribosome function (109, 126, 127). Two major classes of scaRNAs exist with specific sequence elements, structural motifs and guide-specificity. The so-called C/D box scaRNAs carry the C (RUGAUGA) and D (CUGA) box motifs and directs methylation of target RNAs. The H/ACA scaRNAs contain the H (ANANNA) and ACA boxes and promote the isomerization of uridine into pseudouridine (Figure 3) (128). In contrast to snoRNAs that usually contain one of these domains, scaRNAs often consist of two domains that can be two C/D box, two H/ACA box or a mixed domain carrying both types. Analogous to snoRNAs, the scaRNAs contain the sequence complementarity to the substrate RNA but lack catalytic activity. Instead, the C/D and H/ACA core motifs direct the binding of associated proteins that provide the enzymatic capacity. The C/D box scaRNAs associate with the methyltransferase fibrillarin and three additional proteins termed NOP56, NOP58 and 15.5K/NHPX. The H/ACA scaRNAs associate with GAR1, NHP2, NOP10 and the pseudouridine synthase dyskerin. In the form of scaRNPs and snoRNPs, the RNA components direct the enzymes to their target RNA (Figure 3) (127, 128). The exact function of the snRNA post-transcriptional modifications is not fully understood. The modified sites are generally conserved, cluster in functionally
important regions of the target RNAs and might be required for proper RNA folding \((109, 129, 130)\). For instance, pseudouridylation of the U2 snRNA stabilizes its interaction with intronic pre-mRNA and induces a structural change that promotes the first nucleophilic attack of the splicing reaction. U2 snRNAs lacking pseudouridylation and 2'-O'-methylation sites have been suggested to display defective assembly of the functional U2 snRNP particle and/or defective assembly of whole spliceosomal complexes, resulting in deficient splicing \((33-36, 131)\).

**Figure 3**: Simplified figure of box H/ACA scaRNA (A) and box C/D scaRNA (B). Pseudouridylation is marked by \(\psi\), whereas 2'-O'-methylation is indicated by Me. The scaRNA forms base pairs with the complementary region found in the target snRNA (blue). The C and D boxes are found near the 5' and 3' ends of the molecule, respectively. C' and D' represent a second pair of boxes that display lower conservation and are found in the middle of the molecule. Black boxes indicate associated proteins, where the enzyme is marked in bold. WRAP53\(\beta\) associates with the scaRNAs and mediates their Cajal body localization.

ScaRNAs accumulate in Cajal bodies due to the presence of a common RNA element called the CAB box, which functions as a Cajal body localization signal. In H/ACA scaRNAs, the CAB box is four nucleotides long and present either on the 5' or 3' hairpin terminal loops. The tetranucleotide CAB box consensus sequence is uGAG, where the AG dinucleotide is highly conserved and probably found in all H/ACA and mixed domain C/D-H/ACA scaRNAs (Figure 3) \((132)\). Interestingly, the RNA
component of the telomerase enzyme (hTR) is a CAB box-containing H/ACA scaRNA that associates with the scaRNP core proteins and localizes to Cajal bodies (133-135). A CAB-like box has also been identified in Drosophila C/D scaRNAs, consisting of 10 conserved nucleotides (cgaGUAnUg). However, the CAB box sequence of human C/D scaRNAs has not yet been identified.

For a long time, the factor(s) responsible for the Cajal body-specific localization of CAB box-containing scaRNAs remained unidentified. WRAP53β (also denoted TCAB1 or WDR79) was later identified to specifically associate with scaRNAs and promote their Cajal body targeting. WRAP53β associates with human H/ACA (including hTR), C/D and mixed domain C/D-H/ACA scaRNAs (Figure 3). WRAP53β binding to H/ACA scaRNAs requires intact CAB and ACA box sequences, since combined CAB and ACA box mutations drastically reduces WRAP53β binding to these scaRNAs. WRAP53β is suggested to interact with the scaRNA CAB box sequence directly, whereas its association with the ACA box probably requires an additional protein. Interestingly, dyskerin interacts directly with the conserved ACA element and has also been found to interact with WRAP53β in a RNA-independent manner, proposing that WRAP53β binds directly to the CAB box and indirectly to the ACA motif via dyskerin. WRAP53β also associates specifically with the CAB box of Drosophila C/D scaRNAs. Furthermore, scaRNA CAB box mutants unable to bind WRAP53β mislocalize to the nucleoli. In line with this, WRAP53β depletion also results in nucleolar targeting of scaRNAs (136-138).

2.1.6 Telomerase

The telomerase holoenzyme is a RNP complex that synthesizes the addition of TTAGGG telomeric repeats onto the ends of linear chromosomes. The tandem telomeric repeats range from 2-15 kilobases in length and terminates in a 3’ overhang of approximately 200 bases. The minimal catalytic unit of the telomerase enzyme is composed of the internal RNA template (hTR) that contains the sequence copied by the reverse transcriptase (hTERT). Telomerase activity is regulated over human development, with high activity during embryogenesis that is silenced in most adult somatic tissues. However, certain adult stem cell compartments maintain telomerase activity (139-142).

Primary human cells have a finite lifespan of 60-80 population doublings that is controlled by telomere lengths. Due to the end replication problem, telomeres loose about 100-200 bp during each cell division cycle. The progressive telomere shortening eventually results in replicative senescence followed by apoptosis in order to prevent genomic instability (140, 143). To circumvent this problem, cancer cells reactivate telomerase and thus become immortalized. Strikingly, telomere activation has been observed in approximately 90% of all cancers (144-146). A recombination-based telomerase-independent mode of telomere maintenance is also recognized, which is termed alternative lengthening of telomeres (ALT) (147). Exactly how telomerase activity is regulated remains incompletely understood. Most studies have focused on the transcriptional regulation of hTERT, where c-Myc is identified as one hTERT-inducing transcription factor. However, telomerase activity is most likely subjected to multiple levels of regulation, both via hTERT and other components of the telomerase enzyme (148, 149).
Besides the core components hTR and hTERT, telomerase also requires additional factors for its assembly and function. The H/ACA RNP proteins dyskerin, GAR1, NHP2 and NOP10 all associate with telomerase (139). The H/ACA motif of hTR is required for its stability, nuclear localization and RNP assembly, however due to structural deviations and lack of complementary target RNAs it is not believed to direct pseudouridylation (150-152). WRAP53β associates with the enzymatically active telomerase complex and promotes its localization to Cajal bodies by binding to the hTR CAB box (137, 153). The Cajal body does not appear to promote telomerase enzyme assembly, since hTR CAB box-mutants unable to accumulate in Cajal bodies still forms the catalytically active telomerase complex (154). However, Cajal bodies associate with telomeres and are suggested to deliver the telomerase enzyme for telomere elongation during S-phase (155, 156). The interaction of Cajal bodies with telomeres is dependent on telomerase activity, since Cajal bodies do not associate with ALT telomeres (157).

2.1.7 Dyskeratosis congenita

Dyskeratosis congenita (DC) is an inherited multisystem syndrome characterized by bone marrow failure, premature ageing, cancer predisposition and a triad of mucocutaneous features including oral leukoplakia, abnormal skin pigmentation and nail dystrophy. DC displays X-linked, autosomal recessive and autosomal dominant inheritance patterns depending on the disease-causing gene. Several DC genes have been identified, all encoding core components of the telomerase enzyme or telomere capping proteins such as hTR, hTERT, dyskerin, NHP2, NOP10, WRAP53β and TIN2 (158, 159). Therefore, DC is considered to be a disease of defective telomere maintenance and patients display very short telomeres (160, 161). DC has also been described as a stem cell disease, where insufficient telomerase activity impairs stem cell function and limits tissue renewal (162, 163). Mutations in dyskerin cause X-linked DC, which is the most common form of the disease affecting approximately 30% of reported cases (164, 165).

Generally, patients carrying hTERT and hTR mutations present symptoms later in life and have a milder form of the disease compared to DC caused by mutations in dyskerin, NHP2, NOP10, TIN2 and WRAP53β (138, 166, 167). Interestingly, disease severity is not strictly correlated to telomere lengths and DC-associated hTERT mutations have been demonstrated to maintain telomerase activity (167, 168). In addition to telomere dysfunction, additional perturbations might therefore explain the etiology of DC. Consistent with the well-established role of dyskerin in the pseudouridylation pathway, studies of dyskerin-defective mice and zebrafish have revealed rRNA processing defects that also have been observed in DC patient cells (169-172).

In terms of malignant diseases, DC patients most commonly develop myelodysplastic syndrome, acute myeloid leukemia and squamous cell carcinoma of the head and neck. Patients with DC caused by hTR and hTERT mutations are most prone to develop cancer, which is most likely explained by their longer life expectancy (166, 173). The overall increase in cancer incidence of DC patients is estimated to be 11-fold (174).
2.2 WRAP53β

As the name indicates, WD40-encoding RNA antisense to \( p53 \), or \( WRAP53 \), is a gene with dual functions. First, it encodes a regulatory antisense transcript that positively regulates \( p53 \). Second, via alternative transcriptional start site usage, the \( WRAP53 \) gene also gives rise to a protein product belonging to the WD40-repeat containing protein family. These two gene products are now referred to as \( WRAP53\alpha \) and \( WRAP53\beta \), respectively, to emphasize their functional separation (175).

The \( WRAP53 \) gene is located on chromosome 17p13, where it partially overlaps the \( p53 \) gene in a head-to-head orientation. Due to the presence of different starting exons within the \( WRAP53 \) gene, only exon 1\( \alpha \)-containing \( WRAP53\alpha \) transcripts overlap \( p53 \). In contrast, the protein-encoding transcripts originating from the downstream exon 1\( \beta \) lack the region complementary to \( p53 \) (Figure 4A). \( WRAP53\alpha \) transcripts bind to the 5’ UTR of \( p53 \) via the perfectly complementary region, thereby regulating \( p53 \) mRNA stability. \( WRAP53\alpha \) is required for \( p53 \) induction upon DNA damage and \( p53 \)-dependent apoptosis. Knockdown of \( WRAP53\alpha \) transcripts does not affect \( WRAP53\beta \) protein levels, further supporting the distinction between these two transcripts (175).

**Figure 4:** Schematic picture of the \( WRAP53 \) gene locus (A) and the \( WRAP53\beta \) protein (B).

Most studies of the \( WRAP53 \) gene have focused on the function of the \( WRAP53\beta \) protein. Besides the centrally located WD40-domain, \( WRAP53\beta \) also contains a N-terminal proline-rich region and a C-terminal glycine-rich region (Figure 4B). WD40-domain containing proteins are highly abundant in eukaryotes and the WD40-domain is suggested to function as a scaffold that mediates protein and RNA interactions in large molecular complexes (176, 177).

\( WRAP53\beta \) has been identified as a component of the telomerase holoenzyme, required for telomere elongation. \( WRAP53\beta \) associates with all components of the telomerase enzyme, including hTR, hTERT and dyskerin. The interaction between
WRAP53β and hTERT is RNase A-sensitive, suggesting that the CAB box-containing hTR mediates this interaction. Conversely, WRAP53β probably interacts with dyskerin via protein-protein contacts, since this interaction is insensitive to RNase treatment. The combined observations that WRAP53β co-precipitates with nearly all cellular telomerase activity and that WRAP53β does not associate with telomerase assembly factors indicate that WRAP53β is a component of the active telomerase enzyme. Knockdown of WRAP53β disrupts hTR Cajal body targeting and furthermore reduces hTR localization to telomeres. This is not due to altered hTR stability, since WRAP53β depletion does not affect hTR levels. WRAP53β silencing inhibits telomere elongation promoted by hTR overexpression and also results in progressive telomere shortening in cells expressing endogenous telomerase. Similarly, cells containing a CAB box mutated version of hTR also display telomere shortening (137, 178, 179).

Compound heterozygous missense mutations in WRAP53 results in pathogenic telomere shortening and the cancer predisposition syndrome DC. Two identified patients displayed autosomal recessive inheritance of WRAP53 mutations that altered WRAP53β amino acid sequence. One patient had mutations in exon 2 and exon 8 (F164L and R398W, respectively), whereas the other had mutations in exon 7 and exon 9 (H376Y and G435R, respectively) of WRAP53. The mutations reside in highly conserved residues of WRAP53β and were predicted to impair the function of its WD40-domain. Interestingly, overexpression of WRAP53β that contains the DC-associated mutations resulted in its diminished nuclear localization and Cajal body targeting. Patient cells displayed severely reduced endogenous WRAP53β protein levels, especially in the nuclear compartment. This alteration was suggested to occur post-transcriptionally, since WRAP53β mRNA levels were not affected. The patient carrying the F164L/R398W mutation developed squamous cell carcinoma of the tongue associated with reduced WRAP53β levels and Cajal body localization. As a consequence, hTR is not properly targeted to Cajal bodies and instead mislocalize to nucleoli. This is unrelated to RNA levels, since hTR, scaRNA and snoRNA levels are not affected upon WRAP53β knockdown or in DC cells. Therefore, WRAP53 mutations cause DC through impaired intracellular trafficking of the telomerase enzyme (138, 180).

In addition to associating with the CAB box motifs of scaRNAs, WRAP53β binds an additional class of RNAs termed AluACA RNAs. Interestingly, the AluACA RNAs originate from Alu repetitive elements and are processed into H/ACA and CAB box-containing RNAs. Accordingly, these RNAs also associate with dyskerin, NOP10, NHP2 and GAR1, in addition to WRAP53β. Although containing CAB box motifs, the AluACA RNAs accumulate in the nucleoplasm and not to Cajal bodies. However, the 3’ hairpin of AluACA RNAs carries two closely spaced CAB box motifs, instead of the single 3’ or 5’ hairpin CAB box sequence of H/ACA scaRNAs. Whether the differentially arranged CAB boxes contribute to the subcellular localization of AluACA and H/ACA RNAs remains elusive. The function of the AluACA RNAs is not known, however they might guide site-specific RNA pseudouridylation as well (136, 181).

Moreover, WRAP53β is essential for Cajal body maintenance and promotes cancer cell survival (182, 183). SNPs within the WRAP53 gene increase the susceptibility to breast and ovarian cancer (184-186). One of the cancer-predisposing SNPs in
WRAP53 also correlates to defective DNA repair and hematotoxicity in workers exposed to benzene (187). Recently, WRAP53β has been identified as a novel player in the DNA damage response (DDR). WRAP53β promotes DNA repair by mediating the assembly of repair proteins at DNA breaks (Paper III).

