

Aus dem Fachbereich Medizin
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

betreut am
Gustav Embden-Zentrum der Biochemie
Institut für Biochemie II (Kardiovaskuläre Biochemie)
Direktor: Prof. Dr. Ivan Dikic

**Characterization of SPRTN, the first mammalian metalloprotease
that repairs DNA-protein-crosslinks**

Dissertation
zur Erlangung des Doktorgrades der Medizin
des Fachbereichs Medizin
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

vorgelegt von
Stefan Prgomet, dr. med.

aus Düsseldorf

Frankfurt am Main, 2019

Dekan:	Prof. Dr. Stefan Zeuzem
Referent/in:	Prof. Dr. Ivan Dikic
Korreferent/in:	Prof. Dr. Jörg Trojan
Tag der mündlichen Prüfung:	12.05.2020

Contents

1 Summary / Zusammenfassung	6 / 8
2 Introduction	10
2.1 SPRTN regulates a progeria and tumorigenesis axis	10
2.1.1 RJALS-syndrome as the clinical manifestation of SPRTN malfunction	10
2.1.2 SPRTN acts in DNA damage repair	11
2.1.3 SPRTN's essential function remains unknown	13
2.2 DNA damage response (DDR)	16
2.2.1 Types of DNA-lesions	16
2.2.2 DNA damage repair mechanisms	18
2.2.3 Proposed SPRTN function in DNA damage response	23
2.3 DPC and DPC-repair mechanisms	26
2.3.1 Types and genesis of DPCs	26
2.3.2 DPC-repair	29
2.3.3 Proposed SPRTN function in DPC-repair	33
3 Aim	35
4 Results	36
4.1 SPRTN has a proteolytic function	36
4.1.1 SPRTN is a protease with an autocleavage activity	37
4.2 SPRTN interacts with DNA-bound proteins	40
4.2.1 SPRTN is recruited to DNA by DPCs	41
4.2.2 Recruitment of SPRTN to DNA is affected by its proteolytic activity	42
4.3 SPRTN cleaves DNA-protein-crosslinks (DPCs)	44
4.3.1 DPC- removal is dependent on SPRTN	44
4.3.2 SPRTN's proteolytic activity is relevant for DPC-removal	45
4.3.3 Loss of function of SPRTN impairs DPC-removal	48
4.3.4 SPRTN is a promiscuous protease.....	50
4.4 SPRTN is relevant for cellular resistance to DPC-s	53
4.4.1 SPRTN deficiency sensitizes cells towards DPC-causing agents	53
4.4.2 SPRTN facilitates DPC-resistance and enables cell proliferation.....	54

4.5	SPRTN's C-terminus is dispensable for DPC-cleavage but relevant for regulation of SPRTN	57
4.5.1	SPRTN-level is replication-coupled	57
4.5.2	The C-terminus determines the nuclear localization of SPRTN	58
4.5.3	DPC-cleavage is contained in the SprT-domain	61
4.5.4	The C-terminus is important for SPRTN-SPRTN interaction	62
4.6	RJALS-syndrome shows defective DPC-cleavage	64
5	Discussion	68
5.1	Structure and protein interaction of SPRTN	69
5.2	Relation to the supposed yeast orthologue Wss1	74
5.3	Activities of SPRTN and their cellular functions	77
5.3.1	SPRTN exhibits an autocleavage activity	77
5.3.2	SPRTN's main function is the proteolytic cleavage of DPCs	77
5.3.3	SPRTN's proteolytic function is important for DNA-replication	79
5.3.4	SPRTN has an important role in cell cycle regulation	82
5.4	Review of regulation of SPRTN activity	83
5.4.1	SPRTN is activated upon DNA-binding	83
5.4.2	SPRTN's function depends on DNA synthesis	84
5.4.3	The influence of interacting proteins on SPRTN function	85
5.4.4	SPRTN dimerization	87
5.4.5	The impact of autocleavage on SPRTN activity	88
5.4.6	The cell cycle-dependent regulation of SPRTN	90
5.5	Explanation of the clinical findings in RJALS-syndrome	91
5.6	Future perspectives	94
5.6.1	SPRTN as a DNA damage response regulator	94
5.6.2	The role of SPRTN in tumorigenesis and cancer therapy	94
6	Materials and Methods	97
6.1	Cellular biology techniques	98
6.2	Molecular biology techniques	105
6.3	Biochemical techniques	109

7	References	117
8	Abbreviations	146
9	Acknowledgements	155
10	Curriculum vitae	156
11	Schriftliche Erklärung	160

1 Summary

DNA is constantly exposed to various endogenous and exogenous sources causing different kinds of DNA damage. To overcome this threat, cells have evolved various repair mechanisms. Impairments of these repair mechanisms result in diverse diseases. Ruijs-Aalfs syndrome is a monogenic disease characterized by accelerated ageing and carcinogenesis, typical features of impaired DNA repair and was shown to be caused by germline mutations of SPRTN, a newly identified and only partially understood protein. A role of SPRTN in DNA damage response was previously shown and an involvement in translesion synthesis (TLS) proposed. However, later discoveries revealed an essential function of SPRTN, being indispensable for embryonic development of vertebrates and cellular survival, whereby this function is independent of SPRTN's proposed function in TLS. The essential function of SPRTN was proposed to be contained in its protease domain but remained unclear.

In this study we identify SPRTN as the first mammalian metalloprotease that repairs DNA-protein-crosslinks (DPCs). DPCs represent a specific type of DNA-lesions with bulky protein adducts covalently linked to DNA thereby being highly toxic as they potentially stall replication forks and lead to double strand breaks and genomic instability. DPC-repair remains only partially understood despite their frequent appearance and toxicity. With this study we discover and characterize a new mechanism of DPC-repair in mammalian cells - a proteolytic cleavage of the protein adduct by the metalloprotease SPRTN. Accordingly, a proteolytic activity of SPRTN is demonstrated and SPRTN-recruitment to DNA upon DPC-induction displayed. Furthermore, SPRTN exhibits degradation of different proteins covalently bound to DNA in form of DPCs, but not of unbound fractions of the same protein substrates. Consequently, mutations of SPRTN's proteolytic core as well as a mislocalization or depletion of SPRTN result in impaired DPC-repair. The importance of SPRTN-mediated DPC-removal is confirmed by a severely compromised response to DPC-inducing agents for cells with impaired SPRTN function. Additionally to the discovery of SPRTN's essential function this study further provides an explanation of the molecular mechanism underlying Ruijs-Aalfs

syndrome (RJALS), the segmental progeroid syndrome resulting from SPRTN mutation. The effects of the identified clinical mutations on the DPC-repair function of SPRTN are explained and a DPC-accumulation in cells carrying clinical SPRTN-mutation displayed. The obtained data provides sufficient evidence that an impaired DPC-repair is the pathophysiologic cause of RJALS-syndrome, confirming the importance of SPRTN's newly identified function. In conclusion, SPRTN is the first identified mammalian metalloprotease with a DPC-repairing function and the impairment of SPRTN-mediated DPC-removal is the underlying mechanism of RJALS syndrome.

1 Zusammenfassung

DNA ist ständig verschiedenen endogenen und exogenen Quellen ausgesetzt, die unterschiedliche Arten von DNA-Schäden verursachen. Um diese Bedrohung zu überwinden, haben Zellen verschiedene Reparaturmechanismen entwickelt. Beeinträchtigungen dieser Reparaturmechanismen führen zu verschiedenen Krankheiten. Das Ruijs-Aalfs-Syndrom ist eine monogene Erkrankung, die durch beschleunigtes Altern und Tumorentwicklung gekennzeichnet ist, typische Merkmale einer gestörten DNA-Reparatur, und es wurde gezeigt, dass sie durch Keimbahnmutationen von SPRTN, einem neu identifizierten und nur teilweise verstandenen Protein, verursacht wird. Eine Rolle von SPRTN bei der DNA-Schadensantwort wurde bereits gezeigt und eine Beteiligung an der Translesionssynthese (TLS) vorgeschlagen. Spätere Entdeckungen ergaben jedoch, dass SPRTN eine essentielle Funktion hat, die für die embryonale Entwicklung von Wirbeltieren und das zelluläre Überleben unverzichtbar ist, wobei diese Funktion unabhängig von der von SPRTN für TLS vorgeschlagenen Funktion ist. Für die essentielle Funktion von SPRTN wurde vermutet, dass sie in seiner Protease-Domäne enthalten ist, sie verblieb jedoch unklar.

In dieser Studie identifizieren wir SPRTN als die erste Metalloprotease in Säugetieren, die DNA-Protein-Crosslinks (DPCs) repariert. DPCs stellen einen spezifischen Typ von DNA-Läsionen dar, bei denen sperrige Proteinaddukte kovalent an DNA gebunden sind, wodurch sie hochtoxisch sind, da sie möglicherweise Replikationsgabeln blockieren und zu Doppelstrangbrüchen und genomischer Instabilität führen. Die Reparatur von DPCs ist trotz ihres häufigen Auftretens und ihrer Toxizität nur teilweise verstanden. Mit dieser Studie entdecken und charakterisieren wir einen neuen Mechanismus der DPC-Reparatur in Säugetierzellen - eine proteolytische Spaltung des Proteinaddukts durch die Metalloprotease SPRTN. Dementsprechend wird eine proteolytische Aktivität von SPRTN gezeigt und eine SPRTN-Rekrutierung zur DNA nach DPC-Induktion gezeigt. Darüber hinaus zeigt SPRTN einen Abbau verschiedener Proteine, die in Form von DPC kovalent an DNA gebunden sind, jedoch nicht ungebundener Fraktionen derselben Proteinsubstrate. Folglich führen Mutationen

des proteolytischen Kerns von SPRTN sowie eine Fehllokalisierung oder Depletion von SPRTN zu einer Beeinträchtigung der DPC-Reparatur. Die Bedeutung der SPRTN-vermittelten DPC-Entfernung wird durch eine stark beeinträchtigte Reaktion auf DPC-induzierende Mittel von Zellen mit beeinträchtigter SPRTN-Funktion bestätigt. Zusätzlich zur Entdeckung der essentiellen Funktion von SPRTN liefert diese Studie eine Erklärung des molekularen Mechanismus, der dem Ruijs-Aalfs-Syndrom (RJALS) zugrunde liegt, dem segmentalen Progeroid-Syndrom, das aus der SPRTN-Mutation resultiert. Die Auswirkungen der identifizierten klinischen Mutationen auf die DPC-Reparaturfunktion von SPRTN werden erläutert und eine DPC-Akkumulation in Zellen mit klinischer SPRTN-Mutation dargestellt. Die erhaltenen Daten liefern ausreichende Hinweise dafür, dass eine beeinträchtigte DPC-Reparatur die pathophysiologische Ursache des RJALS-Syndroms ist, was die Bedeutung der neu identifizierten SPRTN-Funktion bestätigt. Zusammenfassend ist SPRTN die erste identifizierte Metalloprotease in Säugetieren mit einer DPC-Reparaturfunktion, und die Beeinträchtigung der SPRTN-vermittelten DPC-Verarbeitung ist der zugrunde liegende Mechanismus des RJALS-Syndroms.

2 Introduction

2.1 SPRTN regulates a progeria and tumorigenesis axis

The identification of compromised molecular mechanisms underlying diseases is becoming a growing concept in identifying and understanding cellular processes on one side as well as pathogenesis and potential therapeutic targets for diseases on the other side. Monogenic diseases represent an ideal opportunity to make impactful scientific contributions since they enable a direct connection from a dysfunctional protein to the clinical symptoms. In this way a newly found protein named SPRTN, with only partial previous characterization, was identified to be mutated in a new clinical syndrome and thereby found to have a more important function than previously expected (Fig 1). This main function of SPRTN, essential for genetic integrity, needed to be further characterized.

2.1.1 RJALS-syndrome as the clinical manifestation of SPRTN malfunction

Monogenic disorders are rare diseases characterized by a germline mutation affecting a single gene¹. Hereditary diseases are often classified into groups due to similar clinical presentations. Interestingly, syndromes from the same group are often caused by impairments of connected underlying cellular mechanisms^{2,3}. As an example, a large proportion of progeroid syndromes results from defective DNA-repair mechanisms⁴.

Progeroid syndromes are a class of hereditary diseases characterized by accelerated aging symptoms (name derived from the Greek word meaning “prematurely old”). A classic representative is the Hutchinson-Gilford Progeria Syndrome (HGPS) also known as childhood-progeria, a rare autosomal dominant disease caused by a mutation in the LMNA gene encoding Lamin A⁵. Lamin A is a structural protein supporting the nuclear envelope. Another classical progeroid syndrome is Werner Syndrome, also known as adulthood-progeria (due to its later appearance), an autosomal recessive disorder resulting from a mutation in the WRN gene⁶. Common WRN mutations result in a truncated and impaired WRN helicase activity that is required for repair of DNA double strand breaks⁷.

Classical symptoms of Werner syndrome are short stature, premature cataracts, skin changes similar to scleroderma and hair graying or loss⁶.

There are a several more diseases distinct from the above mentioned due to different underlying mutations but all manifest in similar and at least partially overlapping symptoms of premature aging often with carcinogenesis as well⁸. These diseases, similar but not identical to Werner Syndrome, are called segmental progeroid syndromes and they typically result from defective DNA damage repair⁹. Ruijs et al. described in 2003 a boy of Moroccan origin with an atypical progeroid syndrome and hepatocellular carcinoma, who died with 17 years (Fig. 2)¹⁰. Molecular analysis excluded the clinically supposed Werner syndrome in this case. Therefore, they proposed a new syndrome with a possible mutation in a helicase gene and after the first identifiers, the disease was named Ruijs-Aalfs syndrome (RJALS). Approximately 11 years later, two patients from another family of Australian origin with similar clinical manifestations were identified and described in Lessel et al⁹. A genome-wide linkage analysis with additional exome sequencing of the patients revealed biallelic exome mutations in the SPRTN gene in all three individuals.

2.1.2 SPRTN acts in DNA damage repair

SPRTN, which is also known as C1orf124 or DVC1, was first mentioned in 2012 as a protein with an SprT-like domain at its N-terminus and was therefore named Spartan or SPRTN. SprT is a protein found in *Escherichia coli* and the domain-family was shown to bind zinc and supposed to have protease activity.

SPRTN is a 489-amino-acid protein encoded in chromosome 1 (1q42.2) and encompasses multiple domains¹¹. In addition to the mentioned SprT-like protease domain (amino acids 45–212) it is endowed with a ubiquitin-segregase p97 (VCP16)-interacting motif named SHP (amino acids 253-261), a proliferating cell nuclear antigen (PCNA) interacting box, known as a PIP-box (amino acids 325-332) and an ubiquitin-binding zinc finger domain, known as UBZ4 (amino acids 453–475) (Fig. 1)¹²⁻¹⁴.

The first analysis of SPRTN in 2012 revealed the PCNA-binding (PIP box) and the ubiquitin-binding domains (UBZ), which implicated an involvement of SPRTN in the DNA damage response (as ubiquitylated PCNA is primarily connected to DNA repair)^{12,15}. SPRTN was proposed as a recognizer of ubiquitylated PCNA, thereby supporting post-replication DNA damage repair. The initial results suggested an important role in promoting translesion DNA synthesis (TLS), a mechanism enabling DNA-replication in cases of DNA damaged with ultraviolet-irradiation^{12-14,16}.

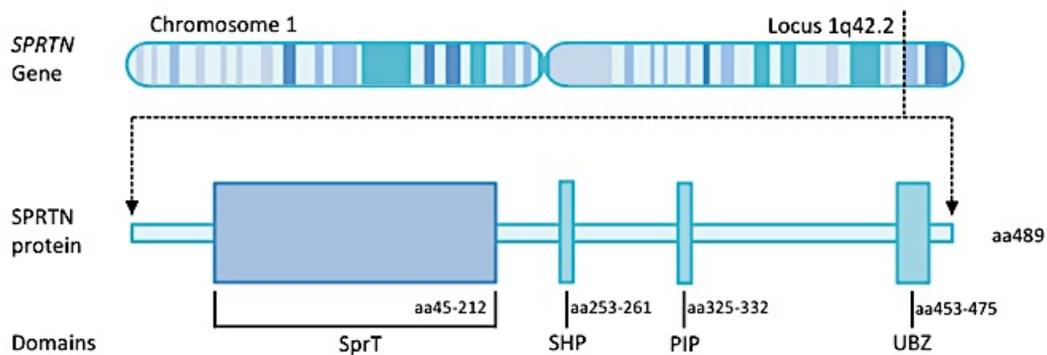


Figure 1 SPRTN's gene location, amino-acid sequence and domain structure. SPRTN is encoded on Chromosome 1 (locus 1q42.2) and is built up of 489 amino-acids. At the beginning of this study 4 different domains have been identified including the N-terminal SprT-domain with a supposed proteolytic function and 3 interacting motifs binding to p97 (SHP), PCNA (PIP) and Ubiquitin (UBZ) respectively.

At the start of our work there were 9 published studies dedicated to SPRTN with different, partially opposing interpretations of its function in the DNA damage response. Details about these findings and SPRTN's supposed function in DNA damage repair will be explained later. Briefly, SPRTN was shown to play a role in providing cells resistance towards different DNA-damage inducing agents like ultraviolet (UV) radiation, Hydroxyurea (HU), Cisplatin or Camptothecin and was primarily studied in response to UV-radiation^{14,17,18}. Several studies showed that SPRTN forms DNA-damage induced foci co-localizing with ubiquitylated PCNA and facilitates the recruitment of TLS polymerase Pol η , thereby functioning in TLS^{12-14,16}. Later studies revealed an interaction with the segregase p97 and suggested that SPRTN recruits p97 to exchange Pol η with Pol ϵ , thereby ending TLS¹⁷⁻¹⁹. Irrespective of the mentioned debate about SPRTN's function in TLS, the

function of SPRTN's N-terminal SprT-like domain (amino acids 45–212) prevailed unclear.

2.1.3 SPRTN's essential function remains unknown

The importance of SPRTN function was first sufficiently described in 2014 with a mouse model²⁰. It was shown that homozygous SPRTN-knockout (KO) mice die at a very early stage of embryonic development (embryonic day 3.5) and give no viable offspring. This proved that SPRTN has a more profound function than initially supposed, showing an essential function (at least in higher vertebrates) since its deletion causes early embryonic lethality. Interestingly, a different mouse model, hypomorphic for SPRTN function, displayed small body size, lordokyphosis, cachexia and cataracts, which all are characteristics typical of progeroid syndromes. A comparison of hypomorphic SPRTN mice with the Ruijs-Aalfs syndrome patients carrying mutations in the SPRTN gene and progeroid phenotypes exhibits crucial similarities determining an evolutionary conserved function of this protein, further implying an important role of SPRTN (Fig. 2B)^{10,20}.

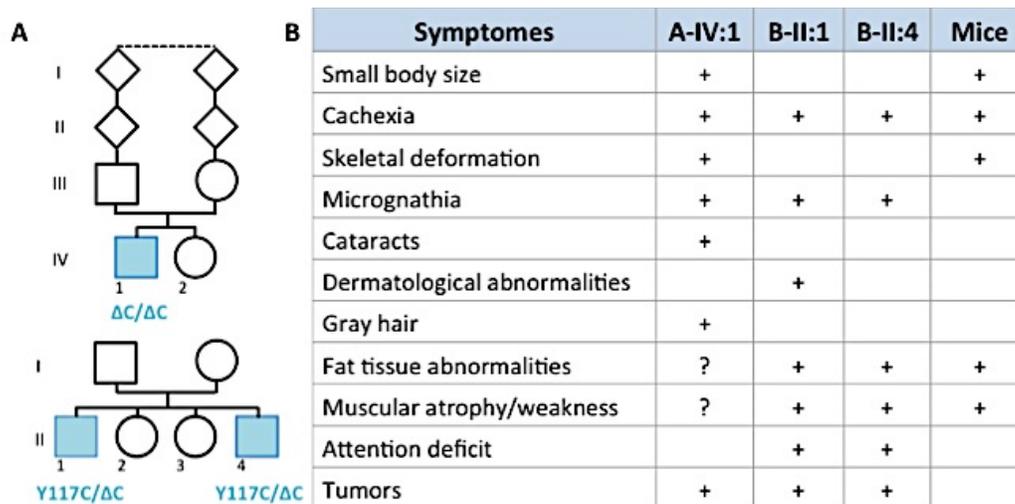


Figure 2 RJALS-syndrome scheme showing heritage and symptoms. **(A)** Family diagrams for the 3 patients with found RJALS-syndrome from 2 families. Patients expressing RJALS-symptoms are colored in blue with the underlying SPRTN-mutations underneath the symbols. The first identified patient carries a homozygous $\Delta C/\Delta C$ -mutation and is named A-IV:1, the other two patients have heterozygous Y117C/ ΔC -mutations and are named B-II:1 and B-II:4. In the family of patient A-IV:1 a consanguinity of unknown degree was found, marked in the figure by an interrupted bound between the

ancestors. (Figure adapted from Lessel et al.2014). **(B)** Comparison of symptoms found in mice with a hypomorphic level of SPRTN-expression and in the 3 RJALS-patients reveals analogous progeroid phenotypes (not all symptoms are listed, figure based on Lessel et al. 2014 and Maskey et al. 2014).

A detailed analysis of SPRTN-mutations in the individuals leads to an important observation. One boy had a homozygous point mutation leading to a premature stop codon resulting in a truncated version of SPRTN missing the C-terminal part (Δ^C SPRTN)⁹. Taking into account the fact that the previously characterized functions of SPRTN depend on its PCNA-, Ubiquitin- and p97-interacting motifs, which are all located in its C-terminal half (missing in the patients truncated Δ^C SPRTN form), suggests that PCNA, p97 or Ubiquitin binding activities of SPRTN are dispensable for human development and raises the possibility that SPRTN's essential function is contained within its proteolytic core (Fig 3)^{9,12-14,16-18}.

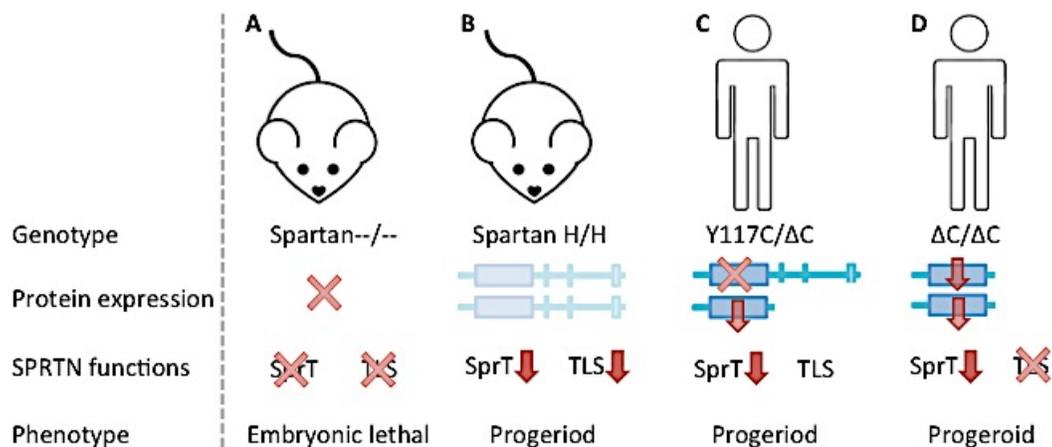


Figure 3 Scheme of SPRTN genotype, protein expression and resulting phenotype observed in patients and mice. Supposed SPRTN-function is demonstrated as the known TLS-function depending on the C-terminus of the protein and a unknown function contained in the SprT-domain. **(A)** Mice with a complete SPRTN-KO are embryonic lethal. **(B)** Mice hypomorphic for SPRTN show a progeroid phenotype. **(C)** Patients B-II:1 and B-II:2 with a weakened function of the SprT-domain and a heterozygously conserved normal TLS-function (since equal heterozygous TLS-functions are found in asymptomatic relatives) exhibits a progeroid phenotype similar to the observations in mice. **(D)** Patient A-IV:1 with a homozygous Δ C/ Δ C-mutation demonstrates a weakened function of the SprT-domain due to a mislocalization (for details see main text) but is completely lacking TLS-function resulting from the missing C-terminus. Since a complete loss of SPRTN is embryonic lethal in mice and a complete loss of SPRTN's TLS function in human is compatible with life, the essential function of SPRTN is supposedly contained in its SprT-domain.

The importance of SPRTN's proteolytic function is further implied by the other mutation found in RJALS-patients (^{Y117C}SPR TN) since the mutation lies in close proximity to the catalytic core of the protease domain^{9,19}. While individuals carrying only a heterozygous ^{ΔC}SPR TN-mutation remained symptom-free, an additional mutation of the protease domain of the remaining SPR TN gene lead to an inhibition of SPR TN's function, exhibited by severe clinical symptoms. The mentioned observations propose that the SprT-like domain carries the essential function of SPR TN, as its complete loss is not compatible with live in vertebrates (in contrast to the missing C-terminus) while its malfunction leads to serious defects. This proteolytic function remains uncharacterized but is supposed to be connected to DNA damage response since its malfunction leads to carcinogenesis and progeria, symptoms related to an impairment of DNA-damage-repair.

2.2 DNA damage response (DDR)

DNA is constantly exposed to potentially harmful endogenous and exogenous agents²¹. Furthermore, physiological processes like DNA-replication can cause damage to DNA endogenously e.g. by mispairing, deletion or insertion of nucleotides²²⁻²⁴. The recognition of damaged DNA and its repair is essential for ensuring genetic integrity and thereby continuation of life²⁵. Therefore, cells have developed different mechanisms to repair or tolerate DNA damage. These mechanisms are collectively known as DNA damage response (DDR). The importance of these processes is best presented by different diseases caused by impairments of DDR-mechanisms showing severe to lethal effects²⁶. For instance, mutations of genes included in DDR-pathways sensitize cells towards different exogenous agents or cause spontaneous mutations leading to cell death or uncontrolled proliferation. This results in cancer development, premature aging, neurodegenerative diseases, and early death^{27,28}. In some cases, impaired DDR-mechanisms are even not compatible with life²⁶. Due to the importance of DDR, a magnitude of different studies over decades has resulted in a comprehensive understanding of DNA damage repair mechanisms although a lot of questions prevail unsolved.

2.2.1 Types of DNA lesions

DNA can be damaged by endogenous sources (through reactive metabolites like reactive oxygen species (ROS) or as errors during replication) or by exogenous agents (including physical agents as ultraviolet/UV or ionizing radiation/IR and chemical agents as aromatic amines, alkylating agents, toxins as well as environmental stress as extreme heat or hypoxia)^{29,30}. The distinct types of DNA damage can be distinguished by their structure (Fig 4).

Deaminated bases (bases without their exocyclic amine) appear spontaneously during DNA-processes especially on single-stranded DNA (ssDNA)²¹. Abasic or apurinic/apyrimidinic (AP) sites are formed by hydrolysis of the base from the sugar-phosphate backbone leaving this part of DNA without a base³¹. DNA bases can become oxidized (8-oxoguanine) as a result of ROS, which can additionally

lead to breaks of the DNA-strand (explained later). Similarly, DNA bases can be alkylated (O⁶-methylguanine) by alkylating agents (as in tobacco or chemotherapeutics) or by normal methylation reactions^{29,30,32}.

Further types of DNA damage include mismatched base pairs (where an error during replication puts a non-matching nucleotide opposite to the original one) or UV-induced pyrimidine dimers^{22-24,33}. An especially dangerous type of DNA damage are breaks of one or both sugar phosphate backbone strands named single strand breaks (SSB) or double strand breaks (DSB) respectively, which are usually caused by IR, ROS or stalled DNA replication. A distinct type of DNA lesions are crosslinks of DNA to diverse molecules^{34,35,36}. As an example, a covalent bond can be formed between DNA strands producing interstrand crosslinks (ICLs). Accordingly, DNA can covalently bind to proteins forming DNA-protein-crosslinks (DPCs), which will be addressed later^{37,38}. Since there are diverse types of DNA damage, cells developed different repair mechanisms specifically acting on each type of lesion.

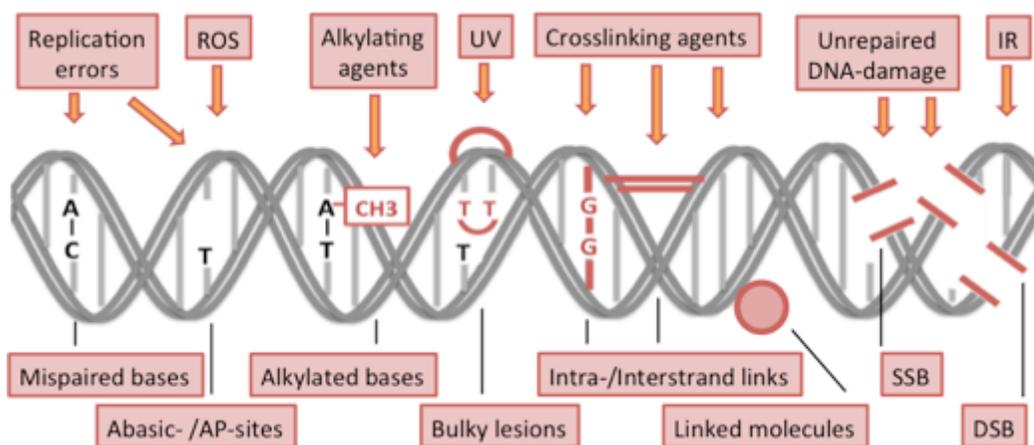


Figure 4 DNA-damage types and causing agents. **(Lower part)** Different types of DNA-lesions range from mismatched lesion between two DNA strands up to breaks of one or both sugar-phosphate backbones. **(Upper part)** Various endogenous and exogenous sources causing DNA-damage. The causes of different DNA-lesions are demonstrated by classical damaging agents, but the causal relations are more complex with one agent leading to various sorts of DNA-lesions simultaneously or successively.

2.2.2 DNA damage repair mechanisms

Repair of damaged bases

The simplest DNA damage repair mechanism is a reversal of the base damage as in the case of alkylated bases (Fig 5). The added alkyl group is released from O-alkylated DNA bases to the repair enzyme, O⁶-alkylguanine-DNA alkyltransferase (AGT) or O⁶-methylguanine-DNA methyltransferase (MGMT), an expensive process as one enzyme can irreversibly repair only one alkylated base³⁹. Similarly, N-alkylated bases are repaired by ALKBH1-8 (Alkylation repair homologues) and FTO (Fat mass and obesity associated) enzymes^{40,41}.

Base excision repair (BER) of damaged bases

Damaged DNA-bases, which cannot be repaired by a reverting enzyme but do not lead to structural distortions of the DNA helix, are excised by 11 different glycosylases thereby leaving an abasic (apurinic/aprimidinic or AP) site, which is further prepared by an Apurinic/aprimidinic AP endonuclease (APE1) and filled in with a nucleotide by Polymerase Pol β ⁴²⁻⁴⁵. This nucleotide is then ligated to the sugar-phosphate backbone by DNA ligase (LIG1) or complex of LIG3 and XRCC1⁴⁶⁻⁴⁸. Alternatively, the transformed abasic site of more than one single base (mostly 2-10) is filled the canonical replication machinery (Polymerase Pol δ/ϵ with PCNA) and further processed by a flap endonuclease (FEN1) and LIG1⁴⁹⁻⁵².

Nucleotide excision repair (NER) of bulky damages

If the base damage causes a distortion of the DNA helix (like UV-induced thymidine dimers), not only the base but a short stretch of nucleotides surrounding the damage is excised with a following re-synthesis and ligation. The lesion is recognized either due to the structural distortion by the heterodimer XPC-RAD23 (Xeroderma pigmentosum group C and Restriction site associated DNA-marker 23) - a pathway named global genome NER (GG-NER) - or due to a stalled RNA-polymerase II transcription by CSA and CSB proteins (Cockayne syndrome A or B) - a pathway named transcription coupled NER (TC-NER)⁵³⁻⁵⁶. After recognition

of a distorted DNA-helix, the TFIIH-complex (Transcription factor II human) is recruited and unwinds roughly 30 surrounding nucleotide pairs (XPB and XPG), incises the sugar-phosphate backbone (XPF and XPG) and re-synthesizes the gap (Pol $\delta/\epsilon/\kappa$ with PCNA, Replication factor C/RFC, and LIG1 or XRCC1-LIG3) by a similar principle as in the BER pathway⁵⁷⁻⁶⁰.