Common to all of the described functions of WRAP53β is the recruitment and proper targeting of factors to cellular sites. WRAP53β controls telomerase localization to Cajal bodies as well as to telomeres and disruption of this intracellular trafficking causes DC (137, 138, 153). Moreover, WRAP53β binds the SMN complex in the cytoplasm and promotes its nuclear entry and Cajal body-localization. SMA patients display defective WRAP53β-SMN binding and reduced SMN accumulation in Cajal bodies (182). In addition, WRAP53β binds scaRNAs and promotes their localization to Cajal bodies (136). Finally, WRAP53β binds the E3 ligase RNF8 and facilitates its interaction with MDC1, which is required for proper localization of RNF8 to DNA breaks and downstream repair events (Paper III). This leads us into the final topic of this introduction, which is the DDR.

2.3 The DNA damage response

The DNA in our nuclei is particularly vulnerable to insults since it carries the genetic information passed between cells and over generations. The integrity of DNA is constantly threatened by potentially damaging agents originating from both endogenous and exogenous sources. Spontaneous chemical reactions and byproducts of normal cellular metabolism underlie endogenous damage to DNA. Environmental and exogenous genotoxic agents such as ultraviolet-light (UV), ionizing radiation (IR) and chemicals all generate DNA damage. Although difficult to estimate, DNA could be subjected to approximately $10^5$ spontaneous lesions/cell/day. The cellular response to DNA damage depends on numerous factors such as type of lesion, dose of the break-inducing agent, cell type, cell cycle position and chromatin context. The DNA double-strand break (DSB) is considered to be the most cytotoxic lesion and two main repair pathways termed homologous recombination (HR) and non-homologous end joining (NHEJ) promote repair of these breaks. Unrepaired or misrepaired DNA damage can induce genomic alterations, for instance in the form of translocations or by introducing mutations in genes. To counteract these potentially life-threatening events, cells activate a cellular signaling cascade collectively termed the DDR that serves to maintain genomic integrity upon DNA damage-induction. The DDR coordinates basic cellular processes such as replication and transcription with DNA repair, eventually determining the destiny of the damaged cell – survival, senescence or apoptosis (188-190).

2.3.1 Genomic instability and carcinogenesis

Cancer is a collective term describing a complex and heterogeneous group of diseases, varying in a range of aspects including tissue origin, genetic background, treatability and clinical outcome. Although diverse, a common set of acquired capabilities is shared by most human cancers. By activating genes called oncogenes, which are involved in promoting carcinogenesis, cancer cells are able to sustain proliferation and angiogenesis. Similarly, by inactivating the genes involved in suppressing tumor formation (i.e. tumor suppressor genes), malignant cells become unresponsive to anti-proliferative signaling and apoptosis. A prominent hallmark of
cancer, that also differentiates a malignant tumor from a benign, is the capacity of the primary tumor to metastasize to distant sites via tissue invasion. Cancer development can be seen as a step-wise process, where normal cells gradually transform into malignant cells via the accumulation of genetic changes conveying proliferative benefits. Substantiation of this hypothesis lies in the fact that cancer incidence is increased with age (191). Genomic instability represents an enabling characteristic in the multistep process that results in the acquisition of the abovementioned capabilities. The term genomic instability in cancer describes all genetic alterations present in the tumor and also includes a rate aspect where the mutational process display increased speed. Genomic instability can be manifested as point mutations, deletions, insertions, gene amplifications, translocations and alterations in chromosome numbers, which contributes to intra-tumor heterogeneity. Such alterations in genes controlling cellular homeostasis drive cellular proliferation and cancer progression. Genomic instability is a characteristic feature of most human cancers (192-194).

The DDR is intimately linked to genomic instability and carcinogenesis. Numerous studies have highlighted the DDR as an essential barrier to cancer development. The DDR is activated already in early lesions of different cancer types. The activated DDR is thought to represent DNA replication stress conferred by oncogene expression. By deregulating oncogenes, pathways normally restricting cellular proliferation can be bypassed resulting in an accelerated S phase entry. However, the unscheduled S phase progression results in an exhaustion of factors required for normal DNA replication, leading to collapsed replication and the induction of DNA breaks. The DDR is then activated to induce the proper protective responses, such as DNA repair, senescence or apoptosis, to suppress genomic instability and carcinogenesis. Given that a functional DDR serves to eliminate DNA damage, cells defective in their responses to DNA damage will continue to proliferate albeit with genetic alterations. Selective pressure towards p53 inactivation represents one such survival response following DNA damage (195-198). Indeed, hereditary human disorders associated with deficient DNA repair commonly predispose to premature ageing and cancer (189, 199, 200). Another source of genomic instability comes from telomere attrition. Upon progressive or pathogenic telomere shortening, the DDR machinery recognizes telomeres as broken DNA ends. In an attempt to repair the DNA break, chromosome ends can fuse to each other resulting in random breakage during mitosis. Further fusion-breakage-bridge cycles continue with each cell division leading up to uneven distribution of genetic material between daughter cells (139, 201). Aberrant DNA repair of non-telomeric DSBs can also induce genomic rearrangements (202). For instance, defects in classical HR and NHEJ DSB repair pathways can shift the balance towards more mutagenic alternatives (203-205).

2.3.2 Cancer therapies and DNA damage

Most available anti-cancer therapies, such as radiotherapy and chemotherapy, rely on the induction of DNA breaks to induce cell death. DNA damaging agents target highly proliferative tissues, which not only include tumor cells but also adversely affect the gastrointestinal tract, bone marrow and hair follicles. However, many cancer cells display deficient DNA repair systems rendering them sensitive to the induction of DNA damage, whereas intact DNA repair pathways of normal cells decrease the toxicity associated with DNA damage (206-208). This phenomenon is exemplified by the highly increased sensitivity of HR- and NHEJ-deficient cells to
DNA damaging agents (209). An enhanced DDR can also be associated with therapy resistance. For instance, the HR protein RAD51 is overexpressed in many cancers and results in enhanced cell survival following the induction of DNA DSBs (210-212). Furthermore, BRCA-deficiency is correlated with genomic instability, predisposition to cancer and increased radiosensitivity due to DNA repair defects. Intriguingly, the intrinsic DNA damage-sensitivity of BRCA2-mutant cells can be reversed by secondary BRCA2 mutations resulting in acquired drug resistance (213, 214). Finally, DDR deficiency can present a challenge in terms of cancer management, since patients with deficient DNA repair systems are predisposed to cancer and at the same time hypersensitive to DNA damage. Radiation mortality has been reported in patients carrying defective DNA repair systems due to highly increased toxicity of the DNA damaging agent (215, 216).

It might seem counterintuitive that anti-cancer treatment is based on cell death induction upon DNA damage exposure, when evasion of apoptosis and selective pressure towards inactivating pro-apoptotic genes is a hallmark of cancer development. In order for anti-cancer therapy to be effective, cancer cells must still retain proficient cell death responses upon induction of DNA damage associated with therapeutic agents. Generally, a reduced DDR is positively correlated with therapeutic outcome, whereas defective pro-apoptotic factors are linked to therapy resistance (199). Given that deficient DNA repair is correlated with therapeutic response, inhibition of DNA repair can increase the efficiency of conventional cytotoxic agents. For instance, inhibition of the apical DDR kinases DNA-PK or ATM results in extreme sensitivity to several DNA damaging agents. The concept of DNA repair inhibitors as cytotoxic potentiators has been expanded to exploit cancer-specific DNA repair defects. Cancer cells may rely on alternative means of DNA repair to compensate for their intrinsic DDR deficiency, such as increased dependence on a specific repair factor or on another repair pathway compared to non-cancerous cells. This is exemplified by the synthetic lethality of PARP inhibition in cancers harboring BRCA1/BRCA2 mutations. The specific killing of BRCA-deficient cells upon PARP inhibition is attributed to the role of PARP in the repair of single-strand DNA breaks (SSBs). Unrepaired SSBs promoted by PARP inhibition are converted to DSBs during replication, which in turn should be repaired by BRCA-mediated HR. However, a direct role of PARP in HR has also been suggested to promote the synthetic lethal interaction. Furthermore, PARP inhibition appears to promote selective killing of cells deficient in other HR components than BRCA1/BRCA2 (209, 217, 218).

2.3.3 DNA double-strand breaks

A DNA DSB is generated when the phosphodiester backbone is simultaneously broken on two closely located sites on opposite DNA strands. DSBs represent a particularly deleterious DNA lesion owing to the loss of genetic information on both DNA strands. Accordingly, SSBs are less challenging since the undamaged strand remains physically intact and can be used as a template to restore the original DNA sequence of the broken strand. Indeed, a single unrepaired DSB in lower organisms can have profound effects on cell survival (190, 219).

DSBs arise through several mechanisms. Programmed DSBs occur naturally during meiotic recombination and upon maturation of the immune system. Endogenous
reactions can produce DSBs in a secondary manner, which for example occurs when the replication fork encounters damaged bases or SSBs. Exogenous sources such as topoisomerase I/II inhibitors, DNA crosslinking agents, radiomimetic drugs and IR directly or indirectly induce DSBs. Although IR is estimated to produce approximately 35 DSBs/Gy/cell compared to 1000 or even more SSBs and damaged bases, the DSB is still considered to be the main cytotoxic lesion (220). As mentioned previously, dysfunctional telomeres can also be a source of DSBs (190, 221-223).

2.3.4 DNA damage signaling

Following the induction of DNA DSBs, a multitude of different proteins are mobilized to the break sites to initiate the signaling cascades required for proper repair. DDR proteins are classically divided into DNA damage sensors, transducers, mediators and effectors. As the names imply, certain factors initially sense and signal the presence of DSBs to proteins that can amplify and transmit the signal to effector molecules that produce the appropriate responses (224). The DDR is regulated both spatially and temporally, where factors accumulate in areas surrounding the DSBs in a hierarchical order. The accumulation of proteins at DSBs can be visualized microscopically as DNA repair centers referred to as foci (or IRIF for ionizing radiation-induced foci). Similar to Cajal bodies, the sequestration of factors into foci allows for the local enrichment of related factors into a specialized compartment. Moreover, the binding of factors to DSBs might protect the free ends from degradation or improper repair events (225-227). Interestingly, there are also sub-compartments within the DSB, which separate specific components of the DDR. Furthermore, certain DSB repair proteins do not form detectable foci, either due to low local concentrations or transient associations (225). Moreover, as observed in H2AX−/− MEFs, the concentration of repair factor into IRIFs is separate from the actual recruitment of factors to DSBs (228). Tethering experiments revealed that even in the absence of DNA breaks, immobilization and local accumulation of DNA damage components is sufficient to activate the DDR. Protein accumulation at chromatin displayed an internal hierarchy, although there are feedback mechanisms functioning to amplify the DNA damage signal (229, 230). The temporal aspect of the DDR is also evident in the recruitment kinetics of repair factors to DSBs, where the upstream components accumulate first to allow for the arrival of downstream DSB factors that indeed display delayed kinetics (227).

A cascade of post-translational modifications that enable specific protein interactions regulates the recruitment of repair factors to DSBs. Several different types of modifications have been identified to play essential roles in the DDR, including phosphorylation, methylation, acetylation, poly(ADP-ribosylation) (PARylation), ubiquitylation and SUMOylation. One of the earliest DDR events is phosphorylation of the histone variant H2AX on serine 139 (referred to as γH2AX). ATM, ATR and DNA-PK of the PI3K-related kinase family all mediate phosphorylation of H2AX (231). ATR appears to be most critical for H2AX phosphorylation in situations of replication stress and in response to UV damage (232). ATM is considered to be the major mediator of H2AX phosphorylation upon DSB induction, however DNA-PK functions in a redundant and overlapping manner to ATM (233, 234).

ATM is activated in response to DNA damage and becomes autophosphorylated on serine 1981, which dissociates inactive ATM dimers into monomers with kinase
activity (235). Exactly how ATM senses the DNA damage and converts into an active form is not fully understood, although reports suggest the involvement of chromatin rearrangements that accompany DNA lesions. The autophosphorylation of ATM appears to be crucial for its stabilization at DSBs (236). The MRN complex, consisting of MRE11, RAD50 and NBS1, can directly sense DSBs and bind DNA. NBS1 of the MRN complex has been reported to interact with ATM and mediate its recruitment to DSBs (237, 238). However, numerous factors have been suggested to regulate ATM activity and localization including TIP60 and MOF. Besides H2AX, activated ATM phosphorylates multiple substrates and mediates DNA damage-induced cell cycle arrest (239).

2.3.5 γH2AX

H2AX is a member of the histone H2A family, which together with H2B, H3 and H4 forms the basic units of chromatin. Depending on the cell type, H2AX constitutes approximately 2-25% of the total H2A molecules. H2AX differs from H2A primarily by a longer C-terminal tail containing a conserved SQ phosphorylation motif. The identification of this serine residue as a DNA damage-responsive phosphorylation site was a key step towards establishing γH2AX as a central DDR player (240).

H2AX is rapidly phosphorylated following the formation of DSBs, reaching maximal amounts within minutes. Phosphorylation of H2AX spreads over large distances of 1-2 megabases flanking the break site (188, 241). Spreading of γH2AX is important for focus formation and retention of proteins at DNA breaks, possibly by generating further binding sites for DDR factors and promoting changes in the chromatin environment that facilitates the access to DSBs (242). Accordingly, H2AX is not essential for the initial recruitment of factors such as 53BP1 and BRCA1 to DSBs, but they fail to form IRIFs and are not retained at the sites (228). H2AX−/− mice are viable and exhibit a relatively mild phenotype with increased radiosensitivity, genomic instability, immunodeficiency, growth defects and male infertility (243, 244). Whereas H2AX−/− mice do not show increased tumor susceptibility, combined loss of H2AX and p53 increases cancer susceptibility and results in solid tumors as well as hematological malignancies associated with translocations (245, 246). H2AX is implicated in both HR and NHEJ repair, but is not essential for either of the two DSB repair pathways (247, 248). This indicates that γH2AX only regulates the repair of a subset of DSBs or serves to optimize repair efficiency. Consistent with the viability of H2AX−/− mice and the redundant role of γH2AX in HR directed repair, depletion of HR components all result in embryonic lethality (247). The focal accumulation of γH2AX has been used as an estimate of DNA DSBs and repair, since the numbers of γH2AX foci appears to directly correlate with the number of DSBs (241, 249). The assessment of γH2AX can also be used in the clinic, both as a cancer biomarker and a predictor of treatment (196, 250, 251).

H2AX is also regulated on the level of dephosphorylation. In addition to serine 139, H2AX is also phosphorylated on tyrosine 142 (Y142). Dephosphorylation of Y142 appears to be required for proper γH2AX-dependent recruitment of downstream DDR factors (252). Several phosphatases are involved in regulating γH2AX, including PP1, PP2A, PP4, PP6 and WIP1. Impairment of these phosphatases results in deficient γH2AX removal (253). γH2AX can also be evicted from nucleosomes, which together
with γH2AX dephosphorylation could represent a mechanism to terminate the DDR and restore chromatin organization to the pre-damaged state (254, 255).

2.3.6 MDC1

Specific binding to γH2AX requires protein domains that recognize phosphorylated amino acids. The FHA and BRCT domains mediate phosphorylation-dependent interactions and are frequently found within DDR proteins (248). One such protein is MDC1, which directly interacts with γH2AX via its tandem BRCT domain. At DSBs, MDC1 functions as an adaptor protein that via multiple interactions recruits downstream DDR proteins to damaged sites. MDC1 amplifies γH2AX signaling via a positive feedback loop, activates checkpoint responses upon IR treatment and might protect γH2AX from dephosphorylation (256-258).

Besides the C-terminal BRCT domain, MDC1 also contains a N-terminally located FHA domain. The MDC1 FHA domain has been described to account for its interactions with ATM, Chk2 and RAD51 (259). Phosphorylation of MDC1 by ATM or CK2 regulates its association with RNF8 and NBS1, respectively. The interaction between MDC1 and NBS1 is important for MRN complex retention at DSBs (260-263). By interacting with γH2AX, ATM and the MRN complex, MDC1 functions as a scaffold protein that enables further ATM-dependent phosphorylation of more distal H2AX molecules and thereby amplifies DNA damage signaling. Interestingly, MDC1−/− mice shows striking similarities to their H2AX-deficient counterparts, with chromosome instability, radiation sensitivity, immune defects, growth retardation and male infertility (264).

MDC1 retention at DSB sites is regulated by proteasomal-dependent degradation, potentially via the SUMO-targeted ubiquitin ligase RNF4. The disassembly of MDC1 foci appears to facilitate the access of other repair proteins to DSBs (265, 266).

2.3.7 RNF8 and RNF168

Over the past years, ubiquitylation in response to DNA damage has emerged as a central signaling pathway in the DDR. Ubiquitin is a small protein that is covalently attached to substrate lysine residues in a process termed ubiquitylation. Ubiquitylation is carried out in a stepwise manner by enzymes catalyzing the activation, conjugation and ligation of ubiquitin (mediated by E1, E2 and E3 enzymes, respectively). Due to the presence of lysine residues within ubiquitin itself, polyubiquitin chains with varying properties can be formed. For instance, ubiquitin chains linked at K48 typically target proteins for degradation whereas K63-linked polyubiquitin chains are particularly important for signaling and in the recruitment of DNA repair proteins. Both classical ubiquitin-mediated degradation and non-proteolytic ubiquitin conjugation play fundamental roles in the DDR (253, 267).

As mentioned above, RNF8 is a DNA damage-responsive interaction partner of MDC1. Due to the presence of a N-terminally located FHA domain commonly present in DDR proteins, RNF8 was postulated to participate in the DDR. Indeed, RNF8 was found to accumulate at sites of DNA damage. Interestingly, RNF8 also contains a C-terminal RING domain. The RING domain is found in over 600 human E3 ligases and mediates ubiquitin ligase activity by transferring ubiquitin from E2 conjugating enzymes onto substrates (268). Several studies have substantiated an
important link between MDC1 and RNF8 in the DDR. The interaction between MDC1 and RNF8 is dependent on the FHA domain of RNF8, whereas the RING domain is dispensable. Strikingly, the localization of RNF8 to DSBs is also FHA-dependent while RING-independent. Silencing of MDC1 disrupts RNF8 localization to DSBs, whereas MDC1 reconstitution restores RNF8 IRIF. Mechanistically, MDC1 contains four TQXF clusters that are phosphorylated by ATM upon DNA damage, which in turn are recognized by RNF8. Thus, following DNA damage, MDC1 binds γH2AX and is subsequently phosphorylated by ATM. The ATM-mediated phosphorylation of MDC1 serves as a docking site for the FHA domain of RNF8, resulting in the targeting of RNF8 to DSBs. Although the RING domain of RNF8 is dispensable for its recruitment to DSBs, both the FHA and the RING domains are required for the accumulation of downstream DDR factors. RNF8 catalytic activity was shown to be required for ubiquitylation of histones H2A and H2AX, which in turn allows for 53BP1, BRCA1 and RAD51 accumulation at damaged sites, representing an ubiquitin-mediated route of protein recruitment (Figure 5) (269-272).

Another E3 ligase, called RNF168, was later identified to be involved in mediating DSB-associated ubiquitylation. RNF168 is mutated in the RIDDLE syndrome, which is characterized by radiosensitivity, immunodeficiency and an inability to recruit 53BP1 to DSBs (273). RNF168 contains a RING domain and two types of ubiquitin-binding domains, where the latter are required for RNF168 DSB targeting (274). RNF8 is the first E3 ligase targeted to DSBs and accumulates independently of RNF168. However, RNF8 silencing results in impaired RNF168 DSB recruitment, placing RNF8 upstream of RNF168. Nevertheless, DSB-ubiquitylation is impaired in the absence of RNF168, suggesting that RNF8 and RNF168 cooperate to maintain ubiquitylation and the downstream recruitment of repair factors. Due to the requirement of its ubiquitin-binding domains for DSB localization, RNF168 has been proposed to recognize ubiquitin chains on H2A and H2AX generated by RNF8 and then function in the amplification of these ubiquitin conjugates (275, 276). Recently, this hypothesis was challenged by a study demonstrating that RNF8 appears to lack ubiquitylation activity towards nucleosomal histones, whereas RNF168 promotes monoubiquitylation of histone H2A/H2AX on lysine 13 and 15. These monoubiquitinated sites are then extended into K63 polyubiquitin chains in an RNF8-dependent manner. Since RNF8 localizes to DSBs prior to RNF168, it is believed that the catalytical activity of RNF8 is first directed towards a non-histone substrate forming the ubiquitin chains required for RNF168 recruitment. RNF168 then primes histones for subsequent RNF8-mediated ubiquitin chain elongation (277, 278). This is in line with previous studies arguing that RNF8 displays di- and polyubiquitylation activities without any substantial effect on monoubiquitylation (271, 279).

E3 ubiquitin-ligating enzymes work in conjunction with E2 ubiquitin-conjugating enzymes. The E2 conjugating enzyme UBC13 functions together with both RNF8 and RNF168 to specifically catalyze K63-linked ubiquitylation at DSBs. RNF8 can also associate with the E2 enzyme UBCH8 to promote K48-linked polyubiquitination (280). Interestingly, the E3 ligase HERC2 was identified to promote the association between RNF8-UBC13 and weaken the binding of RNF8 to UBCH8, representing a regulatory mechanism in terms of paring RNF8 to a specific E2 enzyme (281, 282).

A process termed chromatin remodeling-assisted ubiquitylation was discovered when the chromatin remodeling factor CHD4 was found to facilitate RNF8-mediated
ubiquitylation events. RNF8 recruits CHD4 to sites of DNA damage, which in turn triggers chromatin decondensation that allows for RNF8 ubiquitin conjugation and efficient recruitment of downstream DDR factors (230). Interestingly, MDC1 and RNF168 also promote chromatin remodeling at DSBs by associating with specific chromatin remodeling enzymes (283, 284).

RNF8 deficient mice exhibit increased radiosensitivity, genomic instability, carcinogenesis, immunodeficiency and impaired spermatogenesis (285-287). Similar to the human RIDDLE syndrome, RNF168−/− mice are immunodeficient and radiation sensitive. Moreover, they also display defective spermatogenesis. In contrast to RNF8−/− mice, tumor predisposition is not increased in RNF168−/− mice. However, double knockout mice of both p53 and RNF8 or RNF168 display increased levels of genomic instability and tumor incidence compared to the respective single knockout mice. This indicates that RNF8 and RNF168 functions together with p53 to suppress genomic instability and carcinogenesis (288, 289).

RNF168 can promote its own DSB accumulation by binding the ubiquitylation products catalyzed by itself, potentially generating a feed-forward mechanism of uncontrolled ubiquitin-spreading over undamaged chromatin. Several so-called deubiquitylating enzymes (DUBs) serve to limit and reverse RNF8/RNF168-mediated signaling. The DUBs POH1, BRCC36, USP3, USP16 and USP44 all catalyze the disassembly of RNF8/RNF168-generated ubiquitin chains. The DUB OTUB1 binds and inhibits UBC13, thus acting as a negative regulator of the RNF8/RNF168 pathway. Moreover, histone ubiquitylation can be restricted from excessive spreading by the E3 ligases TRIP12/UBR5 that negatively regulate RN168 protein stability, which in turn can be counterbalanced by USP34 that stabilizes RNF168. An additional E3 ligase termed RNF169 can antagonize RNF168-mediated protein accumulation at DSBs by occupying their binding sites (290-292).

**2.3.8 Ubiquitin-mediated protein recruitment: 53BP1, BRCA1 and RAD51**

As mentioned above, RNF8- and RNF168-mediated ubiquitylation is associated with alterations in chromatin structure and results in the recruitment of downstream repair factors such as 53BP1, BRCA1 and RAD51 to DSBs. However, the precise mode of ubiquitin-dependent DSB accumulation differs between the different factors.

**53BP1**

The NHEJ protein 53BP1 contains a C-terminal tandem BRCT domain, an upstream tandem Tudor domain and an RG-rich sequence. 53BP1 localization to DSBs is dependent on its Tudor domain, the RG-rich motif and homo-oligomerization, whereas the BRCT domain is dispensable (293-295). The tandem Tudor domain of 53BP1 specifically binds to a di-methylated residue on histone H4 (H4K20me2) (296). Since H4K20me2 is a constitutive chromatin mark that is not specific to DNA damage, it was long enigmatic how the DSB-targeting of 53BP1 ensued (297, 298). Moreover, the clear dependence of 53BP1 on RNF8/RNF168 for its DSB localization added an extra layer of complexity, arguing towards crosstalk between histone ubiquitylation and methylation. One study identified the methyltransferase MMSET to locally increase H4K20 methylation upon DNA damage, although the link to histone ubiquitylation and 53BP1 recruitment remained unexplained (299). It has
been hypothesized that the chromatin rearrangements associated with histone ubiquitylation results in the exposure of otherwise hidden methyl groups. Interestingly, two H4K20me2-binding proteins, JMJD2A and L3MBTL1, were identified to physically block the access of 53BP1 to this site. Upon DNA damage, RNF8-mediated ubiquitylation of JMJD2A and L3MBTL1 leads to their degradation or VCP-dependent displacement, respectively, thus exposing H4K20me2 for 53BP1 binding (Figure 5) (269, 300, 301). Recently, a C-terminal ubiquitylation-dependent recruitment motif was identified in 53BP1 that specifically recognizes H2A ubiquitylated on K15. Therefore, 53BP1 can bind nucleosomes modified with both H4K20me2 and the DNA damage-dependent H2AK15Ub, where the latter triggers its retention at DSBs (302). Although dispensable for most NHEJ repair events, 53BP1 is required for heterochromatin repair by promoting ATM chromatin retention (303).

BRCA1

BRCA1 forms at least three distinct sub-complexes by binding to Abraxas, BACH1 and CtIP, forming the BRCA1-A, -B and -C complexes, respectively. Furthermore, BRCA1 interacts with BRCA2 via PALB2. In addition to BRCA1 and Abraxas, the BRCA1-A complex contains BARD1, BRCC36, BRCC45, MERIT40 and RAP80 (304). RAP80 contains ubiquitin interacting motifs that bind to K63-linked polyubiquitin chains generated by RNF8/RNF168 and can therefore direct the BRCA1 complex to ubiquitin-modified DSBs (Figure 5) (305). RAP80- and Abraxas-silencing leads to modest HR defects compared to BRCA1-depleted cells, suggesting that these proteins only mediate a subset of BRCA1 functions (306). RAP80-containing BRCA1 complexes have also been demonstrated to restrict end resection and limit HR by competing with the other BRCA1 complex components (307-309). In agreement with this, depletion of RAP80 or the upstream factor MDC1 only results in BRCA1-dissociation from the DSB-flanking chromatin and does not affect the ability of BRCA1 to associate with single-stranded DNA generated at DSBs upon end resection (225, 308, 310). However, a single amino acid deletion within the ubiquitin interaction motif of RAP80 has been linked to breast cancer predisposition, suggesting functional importance of this BRCA1 recruitment pathway (311). BRCA1-B and -C complex members also localize to DSBs independent of RAP80, suggesting the presence of several independent routes of BRCA1 recruitment (308). Moreover, the BRCA1-A complex component BRCC36 is one of the DUB enzymes identified to antagonize RNF8-dependent modifications, representing an elaborate regulation of DSB ubiquitylation events during DNA repair (312).

RAD51

The link between the assembly of the central HR protein RAD51 and RNF8-mediated ubiquitylation still remains fairly descriptive. RAD51 recruitment to DSBs has been shown to be RNF8-dependent and RNF168-independent. RIDDLE syndrome patient cells display defective 53BP1 recruitment, while RAD51 foci formation is unaffected upon IR treatment (273). Silencing of RNF8 results in deficient RAD51 DSB targeting and impaired HR, whereas RNF168-depletion only show modest RAD51 recruitment defects with increased HR (269, 313). Accordingly, only RNF8 promotes RAD51 accumulation to DNA breaks induced by replication stress (314). The observed physical uncoupling between RNF8 and RNF168 has been attributed to differences in K48- versus K63-linked ubiquitylation activities. RNF8-dependent
generation of K48-ubiquitin chains is proposed to affect RAD51 accumulation, whereas the RNF8/RNF168-dependent K63-linked pathway does not (315, 316). By recruiting the proteasomal component PSMD4 to DNA damage sites, RNF8 together with RNF4 might promote the degradation of RPA, thus promoting the replacement of RPA with RAD51 at DSBs (266). However, others have reported that RNF8 regulates the recruitment of RAD51 to DSBs only in the absence of 53BP1 and BRCA1 (317). Another mechanistic explanation as to how RNF8 mediates the assembly of RAD51 comes from studies on the E3 ligase RAD18. RAD18 is recruited to DSBs by interacting with RNF8-generated ubiquitin conjugates. In turn, RAD18 promotes HR by interacting with the RAD51 paralog RAD51C. Since RAD51 IRIF formation depends on the RAD51 paralogs, both RAD18 and RNF8 deficient cells display impaired RAD51 DSB accumulation and HR repair (318). However, since RNF168 also appears to be required for proper RAD18 recruitment and UBC13 knockdown results in deficient RAD51 foci formation, the mechanism underlying the ubiquitin-dependent assembly of RAD51 is not fully understood (314, 319). Figure 5 summarizes the main steps of RNF8-mediated repair protein recruitment to DNA break sites.