Mismatch repair (MMR)

DNA damage resulting from mismatched bases during DNA replication is typically repaired post-replicatively. MutS (Mutator S) complexes recognize the mismatch with the help of MutL complexes, which incise the new strand⁶¹⁻⁶⁴. The newly synthesized daughter strand containing the wrong base is differentiated from the parental strand by the nicks inside the sugar-phosphate backbone. At these nicks the DNA-strand is excised by EXO1 (Exonuclease 1), the ssDNA stabilized by RPA (Replication protein A) and further re-synthesized and ligated by the canonical machinery including Pol δ , PCNA, RFC (Replication factor C), HMGB1 (High mobility group box 1) and LIG1 as in the previously explained BER and NER pathways⁶⁵⁻⁶⁷.

Single strand break repair

The DNA sugar-phosphate backbone can be disrupted on one or both strands, thereby forming single strand-breaks (SSB) or double-strand breaks (DSB) respectively. In the case of SSB, the DNA backbone nick can block replication or transcription and needs to be repaired. PARP1 (Poly ADP-ribose polymerase 1) is recruited to the lesion causing PARylation, followed by processing of the free ends by APE1, APTX (Aprataxin), PNKP (Polynucleotide kinase 3'-phosphatase) and FEN1 producing a gap of ssDNA further filled by Pol $\beta/\delta/\epsilon$ and ligated by LIG1 or 3⁶⁸⁻⁷⁰. A specific SSB is produced by a Topoisomerase I (Topo I) covalent complex and is processed by Tyrosyl-DNA phosphodiesterase 1 (TDP1) as explained later⁷¹.

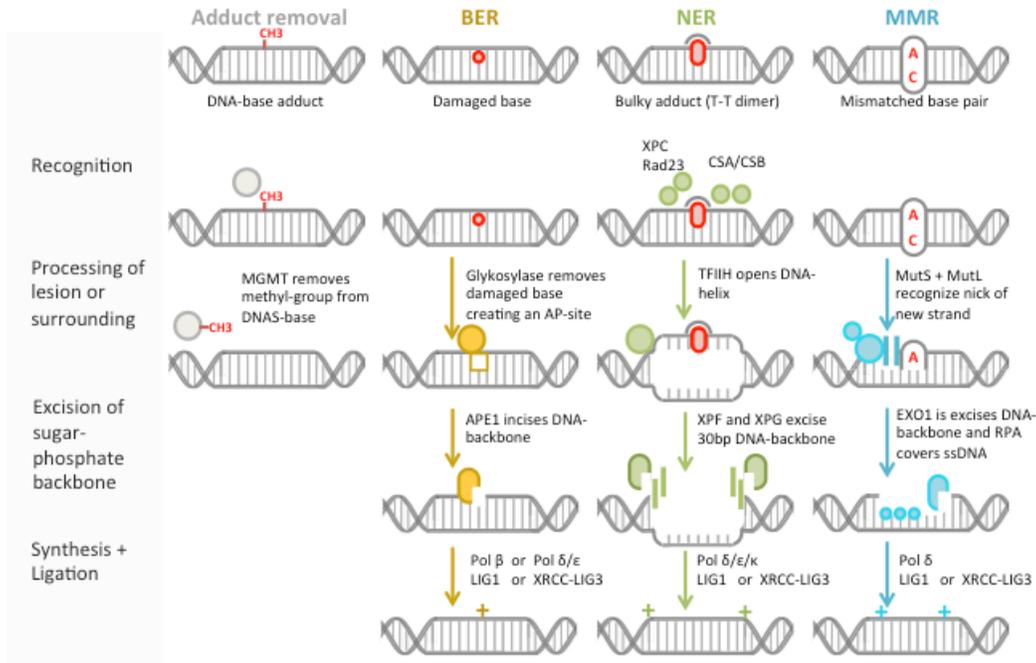


Figure 5 Scheme of DNA-damage repair mechanisms for single strand damage. A comparison of the main mechanisms reveals 4 crucial steps except the most simple adduct removal pathway consisting only of 2 steps. The 4 steps are shown in columns from top to bottom for each mechanism. DNA-damage is shown by red marks. The repair factors are symbolically demonstrated and their names mentioned in the lines for each step. (For details see main text).

Double strand break repair

Double strand breaks represent highly dangerous DNA lesions as they completely destroy the integrity of the DNA double helix and cause severe genomic instability⁷². DSBs are repaired by two major pathways, namely homologous recombination (HR) and non-homologous end joining (NHEJ) (Fig6)

73

Homologous recombination (HR)

Homologous recombination (HR), relies on a homologous DNA strand, found in the sister chromatid⁷⁴. The process is initiated by the MRN-complex activating ATM (Ataxia-teleangiectasia mutated), with further phosphorylation and ubiquitylation of Histone H2AX, thereby recruiting 53BP1 (p53-binding protein 1) and BRCA1 (Breast cancer type 1 susceptibility protein 1)⁷⁵. BRCA1 ubiquitylates CtIP, also known as RBBP8 (RB binding protein 8), which together with MRN

provides nucleolytic cleavage of the free 5'-ends of the DSB thereby producing ssDNA-ends^{76,77}. These ends are covered by RPA, prolonged by EXO1 and then coated by RAD51 in a BRCA2-dependent manner^{78,79}. The RAD51 nucleoprotein chain enables the search for a homologous DNA sequence and invasion of this sequence on the sister chromatid, which serves as a template for synthesis and further extension of the ssDNA⁸⁰⁻⁸². Thereby the 3'-end is extended on the template in a D-loop formation. The process is continued further by two pathways. In a pathway called SDSA (synthesis dependent strand annealing) the synthesis stops soon and the D-loop is dissolved by RTEL1 (Regulator of telomerase elongation helicase 1)⁸³. In another pathway called DSBR (double strand break repair) a Holliday junction is formed, whereby the complementary strand from the DSB-containing chromatid gets bound to its complementary strand from the sister chromatid and also gets further synthesized. This Holliday junction is resolved afterwards by BLM-Topo III-RMI1-RMI2 complex, GEN1, MUS81-EME1, and SLX1-SLX4 causing in most cases a crossover between two sister chromatids⁸⁴.

Non-homologous end joining (NHEJ)

Non-homologous end joining (NHEJ) is a less accurate than homologous recombination, as it is performed without a homologous template and simply binds the free DSB-ends together^{85,86}. The first step is the recruitment of the Ku70/Ku80-complex, which binds XRCC4 (X-ray repair cross-complementing protein 4) to stabilize the ends and activates DNA-PKcs (DNA-dependent protein kinase, catalytic subunit)⁸⁶⁻⁸⁹. In case of incompatible DSB-ends further end-processing is performed by the Artemis-complex leaving naked strands followed by later gap filling by Polymerases Pol λ or Pol μ and subsequent ligation by LIG4/XRCC4 and XLF (XRCC4-like factor)⁹⁰⁻⁹³.

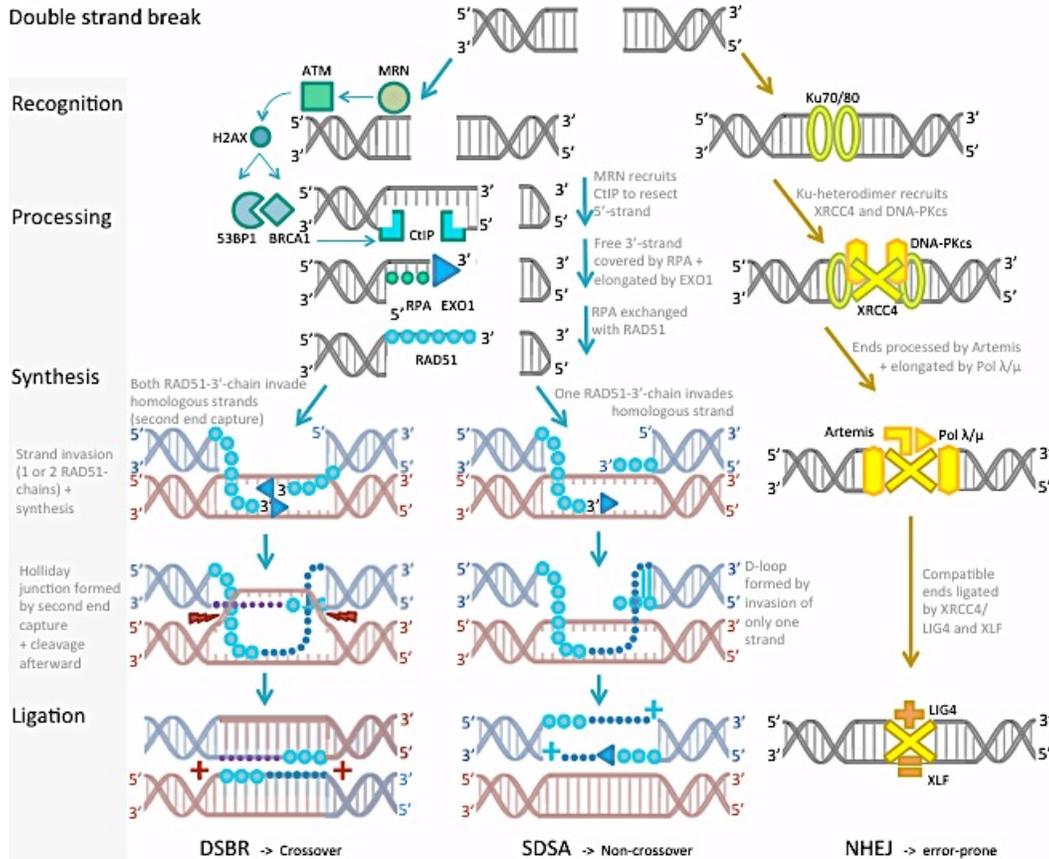


Figure 6 Scheme of mechanisms for repair of double-strand breaks (DSBs). Breaks are either repaired by HR or more error-prone NHEJ (non-homologous end joining). **(Left part)** Homologous repair includes degradation of the 5'-end in order to produce ssDNA, which gets covered with RAD51 and invades a homologous DNA sequence. If both RAD51-ssDNA-chains invade a homologous chromosome, a Holliday junction is formed and cleaved leading to an exchange of DNA, known as crossover. **(Middle part)** In case of invasion by a single RAD51-ssDNA-chain a D-loop is formed and after prolongation of the strand and recognition by the ssDNA-strand of the other side of the break it is annealed to it. **(Right part)** Non-homologous end joining is performed by annealing (optionally preceded by processing of incompatible ends) without a homologous sequence as a template and thereby more error-prone.

Interstrand crosslink (ICL) repair

As ICLs represent a specific type of DNA damage covalently binding two DNA strands, their repair includes several canonical repair pathways explained in other chapters^{94,95}. Briefly, Fanconi anemia proteins (a big family of 21 protein important for crosslink resistance and mutated in Fanconi anemia) are recruited to the ICL, assemble and mono-ubiquitylate the ID complex (FANCI-FANCD2), which initiates incision of the DNA strand by endonucleases as XPF-ERCC1 to unhook the

crosslink⁹⁶⁻⁹⁹. The re-synthesis of the newly formed gap is performed by translesion synthesis (TLS) and the remaining ICL hook is later removed by the NER pathway^{98,100}. The repair mechanism differs between replicating and non-replicating cells. One side of the ICL is repaired as mentioned, while the other side is simultaneously processed into a DSB, which is further repaired by MRN, RAD51 and homologous recombination (HR)^{10,102}. During transcription the NER pathway is important for recognition and initial processing.

Translesion synthesis

The synthesis of DNA upon DNA damage is necessary to prevent replication stalling and highly toxic DNA breaks¹⁰³. Therefore, cells developed a DNA-synthesis mechanism that is able to bypass lesions but also with less accuracy, thereby potentially causing mutations¹⁰⁴. The translesion synthesis (TLS) is performed by 11 polymerases characterized by a more open active site, less contact to DNA and a missing proofreading activity enabling a synthesis over damaged DNA or ssDNA but also making them more error-prone¹⁰⁵. These polymerases replace the canonical DNA polymerases upon collision of the replication fork with damaged DNA enabling fork progression¹⁰³. Additionally, TLS plays an important role as a part of other DNA damage repair mechanisms as mentioned before. TLS is primarily induced via mono-ubiquitylation of PCNA in a RAD6/RAD18-dependent manner. The poly-ubiquitylation of PCNA by RAD5 leads to template switching rather than TLS^{15,103}.

2.2.3 Proposed SPRTN function in DNA damage response

As already mentioned SPRTN was initially described as a factor important in translesion synthesis (TLS), although its exact function remains under debate (Fig 7)¹⁰³. SPRTN's relation to DNA is demonstrated by its ability to form foci primarily during S-phase and at replication sites, thereby colocalizing with PCNA¹⁶⁻¹⁸. Importantly, SPRTN-foci formation is increased upon DNA damage (caused by e.g. UV, microirradiation or MitomycinC/MMC) and colocalizes with DNA damage sites marked by γ H2AX^{13,14,16-18}. The importance of SPRTN in this process is

shown by an increased DNA-damage-sensitivity in SPRTN-depleted cells. Thereby several agents like UV, Methyl methanesulfonate (MMS), Camptothecin (CPT) and Cisplatin showed an increased toxicity, while other agents like IR or MMC showed no effect^{12,14,16-18}.

The recruitment of SPRTN to DNA damage sites depends on its ability to bind to PCNA and ubiquitin in accordance to a supposed function in TLS as ubiquitylated PCNA (Ub-PCNA) serves as a signal for the exchange of canonical DNA-replication machinery with TLS-polymerases^{12-15,17}. The binding of SPRTN to ubiquitin and PCNA was shown by mass spectrometry, pull down and co-immunoprecipitation and is impaired upon mutation of its PIP- or UBZ-domains thereby leading to an impaired UV-resistance for instance^{12-14,16-18}. Furthermore, a direct binding of SPRTN to Ub-PCNA and even a preference to Ub-PCNA in comparison to PCNA was shown^{16,17}. Additionally SPRTN is included in the recruitment of Ub-PCNA to DNA upon DNA damage (like UV), implying a feed-forward loop^{12,14,16}. Accordingly, RAD18, the main ligase for the ubiquitylation of PCNA is connected to SPRTN, and SPRTN-foci formation depends on RAD18 and Ub-PCNA^{12,13,16}. Furthermore, SPRTN and RAD18 act in the same pathway, at least in the case of UV-sensitivity^{12,14}. All this data shows that SPRTN is performing its DNA repair function upon binding to Ub-PCNA and thereby as a part of the TLS-mechanism.

However, the exact role of SPRTN in TLS is not resolved. Some studies suggest a model, in which SPRTN promotes recruitment of Polymerase η , one of the main TLS-polymerases, as SPRTN-depletion or inhibition of its binding to Ub-PCNA reduces Polymerase η -foci^{12,16}. A different model proposes that SPRTN resolves Pol η -foci via its interaction with the segregase p97, as SPRTN binds to this segregase (shown by mass spectrometry and co-immunoprecipitation) and further recruits p97 to DNA damage foci preventing prolonged Pol η -foci persistence^{14,17,18}. In this model the SPRTN-p97-mediated resolving of Pol η is inhibiting a prolonged DNA-synthesis by this error-prone protease. Thereby SPRTN is important for TLS through its prevention of potential TLS-induced mutagenesis. In conclusion, SPRTN is acting with TLS-machinery in repair of

certain DNA damage (like UV- or laser-induced lesions), whereby its function requires its C-terminally located PIP- and UBZ-box.