![Figure 5: The classical view of RNF8-dependent assembly of DNA repair factors at DSBs. 1) DNA DSB induction, 2) The MRN complex assembles at DSBs, 3) MRN promotes ATM recruitment, 4) ATM phosphorylates H2AX (into γH2AX), 5) MDC1 binds γH2AX, 6) ATM phosphorylates MDC1, 7) RNF8 binds to phosphorylated MDC1 and subsequently targets histones for ubiquitylation, 8) RNF168 binds ubiquitylated histones. DSB-associated ubiquitylation events result in BRCA1, RAD51 and 53BP1 recruitment. RNF8 exposes H4K20me2 by JMJD2A degradation and VCP-dependent L3MBTL1 displacement, thereby generating 53BP1-docking sites.](image)

**2.4 Cellular responses to DNA damage: survival or death**

Upon DNA damage, the DDR coordinates several cellular processes such as cell cycle progression with DNA repair or DNA-damage induced cell death in order to avoid genetic alterations and malignant transformation associated with unrepaired DNA breaks. DNA repair promotes cell survival, whereas various modes of cell death results in the removal of cells carrying damaged DNA. In either case, cell cycle arrest
conferred by activated checkpoints allows for DNA repair and prevents the transmission of damaged DNA to daughter cells.

2.4.1 Homologous recombination

HR constitutes several sub-pathways that serve to repair DNA lesions including DSBs and interstrand crosslinks. Moreover, HR functions at collapsed replication forks and maintains chromosome ends by promoting telomere recombination. HR is active during late S and G2 phases of the cell cycle when DNA has been replicated and an undamaged sister chromatid is available as a repair template. HR is therefore considered to be a relatively error-free pathway (190). The gene products of the RAD52 epistasis group are central HR players, originally identified in S. cerevisiae as being required for the repair of IR induced DNA breaks. These proteins are highly conserved among eukaryotes and the human orthologs include RAD51, RAD52, RAD54, the five RAD51 paralogs and all components of the MRN complex (320).

The main steps of HR include homology search and DNA strand invasion into an intact DNA duplex. Homologous sequences used as correct repair templates can be located on a sister chromatid, a homologous chromosome or a non-homologous chromosome in the form of a repeated sequence. However, the sister chromatid appears to be the favored template since these recombination events normally are genetically silent, limiting loss of heterozygosity (LOH) (321). The invading DNA strand is produced from the DNA break by the action of nucleases. Consequently, all different HR sub-modules commence with the 5’ to 3’ processing of the DNA DSB to generate a 3’ ssDNA overhang, which serves as a substrate for the central HR recombinase RAD51. The HR machinery requires over 100 bp of sequence homology for strand invasion (320, 322-325). The combined nuclease activities of the MRN complex and CtIP initiates resection and degrades a stretch of the broken 5’ DNA ends. Further DSB processing involving DNA2, EXO1 and the BLM helicase result in the production of more extensively resected 3’ ssDNA segments (326-330).

The nucleation of RAD51 onto the ssDNA that is created during resection is a slow process and the high affinity ssDNA-binding factor RPA will coat the resected DNA segment. Therefore, RPA has to be displaced from the ssDNA in order for RAD51 to form the nucleoprotein filament responsible for homology search and strand invasion. By physically blocking access of RAD51 to ssDNA and by removing ssDNA secondary structures, RPA can both inhibit and promote recombination, respectively (331). Furthermore, RPA plays an additional role in activating the central DDR kinase ATR via ATRIP, TopBP1 and Claspin by mediating its recruitment to resected DNA ends (322, 332). Eventually, by the help of so-called recombination mediator proteins, RPA is displaced and instead replaced with RAD51. These proteins are involved in the formation and stability of the RAD51 presynaptic filament and can interact with both with the recombinase and DNA. Several recombination mediator proteins assist in the displacement of RPA, including the five RAD51 paralogs (RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3) and BRCA2 (333, 334). In yeast, RAD52 has been identified as one of the main factors responsible for the removal of RPA from ssDNA, however human RAD52 appears to lack mediator activity (335). The complete underlying mechanism as to how the recombination mediator proteins facilitate RAD51-ssDNA filament formation remains unsolved, although it is known that these factors are required for RAD51 IRIF formation (325). Presynapsis is the collective
term describing all events that result in the RAD51 filament assembly. Presynapsis is followed by synopsis, where the RAD51-ssDNA filament invades an intact DNA molecule in the search of sequence homology. RAD51 displays ATPase activity, which is necessary for catalyzing the DNA strand-exchange reaction during HR (336). The binding of the presynaptic filament to the homologous region in a duplex DNA molecule generates a three-way junction known as the displacement (D)-loop. The final step of HR is termed post-synapsis, where DNA is synthesized along the invading 3’ssDNA. DNA polymerase eta (Pol η) is proposed to be responsible for DNA synthesis extending from the D-loop intermediate during HR-mediated repair (337). After this step, separate HR models are described to repair the DNA break by different means (Figure 6).

The double-strand break repair (DSBR) model continues by involving the second end of the DSB, which generates a four-stranded DNA structure called the double Holliday junction (dHJ) (338). The dHJ is resolved in two separate manners. The dHJ can be cleaved by structure-specific endonucleases such as GEN1, MUS81-EME1, SLX1-SLX4 and ERCC1-XPF into non-crossover or crossover products (339-341). Non-crossover products stay in the parental configuration, whereas crossover products entail an exchange of flanking genetic material between the duplex DNA molecules. The two junctions of the dHJ can also migrate towards each other in a process termed branch migration and be dissolved by the joint action of the helicase-topoisomerase complex BLM-TOPOIIIα-RMI1/RMI2 with resulting non-crossover products (342, 343). During meiosis, DSBR is important for producing crossovers between homologous chromosomes to generate genetic diversity and to form the physical connections between homologous chromosomes needed for proper segregation (Figure 6A) (323).

In synthesis-dependent strand annealing (SDSA), the invading strand primes DNA synthesis to restore the lost sequence information and forms a single HJ structure that can slide via branch migration. The newly synthesized DNA end is then detached from its template sequence and anneals with the second DSB end. DNA can then be synthesized from the non-invading 3’ end. In this process, only non-crossover products are generated and the donor sequence remains unaltered. RAD54 and BLM have been implicated in displacing the invading strand to dissociate the D-loop following DNA synthesis, thus promoting SDSA to limit potential LOH (344, 345). In somatic cells, the SDSA pathway is favored over DSBR since crossover products are not formed and therefore genetic rearrangements are avoided (Figure 6B) (338).

Break-induced replication (BIR) is responsible for repairing DSBs that only have one free end, which can result from replication fork collapse or at uncapped telomeres. In the absence of telomerase, telomeres are maintained by recombination and BIR represents one mechanism for ALT. In BIR, the D-loop assembles into a replication fork that can copy the entire distal fragment of the invaded DNA strand or continue until it reaches a converging replication fork. If the template sequence is found in a homologous chromosome, long tracts of LOH can ensue upon BIR. A non-homologous chromosome template can result in gross chromosomal rearrangements. Therefore, BIR is normally suppressed when two-ended DSBs arise, although how this is achieved remains unclear (Figure 6C) (346, 347).
Figure 6: Different modes of homologous recombination. Blue circles illustrate RPA, whereas green circles represent RAD51. The DSBR (A), SDSA (B), BIR (C) and SSA (D) models are illustrated. The dashed lines represent newly synthesized DNA. In (A), endonuclease-mediated resolution of the dHJ structure can result in either non-crossover (black asterisks) or crossover products (red asterisks). The products are marked by black boxes, where the ends of the non-crossover products stay in the parental configuration whereas the crossover products have exchanged flank genetic material. In (D), blue boxes indicate direct repeat sequences.

The single-strand annealing (SSA) pathway of HR repairs two-ended DSBs that are flanked by repeated sequences, such as Alu sequences. Instead of utilizing a template DNA molecule and strand invasion, this RAD51-independent mechanism operates via RAD52 and RPA to anneal homologous ssDNA sequences surrounding the DSB exposed during 5’ end resection (348). Annealing is followed by nucleolytical processing of the remaining 3’ flap overhangs by ERCC1-XPF (349). This process is highly mutagenic, since it results in deletion of one of the repeats as well as the sequences intervening the homologous regions (Figure 6D) (338).

2.4.2 Non-homologous end joining

NHEJ is the major DSB repair pathway in mammalian cells and is active throughout the cell cycle. By simply relying on the re-ligation of broken DNA ends, NHEJ is intrinsically error-prone. If the DNA break is a so-called clean break, with blunt ends or complementary overhangs, no end processing is required before end joining and the original DNA sequence can be restored. However, some DSBs cannot be directly
re-joined and demand nucleolytic processing and DNA synthesis to generate proper substrates for DNA ligation. For instance, IR-induced DSBs usually contain base and sugar damages that must be processed before end joining can proceed. The end processing steps of NHEJ can result in mutagenic events, with loss or gain of genetic material, rendering this DSB repair pathway erroneous (350-352).

The core NHEJ machinery consists of Ku70, Ku80, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), XLF, XRCC4 and DNA ligase IV. Ku70 and Ku80 are highly abundant proteins that bind to DNA in a non-sequence dependent manner and rely on the presence of DSBs. Interaction between Ku70 and Ku80 leads to the formation of a ring-shaped heterodimer, which rapidly binds and encloses both damaged DNA ends with high affinity (353, 354). The Ku70/80 heterodimer (referred to as Ku) bound to DNA is believed to function as a scaffold for the downstream recruitment of NHEJ factors to the break site, including the nuclease, DNA polymerase and ligase activities necessary for end joining (350). Ku targets the catalytic subunit of DNA-PK to DSBs and then moves internally along the DNA. This allows DNA-PKcs to directly contact DNA and the enzymatically active DNA-PK holoenzyme composed of Ku and DNA-PKcs is formed (355). DNA-PK tethers DNA ends and brings them in close proximity in a process termed synapsis, which is essential for end joining (356). DNA-PK is a serine/threonine kinase and this activity is important for the NHEJ process (357). DNA-PK phosphorylates most NHEJ components, including itself (358). Unphosphorylated DNA-PK protects DNA ends

Figure 7: Schematic illustration of non-homologous end joining. Blue circles represent the Ku heterodimer.
from processing enzymes until they are aligned. Upon autophosphorylation, DNA-PKcs interaction with DNA ends is weakened, thus increasing the access of processing enzymes to DNA ends (359-361). The final step of NHEJ is ligation, which is carried out by DNA ligase IV in conjunction with XRCC4 and the accessory factor XLF (Figure 7) (362, 363).

As mentioned above, if the DNA ends produced from a DSB contain chemical modifications or secondary structures they are non-ligatable and require additional processing steps. To resolve incompatible DNA ends, nucleotides are either removed or replaced by cleaving or synthesizing DNA, respectively. The exo/endonuclease Artemis is a key processing protein in the NHEJ pathway. DNA-PKcs phosphorylates Artemis, which results in a conformational change that allows activation of Artemis endonuclease activity. Artemis cleaves DNA structures such as hairpins and 5'/3'-overhangs (364, 365). Furthermore, two polymerases of the Pol X family have been identified to mediate gap filling during NHEJ: DNA polymerases µ and λ. These polymerases contain BRCT domains, which mediate their interaction with Ku and DNA ligase IV-XRCC4 (366, 367). Artemis and DNA pol µ/λ thus generates DNA ends suitable for ligation, which also occurs in conjunction with several other DNA end processing factors.

The rejoining of IR-induced DNA breaks by NHEJ appears to proceed via a two-phase kinetics with differing molecular requirements. The rapid phase of NHEJ repairs approximately 90% of all induced DSBs within 2 h of damage induction and requires DNA-PKcs, Ku and DNA ligase IV but operates independent of ATM and Artemis. However, there is also a slower component of NHEJ that rejoins 10-25% of DSBs dependent on ATM, γH2AX, 53BP1 and Artemis in addition to the classical NHEJ factors. This subset of DSBs appears to be associated with increased chromatin complexity, revealing differences between heterochromatic and euchromatic DSB repair (368, 369). Furthermore, NHEJ is essential for rejoining the physiological DSBs created during B and T cell receptor diversification (V(D)J recombination) and isotype switching between immunoglobulin classes during class switch recombination (CSR). As a consequence, mutations in NHEJ components result in immunodeficiency in addition to radiosensitivity (350, 358, 370).

2.4.3 Regulation of repair pathway choice

In accordance with their implementation over the cell cycle, cells deficient in NHEJ are sensitive to IR in all cell cycle phases whereas HR-deficient cells display increased IR-sensitivity in S/G2. However, although the template requirement of HR is fulfilled during G2, NHEJ appears to be the predominant repair mechanism during this cell cycle phase as well (371-373). This raises the question regarding why NHEJ would be preferred over the more accurate HR-mediated repair and how this decision is regulated.

In contrast to humans, HR is the dominating repair pathway in simpler organisms such as S. cerevisiae. The discrepancies between lower and higher eukaryotes can be partially explained by the genomic organization of the respective species. Approximately 50% of the human genome consists of repetitive DNA sequences, which can be found in the form of Alu elements as well as in centromeric, telomeric and ribosomal DNA. Therefore, it is likely that a DSB would occur in such a
sequence, leaving the possibility of finding a homologous donor sequence in any non-homologous chromosome. A highly repetitive genome can consequently not ensure HR repair to be exact and might direct most repair events to the mutagenic SSA pathway. Yeast genomes, on the other hand, contain almost no repeated sequences making HR a beneficial repair mechanism \((351, 374)\). Budding yeast also lack the key NHEJ proteins DNA-PKcs and Artemis, suggesting that the NHEJ predominance observed in mammals could overlap with the evolutionary appearance of these factors \((375)\).

In mammals, HR is primarily responsible for the repair of replication associated-DSBs \((373)\). Proximity might in part explain why a certain DNA break is directed towards a specific DSB repair pathway. The free ends of a two-ended DSB are usually in close vicinity, thus allowing efficient NHEJ. One-ended DSBs arising during replication lack a second end available for simple re-joining, consequently promoting HR-mediated repair via the available sister chromatid. The multi-protein complex Cohesin provides sister chromatin cohesion following replication, thereby providing proximity between the damaged DNA molecule and the repair template to stimulate sister chromatid recombination \((376)\).

To counteract unwanted repair events and inappropriate repair templates, HR is coordinated with DNA replication. By keeping DNA end resection under strict cell cycle regulation, HR is only to be promoted in the presence of a sister chromatid. This is accomplished by CDK-dependent phosphorylation of resection-associated substrates. CDK enzymatic activity is limited to specific cell cycle phases due to their required interaction with fluctuating cyclins. CDK phosphorylates CtIP on S327 and T847, which promotes CtIP-BRCA1 binding and end resection in S/G2 \((377-379)\). MRE11 has been demonstrated to directly interact with CDK2, thereby promoting CtIP phosphorylation \((380)\). CDK-dependent phosphorylation of the MRN complex subunit NBS1 also promotes end resection and efficient HR \((381)\). Resected DNA ends are poor substrates for NHEJ, whereas Ku bound to DNA ends prevents resection and promotes recruitment of core NHEJ factors.

53BP1 and BRCA1 are two repair factors that appear to be essential for selection of DSB repair pathway usage. Carriers with BRCA1 mutations have highly increased risk of developing breast and ovarian cancer, which is thought to reflect a lack of functional HR. In the absence of BRCA1, replication-associated one-ended DSBs will be left un-repaired or aberrantly repaired by NHEJ. Interestingly, loss of 53BP1 rescues the genomic instability, cancer predisposition and embryonic lethality associated with BRCA1-deficiency. The cell death observed upon PARP inhibition in BRCA1 mutant cells is also alleviated by concomitant loss of 53BP1 \((382, 383)\). 53BP1 has also been found to be lost in triple-negative and BRCA1/2-mutated breast cancers, which might negatively affect treatment response since 53BP1-loss protects these cells from DNA damage hypersensitivity \((384)\). On the molecular level, deletion of 53BP1 restores RAD51 foci formation and HR repair on a BRCA1 mutant background, presumably due to increased DNA end resection. DNA ligase IV-deficiency fails to rescue the HR defect in BRCA1 mutant cells, suggesting that the observed effect is specific to 53BP1 and not a general theme of NHEJ-inactivation \((382, 383)\). Therefore it seems likely that the inhibition of 5’ end resection mediated by 53BP1 needs to be antagonized for efficient HR repair and that BRCA1 is involved in this process.
RIF1 was recently identified as the effector of 53BP1-mediated DSB repair and is recruited to DSBs via its interaction with ATM-phosphorylated 53BP1. An intricate regulatory network exists between 53BP1-RIF1 and BRCA1-CtIP. 53BP1-RIF1 prevents BRCA1 from accumulating into IRIF in G1, thus restricting BRCA1 function to the S/G2 phases of the cell cycle. Conversely, BRCA1-CtIP inhibits RIF1 in S/G2, probably by promoting resection of DSB ends (385, 386). Furthermore, super-resolution microscopy studies revealed differing spatial localization of 53BP1 and BRCA1 in single IRIF over the cell cycle. In G1, 53BP1 localizes to the core of the focus. S phase progression results in exclusion of 53BP1 from the IRIF core to a more peripheral localization, while BRCA1 instead is enriched at the core. This suggests that BRCA1 impedes 53BP1 occupancy proximal to DSBs to promote HR (387).

2.4 Cell cycle checkpoints

DNA damage-induced cell cycle checkpoints arrest damaged cells at G1/S, intra-S or G2/M phases to provide time for completion of repair. The presence of DNA damage induces a signaling cascade initiated from the break site to the actual checkpoint effectors via the apical kinases ATR and ATM. Following their activation, ATR and ATM phosphorylates Chk1 and Chk2, respectively, which in turn modifies effector proteins to initiate cell cycle arrest.

The G1 checkpoint is induced to prevent cells carrying DNA damage from entering S phase. To achieve this, activated Chk1/Chk2 phosphorylates targets such as the phosphatase CDC25A and p53. This results in SCF-β-TrCP-mediated ubiquitylation and proteasomal degradation of CDC25A, while p53 is stabilized. CDC25A can therefore not dephosphorylate and activate CDK2, which is required for G1/S transition and replication initiation. Phosphorylated and stabilized p53 maintains cell cycle arrest by inducing the transcription of p21, a negative regulator of cyclin-CDK complexes (388-390).

The intra-S checkpoint serves to delay cell cycle progression in response to disturbances in DNA replication. Similar to the G1 checkpoint, ATM/ATR-Chk2/Chk1-CDC25A-CDK2 signaling results in inhibition of DNA synthesis by the blockage of the DNA polymerase recruitment factor CDC45. An inability to reduce replication initiation upon IR-treatment, as observed in ATM deficient cells, causes radioresistant DNA synthesis (RDS) that might promote carcinogenesis (391). Furthermore, ATM-mediated phosphorylation of NBS1 also appears to contribute to the intra-S phase checkpoint (388-390).

The G2/M checkpoint inhibits damaged cells from entering mitosis in order to prevent the transmission of damaged DNA to daughter cells. Activated ATM/ATR-Chk2/Chk1 results in the sequestration or degradation of CDC25A, B and C, thereby inhibiting the activation of CDK1 that function in conjunction with cyclin B to promote G2/M progression. Prolonged G2/M arrest is partly p53-independent, since cells lacking p53 still accumulates in G2/M upon DNA damage. Chk1/Chk2 can also activate WEE1, which in turn inhibits CDK1. Upon accomplishment of DNA repair, checkpoint signaling must be terminated for the arrested cells to resume cell cycle progression. This can be regulated by different means, for instance via proteasomal
degradation and dephosphorylation. Both WEE1 and the Chk1-activating protein Claspin are ubiquitylated and degraded upon PLK1-phosphorylation, resulting in deactivation of Chk1-singaling and activation of CDC25 and CDK1. Similarly to γH2AX dephosphorylation upon DDR termination, phosphatases acting on Chk1 and Chk2 potentially reverse checkpoint events. As a result, the G2 arrest is removed and the cyclin-CDK complexes promoting G2/M transition can be activated (388-390, 392).

2.4.5 Apoptosis

An inability to repair DNA DSBs results in cell death, which serves as a protective mechanism to avoid the potentially harmful consequences associated with unresolved lesions. Cell death associated with DNA breaks is clearly manifested by DNA repair-deficient cells that display highly increased apoptotic responses to DNA damaging agents (393, 394). However, it is important to note that decreased cell survival upon exposure to IR is not strictly correlated to increased apoptosis in all cell types (220). Nevertheless, DSB-induction by restriction enzyme cleavage triggers apoptosis, where the frequency of DSBs correlates with the yield of apoptosis. The induction of apoptosis following DNA damage varies greatly between different cell types and depends on severity of the damage (395).

DNA damage most often result in the activation of the extrinsic death receptor-associated pathway or the intrinsic mitochondrial pathway (220). Apoptosis is characterized by morphological features such as cell shrinkage, membrane blebbing, chromatin condensation and DNA fragmentation. Both apoptotic pathways depend on the activation of a family of proteases termed caspases that induce apoptosis by cleaving protein substrates. The extrinsic apoptotic pathway is activated upon binding of extracellular ligands such as FasL/CD95L and TRAIL to death receptors, which triggers a cascade of events including the assembly of the death-inducing signaling complex (DISC) and the activation of caspase-8. In turn, caspase-8 initiates the activation of several so-called effector caspases resulting in protein and DNA degradation. A balance between pro-apoptotic and anti-apoptotic proteins of the Bcl-2 family regulates the intrinsic mitochondrial pathway, which upon activation results in cytochrome c release from the mitochondria. Cytochrome c forms the apoptosome together with Apaf-1 and pro-caspase-9, which activates several caspases serving to implement the death-promoting responses (396).

One important link between DNA damage and apoptosis is p53. ATM, ATR, Chk2 and Chk1 can all phosphorylate p53 following the formation of DSBs. This results in stabilization and increased transactivation activity of p53. Following activation, p53 induces the expression of pro-apoptotic target genes that belongs to both the intrinsic and extrinsic apoptotic pathways resulting in the induction of apoptosis. However, p53 can also be a pro-survival factor by regulating a different set of target genes. There are also p53-independent means of inducing apoptosis upon DNA damage, which is clinically relevant due to the high rate of p53 inactivation in human cancers. For instance, the p53 homolog p73 is transcriptionally induced upon DNA damage in a Chk1/Chk2-dependent manner. Similar to p53, p73 functions as a transcription factor that can induce the expression of pro-apoptotic target genes. Other factors have also been indicated in the induction of apoptosis following DNA damage including caspase-2, PTEN, p38 and JNK, where the latter has been shown to be recruited to
tyrosine 142 phosphorylated H2AX (220, 252, 393, 394, 397-399). DNA damage may also activate cell death via mitotic catastrophe, autophagy and necrosis or cease cellular proliferation via senescence. However, the mechanisms governing which path cells pursue following DNA damage remains incompletely understood (397).
3. AIMS OF THIS THESIS

The overall aim of this thesis was to characterize the function of WRAP53β and its association to cancer. The specific aims of each paper were:

**Paper I:** To uncover the role of WRAP53β in Cajal body formation and maintenance.

**Paper II:** To investigate the impact of WRAP53β on cancer cell survival.

**Paper III:** To elucidate the involvement of WRAP53β in DNA repair.
4. RESULTS AND DISCUSSION

4.1 Paper I

**WRAP53 is Essential for Cajal Body Formation and for Targeting the SMN Complex to Cajal Bodies**

Immunofluorescence staining of WRAP53β revealed that it is enriched within small nuclear bodies in addition to its cytoplasmic distribution. The subcellular localization of WRAP53β initiated the studies collected into Paper I.

Given the size and numbers of the WRAP53β-containing nuclear bodies, we postulated that they might represent Cajal bodies. Indeed, immunofluorescence experiments demonstrated that WRAP53β co-localized with the Cajal body marker protein coilin and that WRAP53β was found in all coilin-positive Cajal bodies. Depletion of WRAP53β resulted in Cajal body disruption and nucleolar mislocalization of coilin and SMN. Similar to coilin, WRAP53β overexpression did not induce Cajal body formation. Ectopically expressed WRAP53β localized to Cajal bodies, however high levels of WRAP53β overexpression resulted in Cajal body disruption and loss of coilin-SMN complex formation. Thus, high levels of exogenous WRAP53β appears to adversely affect endogenous WRAP53β function, potentially via self-association or sequestration of other factors that may facilitate Cajal body formation.

Further biochemical studies revealed that WRAP53β interacts with both coilin and SMN via its WD40-domain and a stretch of its C-terminal region. The same regions were found to mediate WRAP53β localization to Cajal bodies. Interestingly, while WRAP53β was required for efficient complex formation between coilin and SMN, WRAP53β interacts with coilin independent of SMN and with SMN independent of coilin. However, since previous studies demonstrated a direct interaction between coilin and SMN, we hypothesized that WRAP53β might be required for proper targeting of SMN to Cajal bodies and coilin. Since SMN is imported to the nucleus following cytoplasmic snRNP assembly, we examined the relationship between WRAP53β and SMN in the cytoplasmic compartment as well. Subcellular fractionation experiments disclosed that WRAP53β interacts with SMN both in the nucleus and cytoplasm. Knockdown of WRAP53β altered the subcellular distribution of SMN, resulting in an increased cytoplasmic accumulation and decreased nuclear localization. This suggested that WRAP53β could be involved in mediating the nuclear import of SMN. Mechanistically, WRAP53β was shown to mediate the interaction between SMN and the nuclear pore receptor importinβ, which is required for the nuclear translocation of SMN. WRAP53β is not essential for snRNP assembly, however the entire SMN complex and associated snRNPs are dependent on WRAP53β for their Cajal body localization.

Furthermore, we found that SMA type I is associated with defects in WRAP53β-SMN complex formation, which could not be explained by reduced SMN protein levels. The SMA patient cells also displayed reduced accumulation of SMN in nuclear bodies. Deficient WRAP53β and SMN binding might therefore contribute to the nuclear body targeting defects associated with SMA pathogenesis.
Unlike SMN, knockdown of WRAP53β did not alter the nuclear and cytoplasmic distribution of Sm proteins. Sm proteins have previously been demonstrated to enter the nucleus even in the absence of ongoing snRNP assembly, suggesting that unassembled Sm proteins are imported into the nucleus without binding snRNAs (57). In WRAP53β depleted cells, the nuclear pool of Sm proteins might therefore represent Sm proteins that enter the nucleus independent of SMN or stable snRNPs remaining from previous import events.

WRAP53β also directs scaRNAs, including the telomerase RNA, to Cajal bodies in a CAB box-dependent manner. Disruption of this binding by introducing CAB box mutations or depletion of WRAP53β results in scaRNA mislocalization to nucleoli. Deficient WRAP53β-mediated telomerase trafficking results in progressive telomere shortening (136, 137). Mislocalization of hTR to nucleoli is also observed in WRAP53β-associated DC. Interestingly, DC-associated mutations in WRAP53β are believed to impair the function of its WD40-domain. The WD40-domain might therefore be responsible for mediating the RNA-binding activities of WRAP53β, similar to the crucial function of the Gemin5 WD40-domain in snRNA binding. Moreover, WRAP53β DC-mutants display markedly reduced nuclear accumulation of WRAP53β, suggesting that the WD40-domain conveys nuclear translocation of WRAP53β or affects protein stability (58, 138). Therefore, it would be interesting to study whether the WRAP53β DC-mutants are defective in importinβ-binding and nuclear targeting of the SMN complex. Although DC has been associated with rRNA pseudouridylation and ribosome defects, the impact of snRNA modifications and splicing to DC pathophysiology remains to be studied.