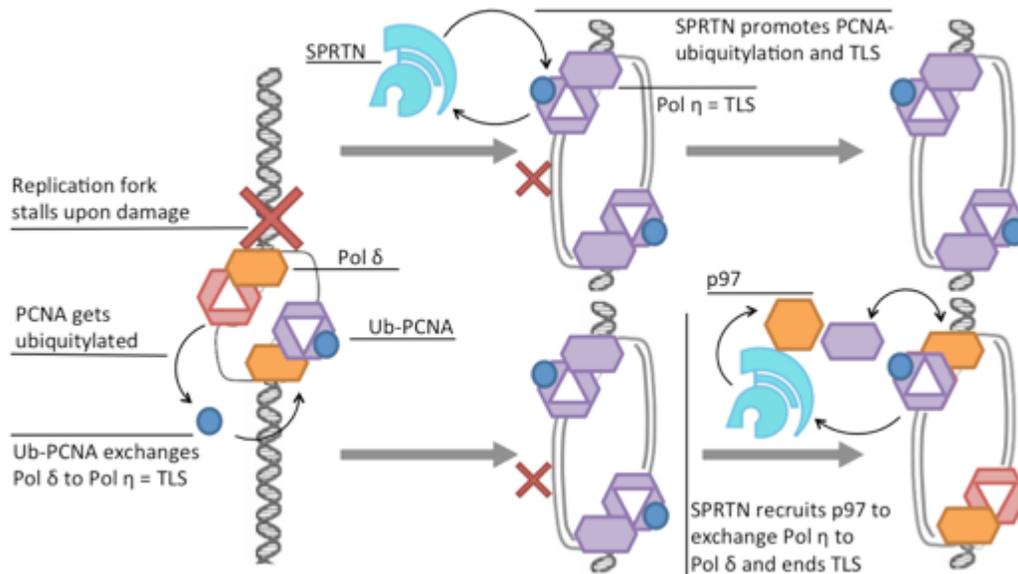


Figure 7 SPRTN's proposed function in Translesion synthesis (TLS). Replication fork arrests upon unresolved DNA-lesions, which cannot be replicated. Fork stalling initiates a RAD6/RAD18-mediated ubiquitylation of PCNA. Ub-PCNA serves as a signal for the exchange of the canonical DNA-Polymerases (Pol δ) with TLS Polymerases (Pol η). SPRTN was shown to be involved in TLS acting in two proposed distinct mechanisms (**Upper part**) SPRTN is recruited to stalled replication forks and promotes RAD18-mediated PCNA-ubiquitylation in a feed-forward loop thereby enabling access of TLS polymerases and stabilizing them. (**Lower part**) SPRTN recognizes Ub-PCNA and recruits the segregase p97, which subsequently removes the potentially error-prone TLS polymerases switching them with canonical DNA-Polymerases. Thereby SPRTN prevents a TLS-associated mutagenesis.

Despite the fact that SPRTN was shown to be implicated in TLS, since this function depends on its C-terminal part, the observed deficits in the patient carrying a clinical mutation of SPRTN (being alive although missing the C-terminus) cannot be explained by SPRTN's function in TLS. Therefore SPRTN has to have another, essential function in DNA damage repair, which is contained in its N-terminal part.

2.3 DPC and DPC-repair mechanisms

A specific sort of DNA damage distinct from the above-mentioned are so called DNA-protein-crosslinks (DPCs). DPCs are complexes of proteins that are covalently and irreversibly bound to DNA. They represent highly dangerous lesions, since they block almost all processes on DNA such as DNA replication or transcription, and additionally cannot be repaired by the canonical DNA repair mechanisms due to their bulkiness¹⁰⁶⁻¹⁰⁹.

2.3.1 Types and genesis of DPCs

Since DNA is a complex structure and constantly subjected to various processes, a high number of proteins (needed for chromatin structure maintenance as well as for DNA repair, replication or transcription) lie in close vicinity to DNA. These proteins are associated and constantly interacting with DNA thereby increasing the propensity of forming different transient connections. As a result, crosslinks between proteins and DNA can appear easily and frequently^{21,109}. This can occur either if transient connections of enzymes to DNA become permanent and irreversible, or if nonspecific proteins become covalently trapped on DNA by endogenous or exogenous sources. DPCs can therefore be divided in two categories by their origin: enzymatic and non-enzymatic DPCs (Fig 8)¹¹⁰.

Enzymatic DPCs

Many enzymes transiently form covalently linked intermediates with DNA while performing their enzymatic function^{111,112}. An example of this is Topoisomerase I (Topo I), an enzyme that prevents supercoiling of DNA. Topo I induces a temporary cut in one DNA-strand (single strand break or SSB) enabling it to rotate around the sister strand^{113,114}. In order to cut the DNA-strand, Topo I is transiently covalently linked to the 3'-end of the cut DNA-strand. Specific conditions like distortion by DNA damage prevent the re-ligation step of DNA and can be induced with compounds such as Camptothecin (CPT)¹¹⁵⁻¹¹⁷. Camptothecin can intercalate

within this connection. Both situations trap Topo I permanently to DNA making it a DPC¹¹⁸.

Another example is Topoisomerase II (Topo II) which similarly to Topo I causes a DNA break, but in this case of both strands (double strand break or DSB) in order to enable the DNA from the sister chromatin to pass through the DSB and thereby prevents supercoiling of the sister chromatids after replication¹¹⁹. As Topo II forms a transient link with the 5'-end of the cut DNA, it can also be permanently trapped to DNA by intercalating agents such as Etoposide (ETO), Anthracyclines (as Doxorubicin) or Anthraquinones (as Mitoxantrone) forming DPCs¹²⁰.

Although not forming a classical covalent link with DNA, PARP1 can also become closely bound to DNA while performing its function of PARylation to signal for the recruitment of repair enzymes. Known PARP-inhibitors like Olaparib (OLA) trap PARP1 to DNA so tightly that it forms a DPC-like structure¹²¹. It was shown that this trapping ability is more important for the effect of PARP1-inhibitors than the catalytic inhibition¹²¹. Furthermore, PARP1 can become covalently linked to abasic sites formed by BER¹²². Further examples of enzymes, which can form DPCs if their usual function is being interrupted are DNA Polymerase β or DNA-glycosylases if problems occur during BER^{110,111,123,124}.

Non-enzymatic DPCs

Not only enzymes which are transiently linked to DNA, but also many other proteins lying in close vicinity to DNA, without interfering with it, have been shown to form DPCs under certain circumstances^{38,124}. These crosslinks can be caused by exposure to exogenous radiation or toxins or even by endogenously produced reactive metabolites¹²⁵. Various reactive intermediates occur frequently due to the intensive molecular activity on or around DNA¹²⁵. As a result, DPCs show progressive accumulation in neuronal and heart tissue¹²⁶. The most common metabolic intermediates that induce DPCs are reactive aldehydes that are produced by processes such as sugar metabolism, lipid peroxidation or ethanol oxidation¹²⁷. An example is alcohol dehydrogenase (ADH), whereby the released carbonyl part of a reactive aldehyde can perform an electrophilic attack on the

primary amine of a DNA base forming a Schiff base which can react with another DNA base (forming an ICL) or a lysine or arginine residue of a protein¹²⁸. Similarly, histone or DNA demethylations, both common processes, produce Formaldehyde (FA) as a byproduct¹²⁹. FA is a very reactive molecule with genotoxic functions principally forming DPCs through a Schiff base stable amide bond between the amino-group of amino acids and DNA bases similar to other aldehydes^{102,130}. DPCs can also appear as a result of DNA damage or as a result of their repair. For instance, abasic sites occurring endogenously or during BER have reactive aldehyde groups, which can easily covalently bind nearby proteins²¹.

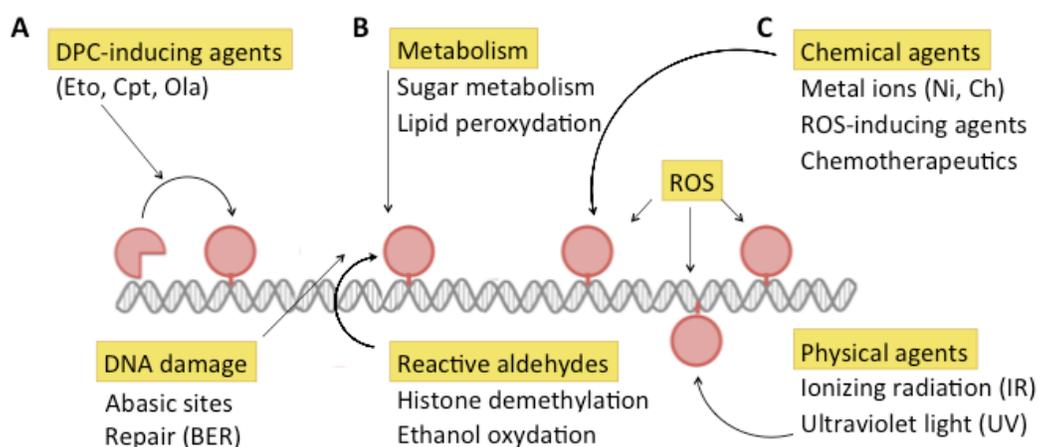


Figure 8 Sorts and induction of DNA-protein-crosslinks (DPCs). **(A)** Enzymatic DPCs appears spontaneously and in higher frequency upon DPC-toxins. **(B)** Endogenous DPC-inducing agents include various metabolic reactions, reactive aldehydes and reactive oxygen species (ROS). **(C)** Exogenous sources as chemical agents or radiation covalently bind proteins to DNA. DNA-damage and its repair can induce DPCs as well.

DPCs commonly result from exogenous toxic sources like radiation or toxins similar to other DNA lesions³⁸. Some examples of widely used chemotherapeutic drugs (Etoposide, Camptothecin) were already mentioned as DPC-inducing agents¹¹⁵. Further, platinum-based anti-cancer drugs (such as Cisplatin) form platinum-DNA complexes which crosslink to proteins (such as Histones)³⁷. 5-aza-dC is incorporated into DNA as a cytosine analogue and covalently traps DNA methyltransferase (DNMT) while it is methylated by this enzyme^{107,131}. Nitrogen-mustard derivatives such as Mustine (HN2) are also used in cancer therapy due to their alkylating function, which induces DPCs¹³². The toxicity of DPCs is already

being used to treat cancers by inducing double strand breaks on a genome-wide scale, DPCs can be caused by environmental factors resulting in similar hazardous effects.

Metal ions such as nickel or chromium can cause chemical crosslinking of proteins to DNA¹²⁷. A wide range of other substances including 1,2-dibromoethane, 1,3-butadiene can lead to DPCs simply as a result of their ROS production¹³³. Free radicals can attack DNA bases causing oxidative base lesions as well as other modifications, which serve as sites for DPC formation. Additionally, radicals attack the DNA backbone causing SSBs and DSBs^{30,134}. The ROS damage to lipids causes lipid peroxidation thereby forming aldehydes, which are known to generate DPCs as well^{135,136}. Ultraviolet (UV) light typically excites DNA bases causing them to react with one another thereby producing covalently linked pyrimidine dimers, or causing them to react with adjacent amino acids, thereby forming DPCs^{33,137,138}. Additionally, UV-light is known to cause DSBs and produces ROS with diverse consequences^{137,138}. Ionizing radiation (IR) damages DNA directly or by producing ROS and free radicals from surrounding water or oxygen molecules which can target DNA or proteins as previously mentioned^{139,140}. As a result, DNA base lesions, SSBs and particularly toxic DSBs as well as protein adducts occur^{34,35}. Interestingly, under normal conditions, IR induces mainly DSB but under hypoxic conditions IR produces preferentially DPCs, rather than DSBs^{38,141,142}. Since DSBs are more dangerous for cells, hypoxia is making cells more resistant towards IR, which causes a problem in clinical treatment of hypoxic tumor regions. A better understanding of DPCs and their repair could be helpful not only in understanding cell physiology but also in explaining the anti-neoplastic effect of cancer therapy, as well as in tumor resistance or cell death.

2.3.2 DPC-repair

The existing knowledge on DPC-repair mechanisms is in comparison to the repair of other DNA-lesion (e.g. DSB) obviously reduced despite the fact that DPCs represent a frequently appearing and highly toxic form of DNA damage¹⁴³.

As explained, there is a range of different DNA damage repair mechanisms, which in most cases act specifically on one type of lesion¹⁴⁴. Whether they also participate in DPC repair needs further investigation. However, some canonical DNA repair mechanisms are shown to provide resistance towards DPC-inducing agents and could therefore also play a role in DPC-removal. These include NER, HR and the Fanconi anemia pathway¹⁴⁵.

DNA damage repair mechanisms included in DPC-repair

The nucleotide excision repair (NER) mechanism was shown to resolve DPCs, with different partially controversial findings¹⁴⁶. Studies show that the ability of the NER system to remove DPCs is strongly dependent on the protein adduct size. An *in vitro* assay performed in bacterial cells showed that NER repairs DPCs of up to 16kDa, while *in vivo* this was possible for adducts under 12-14kDa^{144,147-149}. In mammalian cells, DPCs up to 8kDa were removed *in vitro*, while *in vivo* conflicting results were obtained which most likely depend on the cell model used and on the type of DPCs induced¹⁴⁹⁻¹⁵¹. It is important to mention that the resistance provided by NER towards DPC-inducing toxins (like aldehydes) can be explained by NER-dependent repair of DPCs or by NER-provided repair of other sorts of DNA damage caused by those toxins^{145,148-149}. Taken together, NER may be included in the repair of specific DPCs small enough (although the exact maximal size remains unclear) and thereby making a previous size reduction necessary.

Another canonical DNA damage repair pathway suggested to repair DPCs is homologous recombination (HR)^{145,149}. As an example, in *E. coli* HR is important for resolving DPCs too large to be resolved by NER¹⁴⁸. In mammalian cells, HR is included in repair of FA-induced DPCs regardless of their size¹⁴⁹. Interestingly in yeast, HR is protective against chronic low-dose FA treatment, but harmful under acute high-dose FA¹⁴⁵. Since NER was shown to protect cells against this high-dose FA (which causes primarily SSB), maybe an inadequate repair pathway choice by HR instead of NER could explain this finding¹⁴⁵. It is not excluded that the observed HR-mediated protection against DPC-inducing agents results from a

repair of DSBs or stalled replication forks, which are caused by DPCs, rather than from a direct repair of DPCs.

The Fanconi anemia pathway, primarily important for repairing interstrand crosslinks (ICL), was additionally given a role in DPC repair. This resulted from the increased sensitivity of cells deficient in Fanconi anemia factors towards DPC-inducing aldehydes, 5-aza-dC or PARP-inhibitors, although with opposite findings and no proposed underlying molecular mechanism¹⁵²⁻¹⁵⁷.

Mechanisms acting specifically on DPCs

Due to the mentioned versatility of possible DPCs most probably various distinct repair mechanisms are needed. Theoretically DPCs could be repaired by at least three non-redundant mechanisms acting either on the chemical bond of the protein to the DNA (cutting the crosslink), on the DNA component (cutting the DPC-containing part out of the DNA) or on the protein component (cleaving the protein from DNA) (Fig. 9).

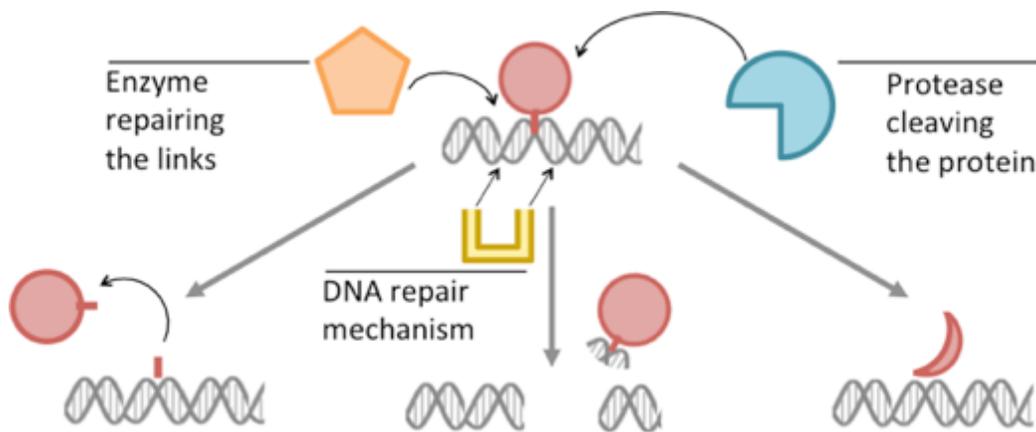


Figure 9 Scheme of DPC-repair. DPs can be resolved by 3 distinct approaches acting on the 3 main parts of a DPC (the linkage point, DNA or protein respectively). **(Left part)** Specific Repair mechanisms resolve the link between the protein and DNA as exemplified by TDP1 and TDP2. **(Middle part)** DPCs can be recognized as damaged DNA and repaired by canonical DNA damage repair mechanisms as NER or HR, which nick the DNA containing the crosslinked protein. **(Right part)** The protein crosslinked to DNA can be degraded by the proteasome or by proteolytic cleavage performed by a potential protease cleaving DPCs.

The first repair mechanism acting specifically on a DPC was discovered in 1996 and showed that Tyrosyl-DNA phosphodiesterase 1 (TDP1) is able to hydrolyze

the phosphotyrosyl bond between the 3'-end phosphate and the Tyrosine-residue in Topo I-DPCs^{158,159}. In order to perform this hydrolysis, Topo I first needs to be partially degraded to a smaller peptide to make it accessible for TDP1¹⁶⁰. After degradation and hydrolysis of the DPC, the remaining SSB is further processed by polynucleotide kinase 3'-phosphatase (PNKP) and repaired by the SSB-repair machinery including PARP1, XRCC1 and LIG3¹⁶¹. A similar mechanism of DPC crosslink hydrolysis is known for Topo II-DPCs and performed by Tyrosyl-DNA phosphodiesterase 2 (TDP2)^{162,163}. The difference to the Topo I-DPC repair is that the hydrolysis is performed on a 5'-end phosphate and that it leaves a DSB, which is then repaired by canonical DSB repair mechanisms (HR or NHEJ)^{162,164}.

Since the crosslinks between DNA and proteins can be chemically very different, it is not very probable that there is a specific repair-mechanism for each sort of crosslink. A more efficacious way to deal with different DPCs would be a nucleolytic cleavage of the DNA part of the DPC, which could be performed by canonical DNA repair mechanisms. In order with this the MRN-complex consisting of meiotic recombination 11 (MRE11), RAD50 and Nijmegen breakage syndrome protein 1 (NBS1), which is usually active in DSB repair, has been shown to remove DPCs containing different proteins (Topo I, Topo II, Spo11) as well^{75,76,162,165,166}. This is done by cleaving both strands of the DNA near the protein adduct in order to remove the DPC-containing DNA part from the intact DNA and thereby forming a DSB, which can be further repaired by HR or NHEJ.

Potential function of proteolysis in DPC-repair

A specific problem regarding DPC-repair is the potentially very large and bulky nature of DPCs since this sterically hinders the repair molecules from recognizing and reaching the inaccessible point of linkage or DNA. This implies that a repair mechanism targeting the protein component would be the most efficient. The protein adduct could be released from DNA completely or at least it has to be reduced first and subsequently enabling access to the linkage or the affected DNA by the above-mentioned or similar repair mechanisms.

The possibility that DPC-repair includes degradation of the protein adduct was shown by several studies. The degradation is thereby mediated by the proteasome¹⁶⁷. In the case of Topoisomerase-DPCs, the Ubiquitin(Ub)/26S proteasome was shown to degrade the protein component to a small peptide remnant thereby revealing the crosslink and making it accessible for hydrolysis by TDP1 or TDP2 as previously explained^{160,168-171}. Additionally, It was shown that proteasomal degradation is important for repair of DNA-methyltransferase-DPCs^{151,172}. Accordingly, the inhibition of proteasome impairs the repair of FA-induced DPCs as well as chromium (VI)-induced DPCs¹⁷³. Although the mechanism on molecular level has not been completely identified, there is enough convincing evidence to support a model of proteasome-mediated degradation of DPCs¹⁴⁵. This model, where the proteasome reduces the bulky protein moiety of DPCs in order to either allow further repair of the crosslink or to enable lesion bypass has been widely accepted but there is evidence that this might not be the only proteolysis-mediated mechanism in DPC-cleavage. A proteolysis-based mechanism included in DPC-repair independent of the proteasome was shown to exist supporting the hypothesis that cells possess proteases that can cleave the protein component of DPCs as well.

2.3.3 Proposed SPRTN function in DPC-repair

The idea that SPRTN could act as a protease that cleaves DPCs emanates from two major findings.

First, a study in cell-free *Xenopus laevis* egg extracts focusing on DPC repair showed that *in vitro* replication is stalled by DPCs but resumes after proteolytic degradation of the protein adduct, whereby the proteolytic cleavage was directly measured for the first time in metazoan and shown to be independent of proteasome¹⁷⁴. However, the protease responsible for the proteolytic cleavage remained unidentified.

Second, a study in budding yeast identified the DNA-dependent metalloprotease Wss1 as the first protease with direct proteolytic activity on DPCs. This study was of special interest to us since Wss1 and SPRTN show similarities with convincing

evidence that Wss1 (Weak suppressor of *smt3*) could be the yeast orthologue of SPRTN¹⁷⁵.

Taking into account that SPRTN is involved in DNA damage repair and that it has a protease-domain and being aware of the fact that its proposed yeast orthologue is a DPC-protease, while a mechanism of DPC-proteolysis has been shown in metazoan, we propose that SPRTN's unknown essential function could be a proteolytic cleavage of DPCs. This would mean that SPRTN presents the first shown DPC-cleaving protease in mammalian cells.

3 Aim

The aim of this study is to identify the essential function of SPRTN and to clarify SPRTN's role in DNA damage response. We hypothesize that SPRTN's main function is contained in its protease domain and that its proposed proteolytic activity is necessary for the maintenance of genomic integrity. We thereby propose a potential involvement of SPRTN in DPC-repair as the underlying mechanism of SPRTN's essential activity.

The initial objective is to determine if SPRTN is a protease and to identify its potential substrates. A further aim is to clarify if SPRTN acts in DPC-repair. Accordingly, we aim to demonstrate an interaction of SPRTN with DPCs and validate a SPRTN-mediated DPC-cleavage. In case of a proven DPC-removal by SPRTN, a further aim is to determine the relevance of SPRTN in cellular response to DPCs. A second objective is to explain the effects of the clinical mutations of SPRTN and subsequently give a molecular explanation of the underlying mechanism of RJALS-syndrome.

4. Results

4.1 SPRTN has a proteolytic function

In order to investigate the function of SPRTN and the effect of SPRTN mutations found in Ruijs-Aalfs syndrome patients, we first wanted to investigate whether SPRTN has proteolytic activity, how it acts as a protease and how this function is compromised by mutations. Therefore, we focused on comparing the function of ^{WT}SPRTN and its mutated variants, especially the catalytic inactive form. We used a SPRTN-variant completely lacking the potential protease activity, ^{E112A}SPRTN (first established by Kim et al.) as the glutamic acid in the HExxH-core of the protease domain is necessary for substrate cleavage (Fig. 10)^{19,176,177}. Since a knockout mouse model of SPRTN was shown to be embryonic lethal, we used an overexpressing system where SPRTN is over-expressed in cell lines without previously silencing endogenous SPRTN in order to avoid additional side effects²⁰.

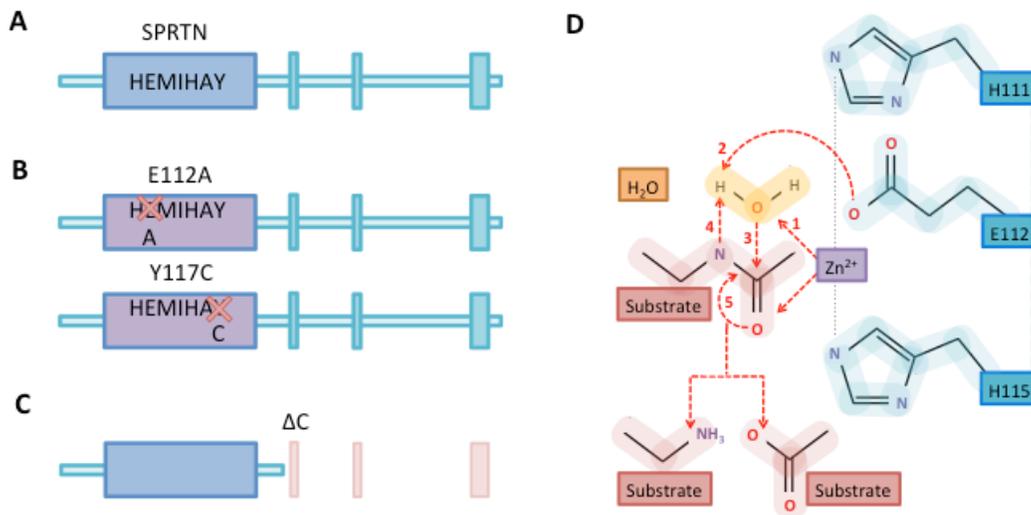


Figure 10 SPRTN as a 489-amino-acid-long Zn²⁺-binding metalloprotease and its mutations. **(A)** WT-SPRTN contains multiple domains, the biggest one being the N-terminal SprT-domain with a HExxH motif typical for zinc-dependent metalloproteases. **(B)** Catalytic E112A-mutant has a mutated glutamic acid in the HExxH motif essential for polarizing H₂O and enabling a nucleophilic attack on the substrate. Clinical Y117C-mutant has a mutation in close proximity to HExxH, which presumably alters the structure of the catalytic core with impaired proteolysis similar to E112A. **(C)** Two clinical ΔC-mutants are truncated versions of 246 or 249 amino acids respectively with a preserved protease SprT-domain. **(D)** Scheme of proteolysis by the HExxH motif demonstrated by subsequent reaction steps: 1 - Zn²⁺ (coordinated by H111 and H115) is activating C=O of the substrate protein and H₂O simultaneously. 2 - Negatively charged E112 deprotonates H₂O by a nucleophilic attack binding a H⁺. 3 - Polarized OH⁻ from

H₂O attacks C=O of the substrate via a nucleophilic addition. 4 – N from the substrate deprotonates E112A by binding of H⁺ previously taken from H₂O. 5 – O⁻oxyanion from substrate initiates elimination and cleavage of the peptide bond of the substrate.

4.1.1 SPRTN is a protease with an autocleavage activity

An investigation of SPRTN's proposed proteolytic function without a known substrate required a different approach. We first looked at the behavior of SPRTN in cells without any interference. Regarding that autolysis or self-cleavage is a commonly seen phenomenon among proteases serving either to promote (as in precursor molecules) or to decrease their activity, we aimed to test whether self-cleavage occurs in the case of SPRTN. With this assumption we compared cells overexpressing the active form ^{WT}SPRTN and cells overexpressing the catalytic inactive ^{E112A}SPRTN (with inactive proteolytic activity) and observed by western blot faster migrating SPRTN fragments in the ^{WT}SPRTN but not in the ^{E112A}SPRTN, presenting cleaved SPRTN fragments (Fig. 11). This finding supports the idea that SPRTN has the ability to cleave proteins and that the activity of its protease domain is needed for its self-cleavage, since mutating the protease domain impairs self-cleavage. The small amount of degraded forms of the inactive ^{E112A}SPRTN is most likely a consequence of cleavage by the endogenous SPRTN.

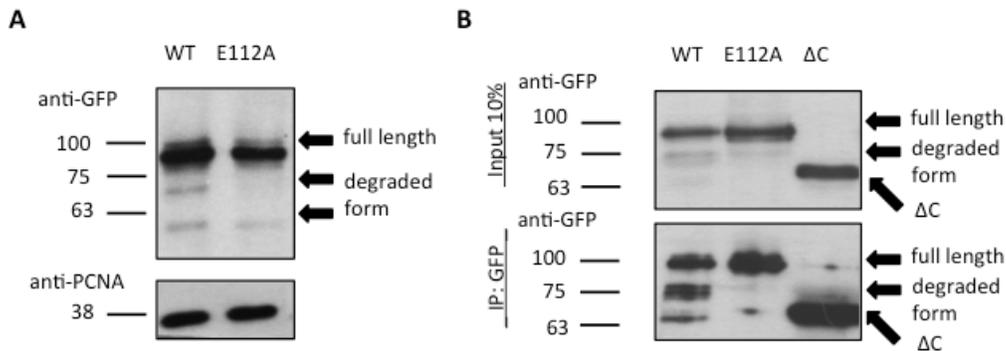


Figure 11 SPRTN has an autocleavage activity impaired upon mutation. **(A)** HEK293T cells were transiently transfected with GFP-tagged WT-SPRTN or the catalytic mutant E112A-SPRTN and left to express the proteins for 72h. Cell-lysates were analyzed by SDS-PAGE followed by immunoblotting with GFP-specific antibodies and compared to PCNA as a loading control. The WT-cell lysate shows cleaved GFP-tagged fragments, which are strongly reduced in E112A. **(B)** Immunoprecipitation with GFP-binding beads from cells expressing GFP-tagged WT, the catalytic inactive E112A or the truncated ΔC-variant validates an impaired self-cleavage in E112A. SPRTN is cleaved from the C-terminus showing different N-terminally labeled truncated fragments. One of the main cleavage products shows a similar molecular weight as the ΔC-variant.

We further wanted additional support of the self-cleavage activity of SPRTN on cellular level by visualizing with additional methods. We took advantage of the fact that the Δ^C SPRTN-mutant (which actually represents the N-terminal part of SPRTN, missing the C-terminal half) is localized in the cytoplasm while the full-length protein remains inside the nucleus, as it was shown in previous studies⁹. As we observed that one of the most abundant cleaved fragments shows a size approximately half the size of the full-length protein (Fig. 11B), we assumed that the cleavage is producing a cleaved SPRTN-form approximately similar to Δ^C SPRTN. Since additional smaller cleavage fragments were seen, there are other cleavage sites in the SPRTN molecule as well, but not of interest for this experiment. Taking this into account, we developed a SPRTN-molecule labeled with GFP at the N-terminus and mCherry at the C-terminus. We expected to see the full-length protein as a yellow signal in the nucleus (as a result of simultaneous green and red fluorescence) as well as a green signal of the GFP-tagged N-terminal cleavage fragment in the cytosol (Fig. 12A). Indeed, as supposed we observed cytosolic N-terminal cleavage product as a proof of SPRTN's auto-cleavage activity *in vivo* (Fig. 12B).

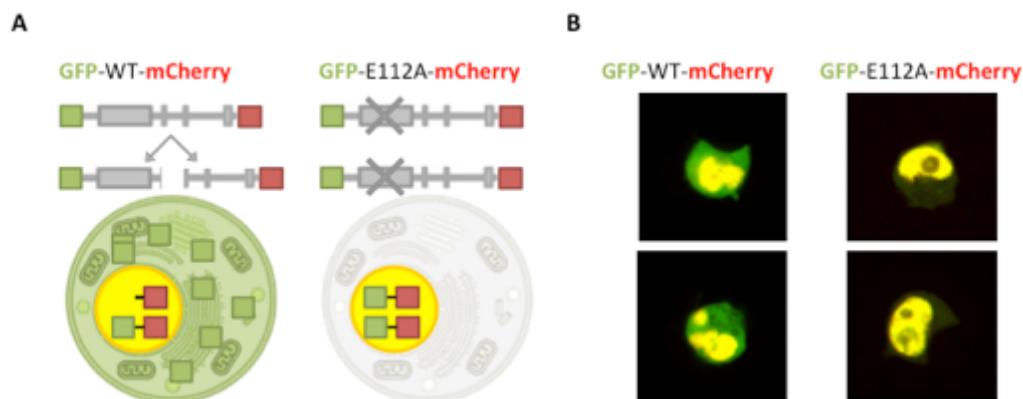


Figure 12 SPRTN's *in trans* self-cleavage shown *in vivo*. **(A)** Scheme of double-labeled SPRTN-molecules and cleavage products and their intracellular localization. The full-length protein is localized in the nucleus while the truncated SPRTN missing the C-terminus shows a cytoplasmic localization. **(B)** HEK293T cells were transiently transfected with a vector carrying either a double-tagged WT-SPRTN or E112A-SPRTN. Microscopic view of cells expressing double-tagged SPRTN-variants shown as a yellow intranuclear signal, resulting from the collective green (GFP) and red (mCherry) fluorescence. Cells expressing WT-SPRTN additionally exhibit a green signal in their cytoplasm from cleaved SPRTN-fragments demonstrating SPRTN-autolysis, which is missing in the cells expressing the catalytic mutant E112A-SPRTN.