The reason why many Cajal body-associated factors accumulate in the nucleoli upon impaired Cajal body function remains unknown. However, scaRNA redistribution to nucleoli might represent a default pathway that emerges when the CAB box is not properly recognized, since the related snoRNAs that lack the CAB box mainly localize to nucleoli. Furthermore, several nucleolar proteins are also present in the Cajal body and might direct factors to nucleoli upon Cajal body loss. Moreover, coilin has been reported to shuttle to and from the nucleoli, which might be regulated via altered coilin phosphorylation or methylation status (400, 401). Interestingly, coilin is also targeted to the nucleoli in motor neurons of SMA type I patients. Nucleolar coilin is suggested to correlate with the degree of neurodegeneration and function as an indicator of severe motor neuron dysfunction (79).

Northern blot analysis following WRAP53β immunoprecipitation revealed that WRAP53β has a strong preference towards binding CAB box-carrying scaRNAs. However, WRAP53β was also found to weakly associate with some snoRNAs and splicing snRNAs, which is likely since WRAP53β promotes Cajal body localization of the SMN complex and associated snRNPs (136, 137). Moreover, we have shown that the binding of WRAP53β to coilin, SMN and importinβ is RNA-dependent since the interactions are disrupted upon RNase A-treatment (data not shown). It is unlikely that scaRNAs are responsible for mediating all of these interactions, since they are not present in the cytoplasm.

Furthermore, we demonstrated that WRAP53β and coilin knockdown results in complete loss of Cajal bodies, whereas cells depleted of SMN still display WRAP53β- and coilin-containing residual Cajal bodies. This suggests that WRAP53β
and coilin are essential for the structural maintenance of canonical Cajal bodies. However, coilin-/- MEFs are reported to contain two distinct types of residual Cajal bodies, one containing snoRNAs and nucleolar proteins and the other snRNAs and scaRNAs. Both types of residual Cajal bodies fail to accumulate the SMN complex, further highlighting the importance of coilin in SMN localization to Cajal bodies. Whether WRAP53β is the factor that promotes snRNA and scaRNA co-localization in residual Cajal bodies remains to be elucidated. The observation that Cajal body numbers are reduced in SMA patient cells while WRAP53β-coilin binding remains intact supports the idea that ongoing snRNP biogenesis is required for Cajal body integrity (40). Nevertheless, deficient snRNP biogenesis does not disrupt all nuclear bodies, since coilin and WRAP53β remain co-localized upon SMN knockdown.

Although further studies are required to determine the precise involvement of WRAP53β in the pathogenesis of SMA, the zinc-finger protein ZPR1 exerts a similar function as WRAP53β in terms of regulating SMN. ZPR1 is localized both to the cytoplasm and Cajal bodies, interacts with SMN and is required for SMN targeting to Cajal bodies. Furthermore, the ZPR1-SMN interaction is disrupted in SMA type I patients. Subsequent mice studies revealed that ZPR1 is essential for survival and for the maintenance of motor neuron axons. Downregulation of ZPR1 in SMA mice leads to an increased loss of motor neurons and reduced lifespan. Thus, disrupted SMN interactions and nuclear targeting contributes to SMA disease progression (45, 402, 403).

In conclusion, we demonstrated that WRAP53β is a constitutive and essential component of the Cajal body. By mediating the interaction between SMN-importinβ and SMN-coilin, WRAP53β promotes SMN nuclear import and localization to Cajal bodies, respectively. Moreover, WRAP53β is involved in targeting additional factors to Cajal bodies. WRAP53β is essential for telomere maintenance, which is one of the established functional roles of Cajal bodies. As such, WRAP53β emerges as a crucial regulator of Cajal body-related processes. However, the functional involvement of WRAP53β in the splicing process remains to be determined.

4.2 Paper II

WRAP53 promotes cancer cell survival and is a potential target for cancer therapy

One of the earliest observations regarding the functional consequences of WRAP53β depletion was the resulting cancer cell death. In paper II, we set out to characterize the anti-apoptotic properties of WRAP53β and determine the mode of cell death induced by its functional absence.

To this end, we studied the expression levels of WRAP53β in a panel of non-transformed primary cells, immortalized but non-cancerous cells and cancer cell lines. Strikingly, WRAP53β expression was significantly higher in all cancer cell lines compared to the primary cells. To investigate if high WRAP53β expression was related to oncogenesis, we performed soft agar colony formation assays to determine if WRAP53β overexpression could promote anchorage-independent cell growth. Indeed, WRAP53β overexpression promoted NIH 3T3 anchorage-independent colony formation, indicative of malignant transformation.
Next, we continued by knocking down WRAP53β expression in several different cancer cell lines. As observed by time-lapse microscopy, WRAP53β-depleted cells displayed morphological alterations associated with apoptosis. To confirm this, we studied the activation of caspase-3. Cells lacking WRAP53β demonstrated caspase-3 activation and cleavage of the downstream caspase-3 target PARP. WRAP53β depletion also resulted in the activation of the pro-apoptotic proteins Bax/Bak and mitochondrial release of cytochrome c. Overexpression of the anti-apoptotic protein Bcl-2 rescued the apoptosis induced by WRAP53β depletion. In contrast, WRAP53β-depleted primary or immortalized cells did not display any significant increase in apoptosis induction. Altogether, we have shown that cancer cells activate the mitochondrial apoptotic pathway upon WRAP53β silencing, whereas normal cells are not dependent on WRAP53β expression for their survival.

Next, we studied WRAP53β expression in primary tumors of the head and neck. High WRAP53β expression was correlated to poor patient prognosis and low intrinsic radiosensitivity, whereas low WRAP53β levels was associated with disease-free survival and high intrinsic radiosensitivity. Taken together, this indicates that WRAP53β might be a prognostic biomarker of head and neck squamous-cell carcinoma (HNSCC), as well as a predictive marker of radiotherapy response.

Although WRAP53β is clearly differentially expressed between primary and cancer cells, the regulation of WRAP53β expression remains largely unknown. WRAP53β expression might be induced at the transcriptional level or display altered mRNA or protein stability in cancer cells. Alternatively, WRAP53β expression might be actively downregulated in normal cells. Either way, WRAP53β expression is essential for cancer cell survival. WRAP53β might promote cancer cell survival by elongating telomeres. However, telomerase-deficient cells still die upon WRAP53β knockdown and apoptosis associated with progressive telomere shortening is usually a slower process (404). Furthermore, the apoptosis induced by WRAP53β silencing is p53 independent, since cell lines differing in their p53 status die to the same extent. Cajal body numbers are generally increased in transformed cells, suggested to account for increased splicing demands of highly proliferative cells (11). Therefore, higher WRAP53β expression might represent its involvement in processes associated with cancer and increased cellular demands. However, the induction of NIH 3T3 anchorage-independent growth upon WRAP53β overexpression suggests a direct involvement of WRAP53β in cellular transformation.

The SMN protein has been suggested to display anti-apoptotic properties that prevent the neuronal cell death associated with SMA pathogenesis. For instance, p53 has been found to accumulate in Cajal bodies upon stress and interact with SMN. SMA is associated with a disrupted p53-SMN interaction and p53 accumulation in nucleoli. It has been suggested that SMN functions to sequester and inhibit p53, while SMNΔ7 fails to associate with p53 thus allowing p53-dependent apoptosis (405). Moreover, SMN prevents caspase-3 activation and interacts with the anti-apoptotic protein Bcl-2 to synergistically protect against apoptosis (406). SMN has been shown to specifically protect neurons from apoptosis, whereas SMNΔ7 displays pro-apoptotic activity. However, how the anti-apoptotic function of SMN is related to snRNP biogenesis and the neuronal phenotype of SMA remains incompletely understood (407). WRAP53β
might also protect neuronal cells from cell death via proper SMN association, however this has not been studied.

In conclusion, WRAP53β expression is clearly linked to cancer. WRAP53β expression is elevated in cancer cell lines and silencing of WRAP53β is associated with cancer cell-specific apoptosis and may therefore be a potential target for cancer therapy. Moreover, WRAP53β may serve as a prognostic marker in HNSCC. However, further studies are required to elucidate the exact anti-apoptotic properties of WRAP53β.

4.3 Paper III

The Cajal body protein WRAP53β regulates RNF8-mediated repair of DNA double-strand breaks

Several collective observations implied that WRAP53β might play a role in the DDR. Two large-scale proteomic screens identified WRAP53β to be enriched in the chromatin fraction upon DNA damage and as a potential DNA damage-responsive ATM/ATR substrate (408, 409). Furthermore, SNPs in the WRAP53 gene is linked to increased breast and ovarian cancer risk and with benzene-induced hematotoxicity associated with deficient DNA repair and genomic maintenance (184-187). Finally, mutations in WRAP53 cause the premature ageing and cancer susceptibility syndrome DC, two phenotypes commonly linked to deficient DNA repair (138).

In order to investigate if WRAP53β is involved in the DDR, we initially performed laser micro-irradiation experiments that induces local DNA damage. Indeed, WRAP53β rapidly accumulated at DNA breaks, reaching a maximum already 5 minutes post-damage. However, the DSB accumulation of WRAP53β was transient and no longer visible after 30 minutes. ATM and PARP inhibition decreased WRAP53β accumulation at damaged sites. Furthermore, we found that WRAP53β interacts with γH2AX at DSBs in an ATM/ATR-dependent manner. Inhibition of DNA-PK or PARP did not affect the WRAP53β-γH2AX interaction and neither did depletion of MDC1, RNF8, RNF168, 53BP1 or RAD51. We also found that WRAP53β bound to the serine 139-phosphorylated C-terminal tail of H2AX rather than the non-phosphorylated variant. By using a series of GFP-tagged WRAP53β deletion constructs, we found that both the N- and the C-terminal regions of WRAP53β were required for mediating binding to γH2AX. In addition, the previously identified ATM/ATR target site located on serine 64 of WRAP53β was required for proper γH2AX interaction. We also validated that WRAP53β was phosphorylated on serine 64 in a DNA damage- and ATM-dependent manner, with a minor contribution of ATR. Depletion of WRAP53β induced spontaneous DNA breaks as observed by an elevated number of γH2AX foci and increased DNA fragmentation in alkaline and neutral comet assays.

Given the previous involvement of WRAP53β in the intracellular trafficking of Cajal body-associated factors, we wanted to see if WRAP53β could promote the accumulation of repair proteins at DNA breaks. WRAP53β silencing did not affect γH2AX or MDC1 IRIF formation, whereas 53BP1, BRCA1 and RAD51 all displayed severely diminished DSB accumulation. Since RNF8 acts downstream of γH2AX and MDC1 to mediate ubiquitylation events required for 53BP1, BRCA1 and RAD51...
recruitment, we next studied the impact of WRAP53β silencing on RNF8 accumulation. Strikingly, RNF8 was not efficiently recruited to DSBs in the absence of WRAP53β and neither was RNF168, resulting in impaired DSB-associated ubiquitylation.

Biochemically, WRAP53β was shown to interact with both MDC1 and RNF8 in a DNA damage-, phosphorylation- and ATM-independent manner. Moreover, WRAP53β associated with MDC1 independent of H2AX and RNF8 and with RNF8 independent of H2AX and MDC1. Interestingly, MDC1 could not interact with RNF8 in the absence of WRAP53β even though ATM-mediated MDC1 phosphorylation was not impaired. As a result, WRAP53β depletion resulted in deficient HR and NHEJ repair and decreased clearance of IR-induced γH2AX foci. Moreover, WRAP53β depleted cells displayed prolonged G2/M cell cycle arrest following irradiation. Conversely, WRAP53β overexpression promoted HR and NHEJ repair and decreased cellular sensitivity to a range of DNA damaging agents. Finally, we demonstrated that WRAP53β expression was correlated with ovarian cancer survival. Low levels of WRAP53β mRNA and protein associated with reduced patient survival and lower expression of other DDR factors. In contrast, patients with tumors that expressed higher levels of WRAP53β had a more favorable outcome that clustered with active DNA repair processes.

In this study, we demonstrated that WRAP53β is dependent on both PARP and ATM activity for proper targeting to laser-induced damage. Upon binding to DNA, PARP catalyzes the synthesis of PAR-chains that are attached onto protein substrates, including PARP itself. The PARylated proteins can be recognized and bound by factors, thereby promoting their PARP-dependent accumulation at DSBs. Interestingly, WRAP53β displays similar dynamics at DSBs as PARP itself and the PARylation-dependent factors ALC1 and CHD4 (410, 411). This suggests that WRAP53β binds PAR-chains, although this has not been studied. However, WRAP53β interacts with both PARP and CHD4 (data not shown). The enzyme PARG primarily governs the short-lived nature of PARylation by rapidly degrading PAR polymers and might regulate WRAP53β DSB dissociation (412). Therefore, it would be interesting to study the effect of PARG on WRAP53β retention at DSBs. ATM mediates the phosphorylation of WRAP53β on serine 64, DSB recruitment and interaction with γH2AX. Due to the fact that ATM inhibition also blocks γH2AX phosphorylation it is hard to draw a conclusion from the interaction studies. However, WRAP53β serine 64 phosphorylation is clearly ATM-dependent and the phosphorylation-deficient S64A WRAP53β mutant is defective in γH2AX-binding. How WRAP53β phosphorylation mediates its interaction with γH2AX remains unknown. In order to specifically interact with γH2AX, WRAP53β has to be able to recognize the phosphorylated epitope. The WD40-domain has been suggested to recognize phosphorylated sequence motifs. For instance, many F-box proteins contain WD40-domains that are implicated in the recognition of phosphorylated substrates in the form of phosphodegrons (413). However, an intact WD40-domain was not sufficient to mediate WRAP53β-γH2AX binding, since the N-terminal and C-terminal regions of WRAP53β was determined to promote their interaction. Moreover, it is unlikely that the C-terminal tail of H2AX specifically recognize phosphorylated WRAP53β. Therefore, phosphorylation of WRAP53β on serine 64 might result in altered structural properties, for instance by inducing a conformational change that exposes regions required for efficient WRAP53β-γH2AX interaction. However, we
do not know if the interaction between WRAP53β and γH2AX is direct. Phosphorylated WRAP53β might interact with an unidentified factor that mediates its interaction with γH2AX. Furthermore, since PARP inhibition only affected WRAP53β recruitment to DSBs and not WRAP53β-γH2AX interaction, there appears to be two parallel paths of WRAP53β recruitment. PARP has previously been linked to DNA damage-associated ubiquitylation responses. PARylation of RNF168 promotes its interaction with the chromatin remodeler SMARCA5, which in turn facilitates RNF168-mediated ubiquitylation and repair factor assembly (283). Furthermore, BBAP and the E3 ligase BAL1 are suggested to promote early DSB ubiquitylation and repair factor recruitment in a PARP-dependent but RNF8-independent manner (414). Finally, CHD4 is recruited to DSBs by two distinct pathways, one dependent on PARP and the other on RNF8. Interestingly, CHD4 is required to promote RNF8-mediated ubiquitylation at DSBs by creating a permissive chromatin environment (230). Accordingly, the short retention time of WRAP53β at DSBs should not affect its ability to facilitate RNF8-mediated events.

WRAP53β recruitment to DNA damage sites is MDC1-independent and WRAP53β depletion does not affect the ability of MDC1 to form foci. On the contrary, RNF8 is dependent on WRAP53β for DSB accumulation. Nevertheless, WRAP53β binds both MDC1 and RNF8 in a DNA damage- and phosphorylation-independent manner. This observation might suggest that WRAP53β forms a rigid constitutive complex with MDC1 and RNF8 that is recruited as a tertiary unit to DSBs. However, this is unlikely given the varying interdependence of the three factors and since MDC1-RNF8 interaction is induced upon DNA damage. Furthermore, stable complex formation would inhibit the transient nature of WRAP53β at DSBs and instead promote its prolonged retention. Thus, the WRAP53β-MDC1-RNF8 complex must be dynamic and subjected to further regulation. ATM inhibition resulted in a marked reduction of MDC1-RNF8 association, while WRAP53β still bound to a similar amount of MDC1 and RNF8. The constitutive nature of the WRAP53β-MDC1-RNF8 interaction might therefore represent a fraction of the proteins forming complexes independent of DNA breakage. This is further supported by the fact that the S64A mutant of WRAP53β, which is deficient in γH2AX-binding, still associates with similar amounts of MDC1 and RNF8. However, since ATM inhibition and WRAP53β depletion gave a comparable negative effect on MDC1-RNF8 complex formation, WRAP53β clearly plays an important role in mediating the DNA damage-associated interaction between MDC1 and RNF8. Exactly how WRAP53β promotes efficient MDC1-RNF8 binding remains uncertain, however following DNA damage WRAP53β might alter the allosteric properties or allow for an as of yet unidentified protein modification of MDC1 or RNF8 that results in efficient FHA-mediated recognition of the ATM-phosphorylated MDC1.

Although WRAP53β is recruited to laser-induced damage, it cannot be detected in foci upon whole cell irradiation. This is probably due to low local concentrations of WRAP53β or the transient nature of its interaction with DSB sites. However, WRAP53β is detected both over the entire region of laser-induced damage and found enriched in microfoci representing resected ssDNA. A similar distribution has been reported for NBS1 and BRCA1 (225). Moreover, we do not detect WRAP53β at the damaged sites in all laser-irradiated cells. However, PLA experiments demonstrated WRAP53β-γH2AX interaction in 100% of cells and although the association between WRAP53β and γH2AX was transient, it was slightly delayed as compared to
WRAP53β DSB dissociation. Therefore, PLA might provide increased sensitivity that allows for improved WRAP53β detection, although on the level of WRAP53β-γH2AX association.

WRAP53β depletion results in the accumulation of spontaneous DNA breaks. This is consistent with a role of WRAP53β in mediating HR. For instance, RAD51 deficient cells accumulate spontaneous DNA breaks prior to cell death (415). Although γH2AX is induced upon DNA fragmentation during apoptosis, apoptotic fragments are not detected by the comet assay implying that WRAP53β depletion indeed induces apoptosis-independent DNA breakage (416, 417). Moreover, the induction of γH2AX precedes the apoptotic phenotype associated with WRAP53β silencing.

WRAP53β overexpression protects against a variety of DNA damaging agents, suggesting its involvement in additional repair events. Interestingly, MDC1 and RNF8 have been implicated to induce H2A ubiquitylation upon UV-treatment associated with nucleotide excision repair (418). Furthermore, the Fanconi anemia pathway plays an important role in repairing DNA crosslinks as those induced by MMC treatment. Defects in this pathway cause Fanconi anemia, a syndrome with striking similarities to DC, associated with bone marrow failure, premature ageing and cancer predisposition (419). Therefore, it would be interesting to determine whether WRAP53β-mutant DC cells display increased sensitivity to any genotoxic agent. The degree of cell death following IR exposure is generally correlated to the levels of induced and unrepaired DSBs (420, 421). Since we did not detect any differences in γH2AX induction between control cells and WRAP53β overexpressing cells, we postulated that WRAP53β overexpression increases DSB rejoining capacity. Indeed, NHEJ and HR repair was found to be more efficient upon WRAP53β overexpression. Similarly, RNF8 overexpression also promoted more efficient DSB repair, suggesting that it is enough to enrich for one factor involved in this cascade to increase the cellular repair capacity. However, we did not study the effect on DSB repair upon overexpression of any other repair protein, so certain rate-limiting factors may still exist in the repair process.

In summary, we have identified WRAP53β as an essential factor in the DDR. WRAP53β is phosphorylated upon DNA damage and is recruited to DSBs in an ATM/PARP-dependent manner. By facilitating the interaction between MDC1 and RNF8, WRAP53β promotes the downstream targeting of repair factors to DNA breaks. WRAP53β depletion results in deficient DSB repair, whereas increased WRAP53β expression can promote these processes. Finally, we identify WRAP53β as a potential tumor suppressor gene in epithelial ovarian carcinoma.

4.4 General discussion and concluding remarks

In this body of work, we have identified the functional involvement of WRAP53β in a variety of biological processes including Cajal body maintenance, cancer cell survival and DNA repair. Many questions remain to be answered, for instance if there are any overlapping properties between these processes, their relative contribution to the phenotype associated with dysfunctional WRAP53β and how WRAP53β itself is regulated to coordinate these processes. Furthermore, the cancer-related properties of WRAP53β might appear inconsistent, with findings suggesting that it both promotes and suppresses carcinogenesis. Clearly, many future studies are needed to address the
specific function of WRAP53β, however in this section some of these issues will be discussed.

Cajal bodies, RNA processing and the DDR

First of all, some Cajal body-components have been linked to the DDR. Coilin has been reported to interact with Ku70/Ku80. The binding of Ku to coilin occurred in a competitive manner with SMN and Sm proteins. Coilin was shown to inhibit NHEJ, presumably by preventing Ku protein recruitment to DNA ends. This was suggested to represent a function related to nucleoplasmic coilin, since Ku70 or Ku80 are not found within Cajal bodies. However, coilin and Ku might interact at telomeres since they are both described to associate with chromosome ends (422). Some DDR factors have also been found to localize to Cajal bodies. For instance in Drosophila, PARP shuttles from chromatin to Cajal bodies upon automodification (2, 423). Similarly, WRAP53β might shuttle between chromatin and Cajal bodies.

Interestingly, UV, cisplatin and IR treatment all result in Cajal body disruption and coilin mislocalization to microfoci and nucleoli. Coilin depletion also results in enhanced cell viability upon cisplatin treatment (424, 425). These observations highlight the Cajal body as a stress-responsive organelle. Coilin and WRAP53β appears to have opposing functions in terms of cell viability and NHEJ repair. On the other hand, the SMN complex members SMN and Gemin2 have been shown to promote RAD51 assembly at DSBs and HR (426). Interestingly, SMN is a chromatin-binding protein that interacts with methylated H3K79 via its Tudor domain, a site previously identified to target 53BP1 to DSBs (427, 428).

A number of large-scale screens aimed at identifying DDR factors have revealed an enrichment of proteins involved in RNA processing, although their exact involvement in the DDR remains incompletely understood. One potential link between transcription, splicing and the induction of DNA damage is the formation of the highly mutagenic R-loop. R-loops are formed when the nascent RNA transcript anneals to the DNA template strand thereby forming RNA-DNA hybrids. The R-loop structure can block DNA replication fork progression and leaves the unpaired DNA strand susceptible to insults. The splicing factor ASF/SF2 is implied to prevent R-loop formation by binding to mRNA precursors and inhibit their association with template DNA during the co-transcriptional splicing process. Depletion of ASF/SF2 results in the accumulation of DSBs (429, 430).

Finally, the neurodegenerative disorder SMA is associated with defective WRAP53β-SMN complex formation. Although WRAP53β remains functionally intact, potential disruption of its nuclear function and organization might associate to deficient DNA repair. Indeed, accumulation of DNA damage in neurons is associated with neurodegeneration (431).

Further studies are required to determine if specific Cajal body components can contribute to DNA repair or if they are involved in mediating WRAP53β-associated DNA repair events. Moreover, we do not know if the RNA-related activities of WRAP53β impact its involvement in the DDR or the phenotypes associated with its deficiency. We also have to determine if SMA cells show signs of increased DNA damage or deficient DNA repair. However, similar to available literature regarding
coilin and DNA damage, we did not observe any repair defect upon coilin depletion. Therefore, structural maintenance of the Cajal body is not directly linked to DNA repair.

**Telomeres and DNA damage**

Telomeres provide an additional important link between WRAP53β, Cajal bodies and the DDR. DDR proteins can be found both at functional and dysfunctional telomeres. At functional telomeres, associated DDR factors promote telomere homeostasis and protect telomeric ends from end joining events. For instance, DNA-PKcs is associated with telomeres and appears to promote telomere capping, since inhibition of DNA-PKcs results in telomere fusions (432). Moreover, Ku70/80 have been shown to directly interact with hTR and promote telomere maintenance (433). Dysfunctional, uncapped telomeres are recognized as DSBs by the DDR machinery and factors are assembled into specific foci associated with telomeres termed TIFs (telomere dysfunction-induced focus) (434). However, we did not observe any apparent co-localization between the spontaneous γH2AX foci induced upon WRAP53β depletion and the telomeric marker TRF2, suggesting that WRAP53β is involved in DNA repair events unrelated to its previously established role in telomere maintenance (data not shown).

HR and NHEJ repair events at telomeres result in chromosome fusions and genomic instability. For instance, RNF8 promotes repair protein assembly at telomeres by ubiquitylating telomeric ends, thereby facilitating chromosome fusion upon telomere dysfunction (435). Moreover, 53BP1 has been identified as an essential regulator of end joining events at dysfunctional telomeres. By increasing the mobility of dysfunctional telomeres, 53BP1 brings chromosome ends into close proximity thereby allowing NHEJ (436). As such, normal DNA repair is associated with genomic stability, whereas DNA repair at telomeres results in genomic instability. Interestingly, similar to telomeres being recognized as DSBs, random DSBs can be detected by telomerase. Although rare, it has been reported that telomerase can generate a telomere at DSBs. Telomere addition to DSBs have to be prevented in order to maintain genomic stability (437).

Telomerase-independent ALT cells are characterized by very long and heterogeneous telomeres that are maintained by HR-mediated events. Analogous to the role of Cajal bodies in telomerase-dependent telomere elongation, a specific subset of telomere-associated PML bodies has been suggested to promote the recombination events of ALT cells. These ALT-associated PML bodies (APBs) contain many DDR proteins in addition to the conventional PML body components (438). NBS1 is essential for their assembly and depletion of NBS1 results in decreased numbers of APBs and telomere shortening in ALT-cells, whereas the telomeres of telomerase-positive cells are unaffected (439, 440). This reveals an interesting link between nuclear body compartmentalization, telomere elongation and DDR proteins.

Regulation of telomere length is important in several aspects of cancer biology; telomere shortening induces chromosomal instability and cancer, whereas cancer progression requires telomere maintenance (441, 442). Therefore, it might appear controversial that DC patients are predisposed to cancer although they exhibit functional loss of the enzyme responsible for cancer cell immortalization. The
inadequate telomere lengths in DC do not allow for proper tissue renewal and results in premature telomere shortening that can lead to genomic instability. Even though DC cells do not completely lack telomerase activity, they may maintain their telomeres via ALT upon malignant progression (166). Since DC-associated mutations in WRAP53β are associated with its nuclear exclusion, all nuclear activities of WRAP53β must be diminished in DC, including its role in DNA repair. Therefore it would be interesting to study DC cells carrying the pathogenic WRAP53β in relation to DNA damage. Cells carrying a DC-associated dyskerin mutation was shown to display an increased number of γH2AX foci upon DSB-induction compared to control cells, which was independent of telomere lengths. Although a fraction of these foci localized to telomeres, the majority of the γH2AX foci localized elsewhere (166, 443). DNA damage accumulation might not be a general theme of DC pathology, although it could explain the clinical diversity between patients with different disease-causing genes.

Oncogenic vs. tumor suppressing properties

High WRAP53β expression has been related to both decreased and increased patient survival, radioresistance, malignant transformation and tumor suppression. This could represent the importance of a fine tuned DDR in many aspects of carcinogenesis.

For instance, both RAD51 deficiency and overexpression promotes genomic instability (205, 415). Moreover, RAD51 is overexpressed in many cancers and can potentiate survival after DSB induction (210-212). Similarly, MDC1 has been identified as an oncogene in cervical cancer due to its increased expression in cancer tissues. In this study, MDC1 knockdown in cancer cell lines resulted in mitochondrial-dependent apoptosis and increased sensitivity to DNA damage (444). Furthermore, classical DSB repair deficiency can promote the usage of more mutagenic alternatives (203, 204). Moreover, the oncogenic properties of the DDR at telomeres can also result in genomic instability rather than genomic maintenance. The DDR has profound consequences for carcinogenesis and therapeutic response since the induction of DNA breaks can both cause and treat cancer.

Although limited sample size, high WRAP53β expression clearly correlated with poor prognosis and radioresistance of HNSCC patients. This could potentially reflect an enhanced DNA repair capacity of these patients, similar to what we observed upon ectopic WRAP53β overexpression. In epithelial ovarian carcinoma, high WRAP53β expression instead correlated with increased overall survival. Low WRAP53β levels might coincide with more chromosomal instability, which is generally associated with drug resistance and poor prognosis (194). It is important to note that the levels of WRAP53β in cancer patient material are compared internally and we do not know how the levels of WRAP53β are correlated between the HNSCC and ovarian cancer samples. In this light, decreased or increased DDR capacities can regulate cancer progression, patient prognosis and response to therapy both positively and negatively. However, the DDR is generally considered to be a protective response that serves to eliminate damage.