We conclude that SPRTN has a proteolytic function exemplified by the self-cleavage of this protein, which is dependent on its conserved protease domain. Importantly, mutations of this part lead to a loss of this function. In addition, we were further able to show that SPRTN-autoproteolysis happens *in trans* since a small but still present cleavage happens on the catalytic inactive ^{E112A}SPRTN as well (Fig. 11A). Since ^{E112A}SPRTN has no proteolytic activity, it has to be cleaved by another protease, and due to the fact that the cleavage products are the same as in self-cleavage of ^{WT}SPRTN (seen in western blot assay) the most likely explanation for this observation is that the catalytic mutant is cleaved *in trans* by other SPRTN molecules, which are endogenously present in cells. This was additionally supported by a slight green signal from the cleaved ^{E112A}SPRTN in the cytosol under microscopic observation.

4.2 SPRTN interacts with DNA-bound proteins

After providing evidence that SPRTN has a proteolytic function and considering that SPRTN plays a role in the DNA damage response we reasoned that SPRTN acts in proteolysis of proteins related to or bound to DNA^{12-14,16-18}. Since there was no known substrate we performed an unbiased screen for potential substrates. For this we used liquid chromatography-mass spectrometry after SILAC-labeling (stable isotope labeling with amino acids in cell culture) to identify proteins with different abundance between ^{WT}SPRTN and ^{ΔC}SPRTN cells (with thanks to J. Lopez-Mosqueda who performed the experiment using the ^{ΔC}SPRTN instead of ^{E112A}SPRTN, as cells expressing the catalytic mutant died prematurely) (Fig. 13A). The search for newly formed amino-terminal amino acids (formed as a consequence of protein cleavage, thereby signaling cleaved substrates) revealed a myriad of proteins as potential ^{WT}SPRTN-substrates.

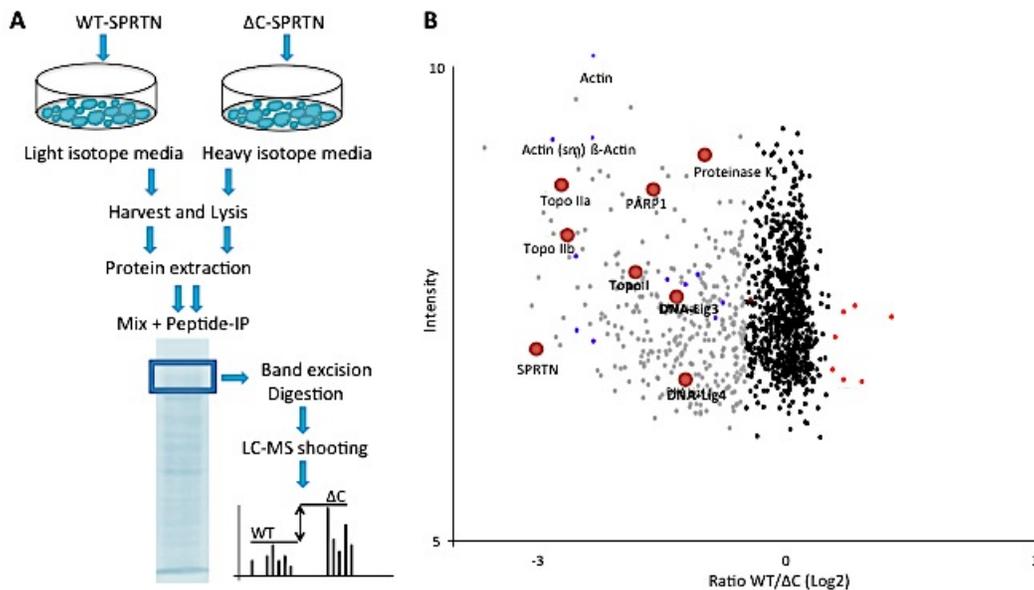


Figure 13 Substrates of SPRTN detected by mass spectrometry. **(A)** Scheme of the SILAC mass spectrometry using cells stably expressing WT-SPRTN or ^{ΔC}SPRTN grown in either light- or heavy-isotope media to label substrate proteins. Cells were lysed after 24 h. After combining both differently labeled lysates, the proteins were separated by SDS-PAGE, purified and digested by trypsin. The prepared samples were separated by liquid chromatography and analyzed by their mass to charge ratio via mass spectrometry. **(B)** SPRTN-interactome shown as a diagram presenting different substrates recognized by newly formed amino-acids and distinguished by their abundance ratio between ^{ΔC}SPRTN and WT-SPRTN. DNA-bound proteins like Histones, PARP1 and Topoisomerases together with SPRTN showed prominently higher prevalence upon SPRTN-mutation and therefore present potential substrates (with thanks to J. Lopez-Mosqueda).

Interestingly, among those proteins there were several prominent ones known to be related to DNA, either functionally or structurally (Fig. 13B). Those proteins were of special interest to us, in view of SPRTN's relation to DNA. We decided to further focus on enzymes known to be transiently bound to DNA but which could be permanently trapped forming DPCs under certain conditions (like PARP1, or Topoisomerase I or II) considering that SPRTN's yeast orthologue Wss1 functions as a DPC-cleaving protease^{118,120,121,175}.

4.2.1 SPRTN is recruited to DNA by DPCs

In consideration of our hypothesis we assumed that the mentioned enzymes would serve as SPRTN's substrates only if they are trapped to DNA and thereby formed DPCs, rather than in their free form. Otherwise, if these proteins would generally interact with SPRTN, SPRTN would function as an inhibitor of the mentioned enzymes with important functions in DNA damage repair thereby having a negative effect on DNA-repair, opposite to the current findings. In order to provide support for our hypothesis, a method that shows primary interaction of SPRTN with the DNA-bound protein fraction, instead of general SPRTN-protein-interaction, was needed. Therefore co-immunoprecipitation, usually used to show protein-protein interactions, was not suitable for this purpose. Instead, we took advantage of the fact that DPCs are by definition bound to DNA, and if they are supposed substrates of SPRTN, their presence should track SPRTN to DNA in order to enable cleavage. Accordingly we isolated chromatin and analyzed the level of SPRTN on DNA in dependence of DPCs. The induction of DPCs should serve as a signal for recruiting SPRTN to its newly formed substrate and thereby to DNA. With this presumption we tested whether DPC-induction has an effect on the amount of SPRTN bound to DNA. As assumed, we could detect a clear increase of SPRTN-level on DNA upon DPC-induction indicating that DPCs recruit SPRTN to DNA (Fig. 14).

We conclude that the chromatin-recruitment of SPRTN is dependent on the induction of DPCs. This experiment was performed using various known DPC-inducing agents (Etoposide, Camptothecin, Olaparib), each one targeting a

specific enzyme, which was found in the mass spectrometry screen as a potential substrate of SPRTN. Given the fact that all used DPC-inducing drugs had the same effect, regardless of the enzyme identity, this finding indicates a broad role of SPRTN function in DPC-repair. Since the employed DPC-inducing agents act selectively trapping one protein to DNA respectively, this increase in SPRTN-level on DNA upon DPC-induction treatment is explained either through a direct recruitment of SPRTN by its potential substrate or alternatively, through an indirect recruitment by SPRTN interacting proteins, namely PCNA, ubiquitin or VCP/p97. In either case, SPRTN function is implicated in DPC-repair, but the exact mechanism has to be further analyzed.

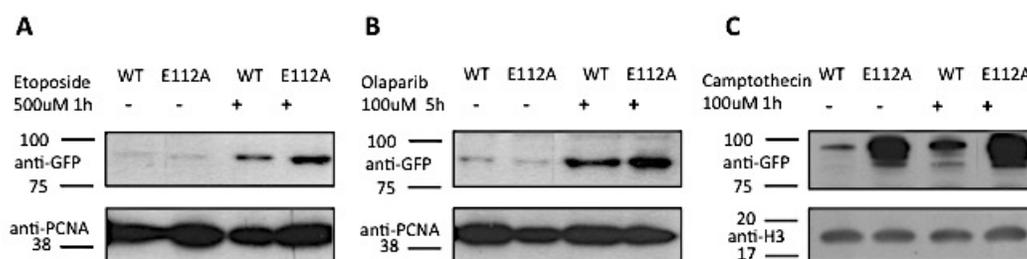


Figure 14 SPRTN's chromatin bound level is increased upon DPC-induction. HEK293T-cells transiently expressing either GFP-WT-SPRTN or inactive E112A-mutant were treated with different DPC-inducing agents and compared to untreated cells. Cell lysates were fractionated to isolate chromatin-bound proteins and afterwards analyzed by immunoblotting with GFP-antibodies compared to a loading control. **(A)** Olaparib treatment with 100µM for 5h inducing PARP1-DPCs. **(B)** Etoposide treatment with 500µM for 1h inducing Topoisomerase II-DPCs. **(C)** Camptothecin treatment with 100µM for 1h inducing Topoisomerase I-DPCs.

4.2.2 Recruitment of SPRTN to DNA is affected by its proteolytic activity

Analyzing SPRTN recruitment upon DPC-induction, we observed that the catalytic inactive ^{E112A}SPRTN showed an increased level on DNA upon DPC-induction. An increase in protein levels on chromatin can generally be caused either by an increase in SPRTN transcription, by a stronger binding affinity or by a sustained binding. The only difference between ^{E112A}SPRTN and ^{WT}SPRTN is a point mutation in the protease domain thereby only affecting its proteolytic function without major effects on its structure¹⁹. This fact makes a stronger DNA-binding affinity unlikely and indicates that the increased amount of ^{E112A}SPRTN on DNA is rather a result of its inability to remove DPCs, subsequently remaining on DNA

without performing its function. This is further supported by the fact that the difference between E^{112A} SPRTN and WT SPRTN abundance is more pronounced upon DPC-induction. Yet, to exclude that the higher E^{112A} SPRTN level is caused by a higher expression level of E^{112A} SPRTN in total, a simultaneous analysis of the chromatin fraction (CF) and total cell lysate (TCL) was performed in physiological and DPC-inducing conditions. We could see that the total level of E^{112A} SPRTN is only slightly higher compared to WT SPRTN but clearly more pronounced in the chromatin fraction, especially under DPC-induction, confirming that the inactive E^{112A} SPRTN is preserved on DNA more than WT SPRTN (Fig. 15).

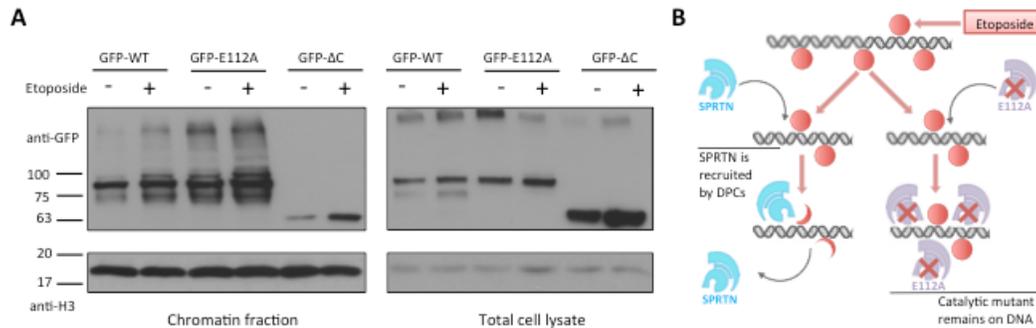


Figure 15 SPRTN remains chromatin-bound upon inhibition of its proteolytic function. **(A)** Comparison of chromatin bound SPRTN-level with total cell level in HEK293T-cells transiently expressing GFP-tagged SPRTN-WT, E112A or ΔC . An increased chromatin bound SPRTN-level is observed upon Etoposide treatment (250 μ M for 24h) with a low effect on the total cell level. The ΔC -variant shows a cytosolic localization with a raised chromatin-level after Etoposide-treatment. **(B)** Scheme of SPRTN recruitment to chromatin upon induction of Topo II-DPC by Etoposide with an increased level of the catalytic mutant supposedly due to its impaired cleavage of the Topo II-DPCs.

We used this system to measure the extent of SPRTN recruitment to DNA upon DPC-induction and how the total level of SPRTN is affected by DPC-induction. Further we wanted to simultaneously compare both known mutant versions (ΔC SPRTN and E^{112A} SPRTN) to WT SPRTN. Thereby we could conclude that DPC-induction is recruiting all SPRTN variants to DNA, while the mutated forms remain on DNA to a higher extend. This is especially obvious in the case of ΔC SPRTN, since this mutant is primarily located in the cytoplasm but under DPC-induction becomes recruited to DNA. This observation gives further and stronger evidence for the connection of SPRTN and DPCs.

4.3 SPRTN cleaves DNA-protein-crosslinks (DPCs)

Subsequently to the evidence that SPRTN is a protease and that it is recruited to DNA by DNA-protein-crosslinks (DPCs) in a manner dependent on its proteolytic activity, the next obvious objective was to demonstrate proteolysis of DPC-substrates.

4.3.1 DPC-removal is dependent on SPRTN

In order to demonstrate DPC-cleavage, it is necessary to visualize complexes of proteins bound to DNA separately from the free proteins and additionally to detect changes in protein-level, not in total, but only for the fraction of proteins bound to DNA (the DPC-fraction), while simultaneously proving that the free protein portion is left unchanged. While searching for a way to distinguish cleavage of DPCs from cleavage of free proteins, we found a cesium chloride (CsCl) density gradient assay enabling us to isolate DNA-bound proteins from free ones (regardless if they are in the cytosol or nucleus)¹⁷⁸. This was done by an ultracentrifugation of cell lysates through a CsCl density gradient and further analysis of the amount of the protein of interest, in our case Topoisomerase II (Topo II), DNA-bound and unbound separately (Fig. 16A)¹⁷⁹. The finding that upon Etoposide treatment most of Topo II is DNA-bound (moving from the free fraction to the DNA-bound fraction) correlates with the known DPC-inducing function of Etoposide and simultaneously confirms the accuracy of this experimental method (Fig. 16B).

Since the DNA-bound fraction representing Topo II-DPCs is strongly reduced upon overexpression of ^{WT}SPRTN, this experiment gives the first clear evidence that DPC-removal is depending on SPRTN. This removal was shown to be a result of the proteolytic activity of SPRTN validated by the obvious difference in the amount of Topo II left DNA-bound (representing DPCs) between the proteolytic inactive ^{E112A}SPRTN (with a high level of Topo II-DPCs) and the active ^{WT}SPRTN (with almost no Topo II-DPCs detectible) (Fig. 16B). Importantly, this difference was only shown in the DNA-bound fraction of Topo II, not in free Topo II-fraction, which basically remained unaltered, supporting that SPRTN cleaves Topo II only in its DPC-form, as presumed. With these results, we obtained clear evidence that

SPRTN cleaves the DNA-bound fraction of its substrate protein (DPCs) and that this function depends on its protease domain.