WRAP53β is recruited to DNA damage sites in fibroblasts, however WRAP53β depletion does not significantly affect spontaneous or residual γH2AX foci (data not shown) or apoptosis in these cells. This might suggest that normal cells have better
backup systems for DNA repair, whereas cancer cells are DDR-deficient and sensitive to the depletion of single factors. SMN, coilin and RNF8 are all expressed at higher levels in cancer cells compared to normal cells (data not shown). This might reflect the increased metabolic demand and DNA damage load of cancer cell lines and not a direct oncogenic property of the specific proteins. However, we did observe that WRAP53β overexpression transforms NIH 3T3 cells. The signals governing this process remain unknown, although it highlights the need for proper regulation of WRAP53β expression in carcinogenesis.

Findings that inherited mutations in WRAP53 cause the cancer predisposing syndrome DC argues that WRAP53β mainly functions as a tumor suppressor gene. Further insights into the physiological role and contribution of WRAP53β to cancer development might come from the generation of WRAP53β knockout animal models. Nevertheless, WRAP53β could be an important target in the treatment of cancer. Cajal body disruption is expected to decrease the production of mature snRNP's resulting in inefficient splicing. Indeed, compounds targeting the spliceosome have shown anti-tumor activities (445). Moreover, inhibition of proper snRNP assembly by interfering with SMN and Sm protein interaction results in decreased cell proliferation and apoptosis induction (446). Telomerase may serve as a cancer-specific target and deficient WRAP53β-mediated telomerase trafficking would result in telomere shortening (137, 447, 448). Inhibition of DNA repair pathways can be used as monotherapy or in combination with DNA damaging agents (206). Although

**Figure 8:** Model figure of the various roles of WRAP53β. Green arrows represent WRAP53β-mediated events and interactions. Red inhibitory marks reflect WRAP53β dysfunction and the red arrow represents one consequence of WRAP53β-deficiency. The blue boxes state the potential outcomes of defective WRAP53β function.
silencing of several DDR factors results in severely deficient DNA repair, we have not observed the same amount of apoptosis as compared to WRAP53β knockdown (data not shown). Moreover, coilin depletion does not induce significant cancer cell apoptosis. Therefore, the combined inability to repair DSBs, dysfunctional telomere elongation, Cajal body disruption and additional undiscovered functions of WRAP53β might contribute to the cancer cell-associated death observed upon its depletion (Figure 8).
5. ACKNOWLEDGEMENTS

Finally it is time for some acknowledgements! This thesis would not have been possible without the combined support of several people. I would like to express my deepest appreciation to all of you and in particular:

My main supervisor Marianne. Your devotion and passion for science is really amazing! I always had this feeling that I could call you any day of the week, any time of the day and you would be happy to provide an answer to any lab-related issue. You are truly solution-oriented, hands-on and eager to make things (and people) work as fast and efficient as possible. Your tendency to micromanage is just a reflection of your immense interest in science, your great curiosity for data and the ambition to find out how things actually work. Thank you for all these years and all your support!

My co-supervisors Klas and Per. For agreeing to be so and to Klas for originally accepting me into your lab before the Farnebo group era.

My co-supervisor (!) Salah. For supervising me in the beginning of my studies. When looking back at that time, I am amazed by the time and energy you spent on teaching me basically everything. I’m convinced that I wouldn’t have continued on this track if not for you. You really have this unique ability of motivating people around you. I am convinced that you will be an excellent group leader when the time comes! I really miss you, things have not been the same since you left!

Elisabeth a.k.a DJ smooth. For mixing your music to perfection without anyone listening. If we were to vote, you would definitely win the title as lab friend! Just remember that any minor slipup can be rescued by a bear story. It is empty in the lab since you ran off with Ramon, come back soon and we will watch El Clásico together!

Hanif or do you mean Janne? Since you came along I never once had to open a door or lift anything. But I guess you have to make use of those amino acids in some way… And of course for all the great times at the laser micro-irradiation facilities, with me freezing and you chillaxing with red bull and snickers. You’re a really dedicated person who always brings a great atmosphere to the lab. And if science doesn’t work out, you know that you can always fall back on collecting cans!

Summer-Alex, or even better, all-year-round-Alex. For partially supporting me when I dismiss everyone and everything, although you are too nice to go all in. And for bringing key solutions to our projects already as a summer student! You are also one of the rare people who actually get things, kind of like a modern Voldemort. I look forward to your return!

Christos aka the Greek guy. For being a great, outgoing and smart person! I am convinced that your studies will go great! Maybe one day we will meet up in Genève (French accent) for some bryggkaffe!
Fredrik for bringing the countryside and Östgötska (uttalas [ˈœjœtska]) to the big city (Solna). Also for being a genuinely great person! What happened to that Flong guy? I’m sure you will do an amazing job in your new lab!

Adam, although hard to understand your broken Swedish (?) you definitely left your mark on the lab! Your classical expressions such as “men det går ju faktiskt inte” and “vad är en bal på slottet” live on to this day. Come by sometime, we miss having you around!

Past and present colleagues, friends and family for all great times!

Lastly to Per, for everything you always do! This would definitely not have been possible without you. I cannot imagine anyone who could be more supportive than you. Meeting you is the best thing that came from my PhD studies, I’m very glad we relocated our extensive coffee sessions outside the lab... I’m so happy for everything we have together. Jag älskar dig!
6. REFERENCES

5. T. Carvalho et al., The spinal muscular atrophy disease gene product, SMN: A link between snRNP biogenesis and the Cajal (coiled) body. The Journal of cell biology 147, 715 (Nov 15, 1999).
7. M. Carmo-Fonseca, J. Ferreira, A. I. Lamond, Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. The Journal of cell biology 120, 841 (Feb, 1993).
10. E. Pena, M. T. Berciano, R. Fernandez, J. L. Ojeda, M. Lafarga, Neuronal body size correlates with the number of nucleoli and Cajal bodies, and with the organization of the splicing machinery in rat trigeminal ganglion neurons. The Journal of comparative neurology 430, 250 (Feb 5, 2001).
12. S. M. Hearst et al., Cajal-body formation correlates with differential coilin phosphorylation in primary and transformed cell lines. Journal of cell science 122, 1872 (Jun 1, 2009).
13. P. J. Young et al., Nuclear gems and Cajal (coiled) bodies in fetal tissues: nucleolar distribution of the spinal muscular atrophy protein, SMN. Experimental cell research 265, 252 (May 1, 2001).


41. K. E. Tucker et al., Residual Cajal bodies in coilin knockout mice fail to recruit Sm snRNPs and SMN, the spinal muscular atrophy gene product. The Journal of cell biology 154, 293 (Jul 23, 2001).
44. I. Lemm et al., Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. Molecular biology of the cell 17, 3221 (Jul, 2006).
45. L. Gangwani, R. A. Flavell, R. J. Davis, ZPR1 is essential for survival and is required for localization of the survival motor neurons (SMN) protein to Cajal bodies. Molecular and cellular biology 25, 2744 (Apr, 2005).
47. B. Schrank et al., Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. Proceedings of the National Academy of Sciences of the United States of America 94, 9920 (Sep 2, 1997).
52. R. Morse, D. J. Shaw, A. G. Todd, P. J. Young, Targeting of SMN to Cajal bodies is mediated by self-association. Human molecular genetics 16, 2349 (Oct 1, 2007).
53. R. J. Cauchi, SMN and Gemins: ‘we are family’ ... or are we?: insights into the partnership between Gemins and the spinal muscular atrophy disease protein SMN. BioEssays : news and reviews in molecular, cellular and developmental biology 32, 1077 (Dec, 2010).
56. W. Feng et al., Gemins modulate the expression and activity of the SMN complex. Human molecular genetics 14, 1605 (Jun 15, 2005).
57. K. B. Shpargel, A. G. Matera, Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. Proceedings of the National Academy of Sciences of the United States of America 102, 17372 (Nov 29, 2005).
60. U. Fischer, Q. Liu, G. Dreyfuss, The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. Cell 90, 1023 (Sep 19, 1997).


66. S. Lefebvre *et al*., Correlation between severity and SMN protein level in spinal muscular atrophy. *Nature genetics* 16, 265 (Jul, 1997).


68. S. Lefebvre *et al*., Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80, 155 (Jan 13, 1995).


71. S. Cho, G. Dreyfuss, A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes & development* 24, 438 (Mar 1, 2010).


74. M. Feldkotter, V. Schwarzer, R. Wirth, T. F. Wienker, B. Wirth, Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *American journal of human genetics* 70, 358 (Feb, 2002).


77. S. Jablonka, B. Schrank, M. Kralewski, W. Rossoll, M. Sendtner, Reduced survival motor neuron (Snm) gene dose in mice leads to motor neuron degeneration: an animal model for spinal muscular atrophy type III. *Human molecular genetics* 9, 341 (Feb 12, 2000).

78. U. R. Monani *et al*., The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn(-/-) mice and results in a mouse with spinal muscular atrophy. *Human molecular genetics* 9, 333 (Feb 12, 2000).

79. O. Tapia *et al*., Reorganization of Cajal bodies and nucleolar targeting of coilin in motor neurons of type I spinal muscular atrophy. *Histochemistry and cell biology*, (Feb 1, 2012).

82. F. Gabanella et al., Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. PloS one 2, e921 (2007).
83. L. Pellizzoni, B. Charroux, G. Dreyfuss, SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. Proceedings of the National Academy of Sciences of the United States of America 96, 11167 (Sep 28, 1999).
84. C. Winkler et al., Reduced U snRNP assembly causes motor axon degeneration in an animal model for spinal muscular atrophy. Genes & development 19, 2320 (Oct 1, 2005).
85. K. Praveen, Y. Wen, A. G. Matera, A Drosophila model of spinal muscular atrophy uncouples snRNP biogenesis functions of survival motor neuron from locomotion and viability defects. Cell reports 1, 624 (Jun 28, 2012).
86. S. Pagliardini et al., Subcellular localization and axonal transport of the survival motor neuron (SMN) protein in the developing rat spinal cord. Human molecular genetics 9, 47 (Jan 1, 2000).
88. T. Kiss, Biogenesis of small nuclear RNPs. Journal of cell science 117, 5949 (Dec 1, 2004).
94. C. Branlant et al., U2 RNA shares a structural domain with U1, U4, and U5 RNAs. The EMBO journal 1, 1259 (1982).
96. W. J. Friesen, S. Massenet, S. Paushkin, A. Wyce, G. Dreyfuss, SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. Molecular cell 7, 1111 (May, 2001).
97. J. Mouaikel et al., Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. EMBO reports 4, 616 (Jun, 2003).
100. U. Fischer, V. Sumpter, M. Sekine, T. Satoh, R. Luhrmann, Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. The EMBO journal 12, 573 (Feb, 1993).


114. G. Tanackovic, A. Kramer, Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. *Molecular biology of the cell* 16, 1366 (Mar, 2005).


121. K. P. Smith, J. B. Lawrence, Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies. *Molecular biology of the cell* 11, 2987 (Sep, 2000).


134. R. L. Tomlinson et al., Telomerase reverse transcriptase is required for the localization of telomerase RNA to cajal bodies and telomeres in human cancer cells. *Molecular biology of the cell* 19, 3793 (Sep, 2008).


151. C. Trahan, F. Dragon, Dyskeratosis congenita mutations in the H/ACA domain of human telomerase RNA affect its assembly into a pre-RNP. Rna 15, 235 (Feb, 2009).
154. G. Cristofari et al., Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. Molecular cell 27, 882 (Sep 21, 2007).
164. A. J. Walne et al., Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. Human molecular genetics 16, 1619 (Jul 1, 2007).
165. N. S. Heiss et al., X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nature genetics 19, 32 (May, 1998).


234. T. Stiff et al., ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer research* 64, 2390 (Apr 1, 2004).


242. V. Savic et al., Formation of dynamic gamma-H2AX domains along broken DNA strands is distinctly regulated by ATM and MDC1 and dependent upon H2AX densities in chromatin. *Molecular cell* 34, 298 (May 15, 2009).


244. C. H. Bassing et al., Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proceedings of the National Academy of Sciences of the United States of America* 99, 8173 (Jun 11, 2002).


247. J. Yuan, R. Adamski, J. Chen, Focus on histone variant H2AX: to be or not to be. *FEBS letters* 584, 3717 (Sep 10, 2010).


255. M. Altmeyer, J. Lukas, To spread or not to spread--chromatin modifications in response to DNA damage. *Current opinion in genetics & development* 23, 156 (Apr, 2013).


G. S. Stewart et al., The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell* **136**, 420 (Feb 6, 2009).

C. Doil et al., RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell* **136**, 435 (Feb 6, 2009).


310. B. Sobhian et al., RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198 (May 25, 2007).


362. P. Ahnesorg, P. Smith, S. P. Jackson, XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 124, 301 (Jan 27, 2006).


375. S. J. DiBiase et al., DNA-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. *Cancer research* **60**, 1245 (Mar 1, 2000).


400. C. E. Lyon, K. Bohmann, J. Sleeman, A. I. Lamond, Inhibition of protein dephosphorylation results in the accumulation of splicing snRNPs and coiled bodies within the nucleolus. *Experimental cell research* 230, 84 (Jan 10, 1997).
403. S. Ahmad, Y. Wang, G. M. Shaik, A. H. Burghes, L. Gangwani, The zinc finger protein ZPR1 is a potential modifier of spinal muscular atrophy. *Human molecular genetics* 21, 2745 (Jun 15, 2012).


412. M. Y. Kim, T. Zhang, W. L. Kraus, Poly(ADP-ribose)ylation by PARP-1: 'PAR-lying' NAD+ into a nuclear signal. *Genes & development* 19, 1951 (Sep 1, 2005).


414. Q. Yan et al., BAL1 and its partner E3 ligase, BBAP, link Poly(ADP-ribose) activation, ubiquitylation, and double-strand DNA repair independent of ATM, MDC1, and RNF8. *Molecular and cellular biology* 33, 845 (Feb, 2013).


422. V. Velma, Z. I. Carrero, A. M. Cosman, M. D. Hebert, Coilin interacts with Ku proteins and inhibits in vitro non-homologous DNA end joining. *FEBS letters* 584, 4735 (Dec 1, 2010).