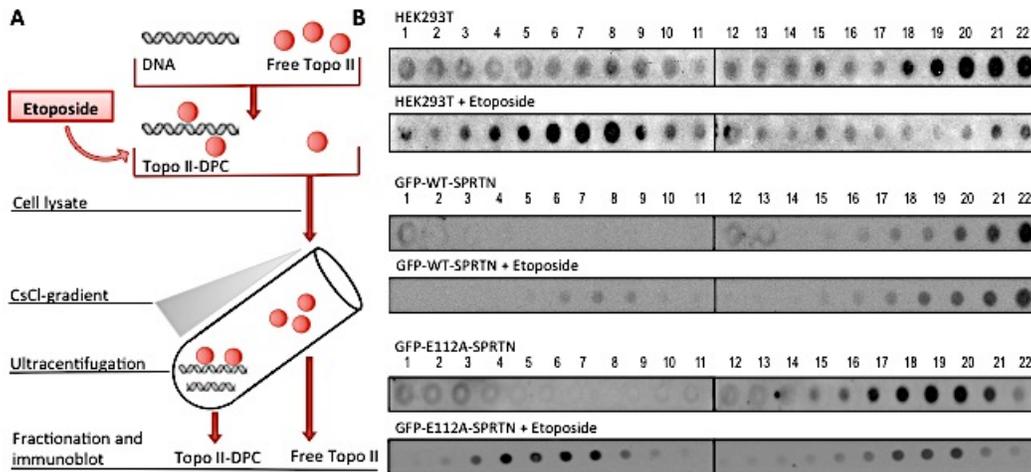


Figure 16 SPRTN cleaves proteins covalently bound to DNA. **(A)** Illustration of the CsCl-gradient assay. Cells were treated with Etoposide (250 μ M 1h) to induce Topo II-DPCs and compared to untreated cells. Cell lysis under denaturing conditions using Laemmli buffer was followed by ultracentrifugation in a CsCl-gradient column in order to separate free proteins left in the upper part of the column from the DNA-fraction in the bottom part of the column containing covalently bound proteins (DPCs). **(B)** Comparison of Topo II-level upon Etoposide treatment with an untreated control, in endogenous HEK293T-cells, cells transiently transfected with GFP-tagged WT-SPRTN or with the catalytic mutant E112A respectively. Distinct fractions from the CsCl-gradient ultracentrifugation were separately transferred by Dot blot (lower numbers presenting lower gradient fractions containing DPCs and higher numbers showing fractions with free proteins) and analyzed by immunoblotting with Topo II-specific antibodies.

4.3.2 SPRTN's proteolytic activity is relevant for DPC-removal

After demonstrating that DPC-cleavage depends on SPRTN, we aimed to further validate this finding by additional analysis of the protein-cleavage. First, we wanted to simultaneously measure the amount of SPRTN and its DPC-substrate in order to rule out the unexpected possibility that the observed difference in DPC-cleavage is a result of a lower expression level of ^{E112A}SPRTN compared to ^{WT}SPRTN. Furthermore, since the CsCl density gradient gives only information about the amount of Topo II in total without dichotomizing between full-length or fragments (as it shows all parts reacting with the used anti-Topo II-antibody in the probed fraction) we further wanted to confirm that the demonstrated cleavage happens on the full-length protein. For that reason we utilized SDS-PAGE to segregate the Topo II from its potential fragmented forms and to determine

SPRTN-level at the same time. Since only DPC-cleavage was of interest, we previously isolated the chromatin-bound fraction of the cell lysates. Consistent with our finding from the CsCl-assay, a clearly higher level of unresolved full length DNA-bound-Topo II upon Etoposide treatment in ^{E112A}SPRTN compared to ^{WT}SPRTN could be determined. At the same time, the level of ^{E112A}SPRTN was raised in comparison to ^{WT}SPRTN, thereby confirming that the observed higher cleavage of Topo II-DPCs in ^{WT}SPRTN is directly reliant on SPRTN's proteolytic function, not its level (Fig. 17A). We repeated the assay with ^{ΔC}SPRTN as well and observed a similar, yet less pronounced deficit in the cleavage of Topo II-DPCs compared to ^{WT}SPRTN. Notably, there was an obviously lower expression of ^{ΔC}SPRTN on DNA compared to ^{WT}SPRTN (consistent with our previous results and the known cytosolic localization of this mutant) (Fig. 17B). Therefore the observed higher level of Topo II-DPCs in cells expressing ^{ΔC}SPRTN could be explained by the mislocalization of this mutant, rather than a weakness in its proteolytic activity. Nevertheless, both mutations lead to a greater accumulation of DPCs as a final result, compatible with the deficits observed clinically and experimentally for both variants^{9,10}.

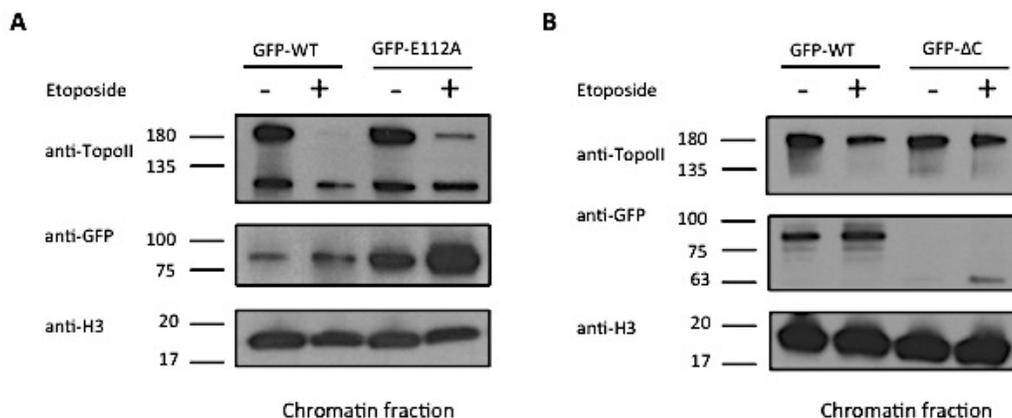


Figure 17 DPC-cleavage by SPRTN depends on its proteolytic function and localization. HEK293T-cells transiently expressing different GFP-SPRTN variants were treated with Etoposide (250μM 24h) to induce Topo II-DPCs and cell lysates fractionated to isolate chromatin-bound proteins followed by immunoblotting with different antibodies to detect the amount of Topo II-DPCs and DNA-bound SPRTN-level in parallel. **(A)** Comparison of Topo II-DPC cleavage and SPRTN-level in cells expressing either GFP-tagged WT-SPRTN or the catalytic inactive E112A. Higher Topo II-DPC-level after Etoposide treatment is seen in the catalytic mutant variant. **(B)** Analogous comparison of cells expressing GFP-ΔC-variant or GFP-WT. An impairment of Topo II-DPC cleavage is seen in the ΔC-variant (similar but less pronounced as in E112A, despite weaker chromatin recruitment).

Surprisingly, the level of full length Topo II was found to be decreased upon Etoposide treatment in western blot assays regardless of the SPRTN-variant expressed. This could be observed in all experimental settings, regardless of exposure time or concentration of Etoposide. As Etoposide is known to bind Topo II to DNA, Topo II-level is expected to be higher, especially on DNA¹²⁰. The most reasonable explanation for this adverse finding is that under Etoposide Topo II is forming DPCs, which are toxic and quickly become cleaved thereby reducing Topo II-level. To investigate the effect of Etoposide on chromatin-bound Topo II, we exposed cells to Etoposide for increasing time points and observed a decline in Topo II-level upon time, supporting our hypothesis (Fig. 18). As the Topo II-reduction upon Etoposide is more pronounced in the chromatin fraction (Fig. 19), this explanation seems even more feasible. Additionally, the same effect of protein-level reduction upon DPC-formation was observed for other DPC-inducing agents (data not shown), further supporting this theory. Regardless of the underlying mechanism, since this observation was consistent in all experimental settings regardless of the used SPRTN variants, it had no relevance for our study of SPRTN and needed no further investigation.

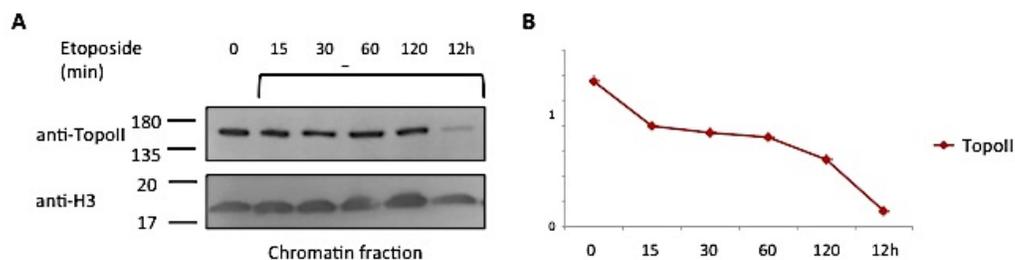


Figure 18 Time dependent decline of the DNA-bound fraction of Topoisomerase II upon Etoposide treatment. **(A)** Immunoblotting of chromatin bound Topo II with loading control from HEK293T-cells treated with Etoposide 250 μ M for the indicated time. **(B)** Diagram showing Topo II-level normalized to the loading control showing a time dependent decrease of chromatin bound Topo II after Etoposide treatment.

After explaining the reduced Topo II-level on DNA upon Etoposide treatment as a result of Topo II-DPC repair, we wanted to compare Topo II-DPCs and the total level of Topo II in order to estimate SPRTN's proteolytic function on DNA-bound and total Topo II in parallel. We made use of the previously performed

$^{WT}SPRNT/E^{112A}SPRNT$ comparing assay but with simultaneous analysis of the chromatin-bound fraction (CF) and the total cell lysate (TCL) (Fig. 19A). Indeed by comparing the Etoposide-induced DNA-bound and total Topo II-level (presented as ratios of Topo II after and before Etoposide treatment) between $^{WT}SPRNT$ and $E^{112A}SPRNT$ we found that an impaired proteolytic function of SPRNT leads predominantly to a higher abundance of Topo II bound to chromatin (DPC), rather than Topo II in total thereby confirming SPRNT as specific DPC-cleaving protease (Fig. 19B).

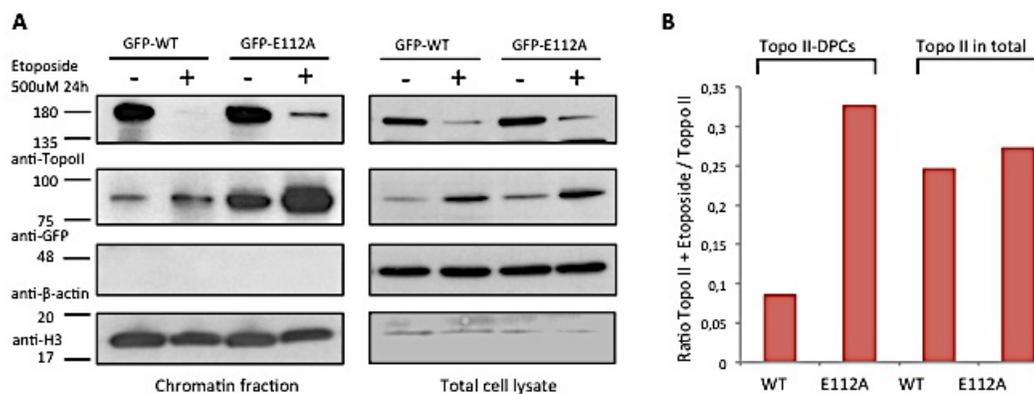


Figure 19 SPRNT cleaves the DNA-bound fraction of Topoisomerase II. **(A)** Simultaneous immunoblotting of chromatin bound proteins and total cell proteins. HEK293T-cells expressing GFP-WT or E112A with/without Etoposide treatment (250μM 24h) were harvested and divided to equal amounts for direct total cell lysis and for fractionation in chromatin bound and free proteins. Topo II-level upon Etoposide treatment is stronger decreased by WT than by inactive E112A. The difference in Topo II-level between WT and E112A (as a marker of SPRNT-activity) is pronounced in the chromatin bound fraction demonstrating a proteolysis of primarily Topo II-DPCs. **(B)** Diagram showing Topo II-ratios (after and before Etoposide treatment) for chromatin bound Topo II-DPCs and total cell Topo II in cells expressing the proteolytic active WT or inactive E112A. The comparison of chromatin fractions and total cell lysates reveals a primary effect of WT-SPRNT on DNA-bound Topo II-DPCs, less on Topo II in total. This proteolytic effect is missing in the catalytic E112A-mutant.

4.3.3 Loss of function of SPRNT impairs DPC-removal

After confirming that SPRNT functions as a DPC-cleaving protease, we wanted to further validate that this DPC-cleavage is not only a result of SPRNT overexpression, but happens as a result of endogenous SPRNT activity as well. In order to demonstrate this, we needed to compare a model with unaltered SPRNT-function to a model where SPRNT-function is lost, not only mutated. Since a

complete SPRTN-knockout (KO) is lethal we used a conditional knockout system established by Maskey et al. 2014²⁰. In this model mouse embryonic fibroblasts (MEFs) carrying a floxed SPRTN-allele (SPRTN^{fllox/-}) were transduced with Cre-ER^{T2}-retroviruses introducing a Cre-recombinase, which can convert the floxed SPRTN-allele into a KO-allele (thereby completely terminating SPRTN-expression) but only upon 4-hydroxytamoxifen (4-OHT) treatment (SPRTN^{fllox/-};Cre-ER^{T2}-MEFs) (Fig. 20A). We used mouse embryonic fibroblasts (MEFs) and treated them with 4-hydroxytamoxifen (4-OHT) to induce SPRTN-KO. According to our finding that SPRTN cleaves DPCs we wanted to investigate the effect of SPRTN-KO on DPC-levels and treated these MEFs with Olaparib in order to induce PARP1-DPCs¹²¹. Using a chromatin fractionation assay we determined the prevalence of DPCs by measuring PARP1-level in the DNA-bound fraction and compared this between the 4-OHT-treated MEFs (SPRTN-KO) and the untreated MEFs (SPRTN-WT). As suspected, we could observe that a knockout of SPRTN leads to an accumulation of DPCs (shown by a raised PARP1-level on DNA) (Fig. 20B).

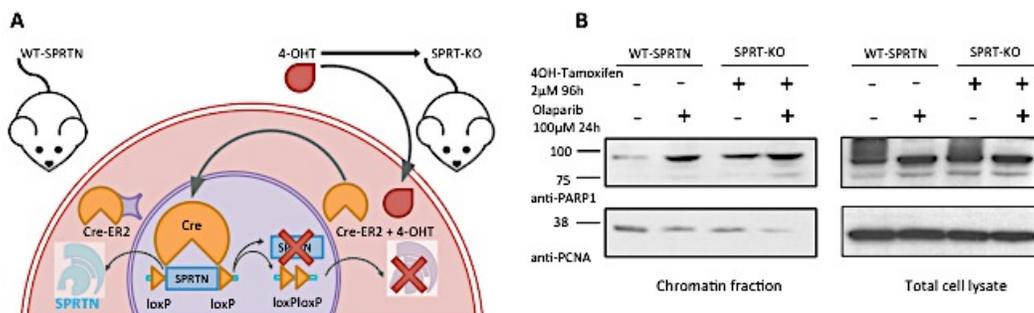


Figure 20 SPRTN-deficiency causes impaired DPC-cleavage similar to SPRTN-mutation **(A)** Scheme of inducible KO of SPRTN. MEFs were treated with 2µM 4-OHT for 96h to activate a Cre-recombinase by allowing it to enter the nucleus. After entering the nucleus the Cre-recombinase subsequently excises the floxed SPRTN allele thereby transforming the cells to SPRTN-KO MEFs (right part of the picture). Untreated MEFs contained the Cre-recombinase in the cytosol and thereby inactive resulting in an unaltered SPRTN-gene and a normal SPRTN-expression (left part). **(B)** Immunoblotting of PARP1 upon Olaparib treatment. Control SPRTN and SPRTN-KO MEFs were treated with Olaparib 100uM for 24h or left untreated. Cells were simultaneously subjected to chromatin fractionation and direct lysis respectively followed by an analysis by immunoblotting.

Since this result resembles the observed increase of DPC-level caused by a mutation in SPRTN's protease domain (^{E112A}SPRTN), we concluded that inhibition

of SPRTN's proteolytic activity has the same effect on DPC-cleavage as a complete loss-of-function of SPRTN (Fig. 19). Together with our previous findings this supports that the relevant function of SPRTN pertaining to DPCs is contained within its proteolytic core and its malfunction leaves DPCs unrepaired, thereby causing genetic instability.

4.3.4 SPRTN is a promiscuous protease

Considering the fact that different DPCs can be cleaved by SPRTN (e.g. Topo II, PARP1) we reasoned that SPRTN could have an unspecific proteolytic function cleaving various distinct DPCs. Yet, the DPCs we studied were enzymatic DPCs, formed by trapping enzymes permanently bound to DNA (by exogenous sources like toxins), so we wanted to test if SPRTN is also able to cleave non-enzymatic DPCs, formed by proteins in close vicinity to DNA. These proteins can be trapped and form DPCs due to external influences or, of more importance, endogenously during cell metabolism^{38,110,125}. In search for a candidate protein with DPC-forming potential we chose histones due to their abundance and their close connection to DNA, making the possibility of spontaneous Histone-DPC formation high. Furthermore, post-translational modifications are shown to occur very frequently in all four core histone proteins. In an chromatin fractionation assay, similar to the previously explained but using a gradient gel to better separate low molecular weight proteins like histones, we discovered cleavage products of DNA-bound Histone H3 exclusively upon overexpressing ^{WT}SPRTN, not its mutated versions (Fig. 21). This result clearly demonstrates that Histone H3-DPCs become cleaved by SPRTN, and further supports the previously found impairment of DPC-cleavage in mutated SPRTN variants.

Interestingly, by analyzing SPRTN autocleavage, we could observe that the cleaved fragments of SPRTN were detected primarily in the total cell level and only at a low amount in the chromatin fraction, thereby showing that the cleaved SPRTN fragments are released from DNA (Fig. 21A+B). This finding is in accordance to the cytosolic localization of SPRTN-degradation products (Fig. 12).

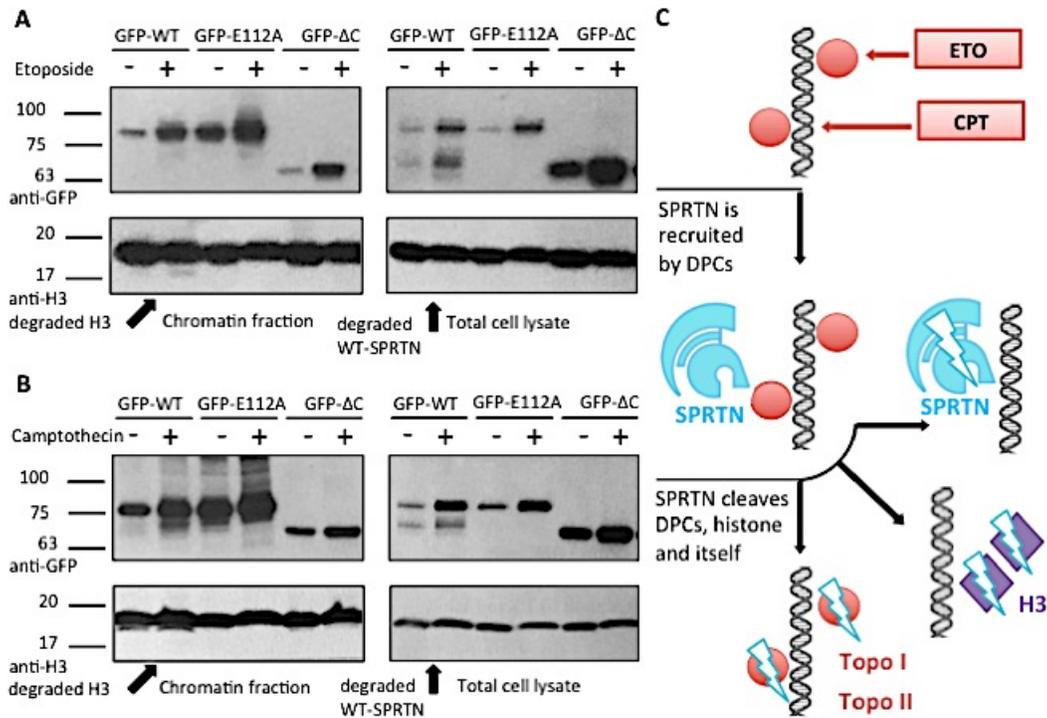


Figure 21 SPRTN has a promiscuous DPC-cleaving activity. **(A)** Immunoblotting of cells expressing GFP-tagged-WT, the catalytic inactive E112A or the truncated Δ C-variant treated with Etoposide 250 μ M for 24h or left untreated. Cell lysates were subjected to chromatin fractionation and cleavage of histone H3-DPCs and SPRTN was detected by simultaneous immunoblotting with H3- and GFP-specific antibodies. Cleaved products were detectable only in cells expressing WT, promoted upon Etoposide treatment. **(B)** Immunoblotting for SPRTN- and H3-cleavage upon treatment with Camptothecin 100 μ M for 24h. **(C)** Illustration of effect of DPC-induction (e.g. Etoposide treatment) on SPRTN function. After recruitment autocleavage and substrate cleavage of the induced DPC, as well as of surrounding DPCs is induced.

An additional aim was to investigate whether SPRTN acts on endogenously appearing DPCs (such as H3-DPCs) under normal conditions or only upon activation by DPC-inducing treatment. Furthermore we wondered if this SPRTN-activation causes cleavage of one specific DPC-substrate or promotes cleavage of other DPC-substrates in general. In order to test this we used DPC-inducing agents, which are not acting on histones (Etoposide or Camptothecin trapping solely Topo II or Topo I respectively) instead of an unspecific DPC-causing agent (as Formaldehyde) since a formation of H3-DPCs would naturally raise H3-DPC-cleavage and give no information about the specificity of SPRTN-activation^{102,118,120,130}. We found that H3-DPCs become cleaved endogenously. However we observed simultaneously that the H3-cleavage was noticeably higher

upon Etoposide treatment, as well as the auto-cleavage of SPRTN, indicating that SPRTN-activation (regardless by which inducing agent) promotes SPRTN activity and DPC-cleavage of various substrates generally and unspecifically (Fig. 21C). With these results we conclude that SPRTN is a DPC-cleaving protease acting on various different proteins crosslinked to DNA. This promiscuous proteolytic function, whereby SPRTN once activated cleaves various substrates could have dangerous consequences and therefore needs to be regulated and coordinated, a topic we also wanted to address.

4.4 SPRTN is relevant for cellular resistance to DPC-s

In view of the validation that SPRTN cleaves DPCs proteolytically we wanted to investigate the significance of this newly found function. Accordingly, we wanted to assess whether SPRTNs DPC-cleaving function is relevant for cell activity and to which extent it affects cell survival and growth.

4.4.1 SPRTN deficiency sensitizes cells towards DPC-causing agents

SPRTN's essentiality for cellular survival and embryonic development was already demonstrated by the fact that its complete loss of function is embryonic lethal in mice and that its impairment leads to growth deficits and progeroid features^{9,20}. On cellular level, SPRTN-KO reduces cell proliferation and induces cell death as already published and seen in our experiments as well (data not shown)⁹. Despite the multiple evidence of SPRTN's crucial role in cells, we wanted to address whether SPRTN provides tolerance to DPC-inducing agents. Therefore we analyzed the effect of DPC-inducing agents on cells with intact or deficient SPRTN-function by measuring ATP release as a marker for cell viability (Fig. 22A)¹⁷⁸. For this experiment we chose a SPRTN-knockout system instead of the previously used catalytic mutation to overcome any potential effects of higher expression levels of the ^{E112A}SPRTN mutant (the analogous effect of catalytic mutation and expression-deficiency of SPRTN on DPC-cleavage was already shown). As presumed, an evidently lower cell survival rate after exposure to DPC-causing agents is observed in SPRTN-KO-cells, thereby showing that SPRTN-deficiency increases cell sensitivity towards DPCs (Fig. 22B).

Additionally to the previously used DPC-inducing agents, we used Formaldehyde in order to provide support that SPRTN's function is important for cell survival not only due to its ability to cleave enzymatic DPC, but also due to its cleaving of non-enzymatic DPCs (e.g. histones) after gaining evidence of a SPRTN-mediated Histone H3 cleavage (Fig. 21A). FA is an agent known to induce a high level of crosslinking thereby causing various different and unspecific DPCs^{102,130}. It was of special interest additionally because Formaldehyde was shown to be endogenously produced in the vicinity of chromatin as a byproduct of

histone demethylation¹²⁹. As we could see that SPRTN deficiency increases cell sensitivity towards Formaldehyde, this indicates that SPRTN's function is important for cells in various conditions where they are exposed to DPCs (like alkylating agents, radiation etc.). Of course, this presumption needs to be further proven.

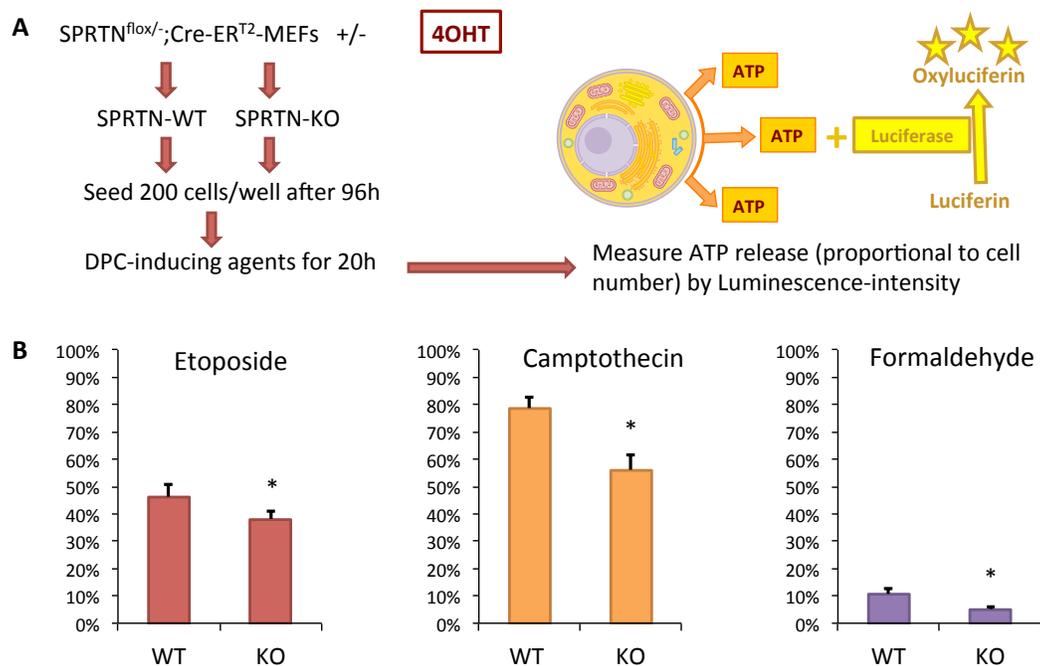


Figure 22 SPRTN-deficiency impairs cellular response to DPCs **(A)** Scheme of the survival assay comparing SPRTN-KO MEFs and MEFs with endogenous SPRTN-expression. Cells were treated with 4-OHT for 96h to induce KO of SPRTN or left untreated. After seeding of 200 cells and 20h treatment with Etoposide 50µM, Camptothecin 1000µM or Formaldehyde 500µM cells were subjected to analysis of ATP-level as a marker of surviving cell number. Relative cell number was measured by Luminescence of ATP-dependent Luciferase reaction. **(B)** Demonstration of SPRTN-dependent DPC-toxicity by comparing survival fractions of WT and KO cells. Survival fractions were determined as ratios of cells treated with DPC-inducing toxins and untreated cells. Data shown as Mean+SEM from 4 experimental replicates analyzed by paired t-test. An increased DPC-sensitivity is found in SPRTN-deficient cells.

4.4.2 SPRTN facilitates DPC-resistance and enables cell proliferation

As we wanted to examine the relevance of SPRTN's function to cleave DPCs, we further looked whether SPRTN can provide resistance against DPCs enabling not only temporary cell survival, but subsequent proliferation as well. To analyze proliferation ability we performed a colony formation assay, which shows the ability

of a single cell to form a colony (Fig. 23A). In this assay cells deficient for SPRTN were compared to normal cells to examine the negative effect of SPRTN-deficiency on resistance towards DPC-inducing agents. Therefore cells were exposed to DPC-inducing agents and allowed to form colonies in order to detect the fraction of cells with retained capacity to divide and proliferate.

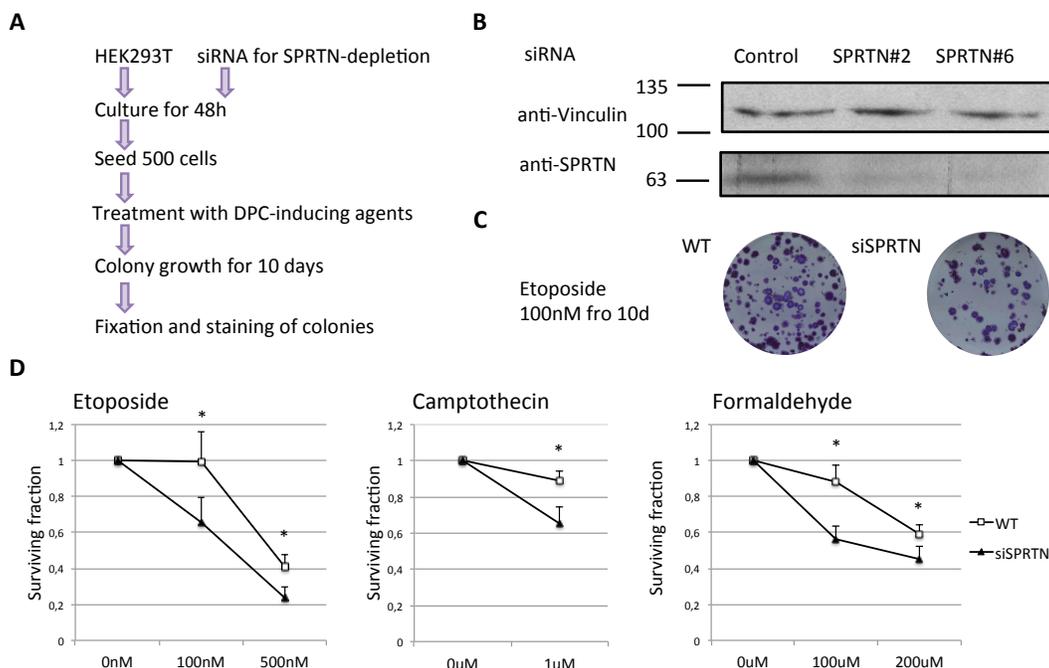


Figure 23 Cellular survival upon DPC-induction depends on SPRTN. **(A)** Scheme of the clonogenic experiment. SPRTN-deficiency was induced by siRNA and compared to endogenous SPRTN-expression. Cells were diluted and after treatment with DPC-inducing agents left to grow for 10 days. **(B)** Immunoblotting proving SPRTN-depletion upon siRNA-treatment. **(C)** Images of colonies formed from 500 cells after 10 days and stained with 0.5% crystal violet. **(D)** Survival curves upon different concentrations of DPC-toxins comparing SPRTN-deficient and endogenous cells. Data shown as Mean+SEM, from 3 replicates, analyzed by paired t-test. A decline in colony number upon increased DPC-toxin-concentrations is evident. A consistently reduced survival in SPRTN-deficient cells compared to control cells (statistically significant in all settings) is particularly obvious upon intermediate DPC-toxin-level proving SPRTN-dependent DPC-resistance. Higher concentrations of DPC-toxins resulted in raised cell death in both, SPRTN-deficient and control cells, with a lower dependence on SPRTN.

To specifically measure sensitivity towards DPC-inducing agents we determined survival fractions as ratios between colonies after treatment and untreated colonies, for cells intact or deficient for SPRTN respectively. SPRTN-deficiency was induced using a siRNA-system since the SPRTN-KO system leads to cell

culture decline after 6-7 days without any treatment (Fig. 23B). Consistent with our assumption, cells with deficient SPRTN-function are more sensitive to DPC-inducing agents compared to controls showing a significantly stronger survival- and proliferation-impairment upon treatment (Fig. 23c+D). This leads to the conclusion that SPRTNs DPC-cleaving ability has a dominant and important role in cellular resistance against DPCs.

4.5 SPRTN's C-terminus is dispensable for DPC-cleavage but relevant for regulation of SPRTN

In accordance to the importance of SPRTN's DPC-cleaving function being essential on one side, but additionally being potentially harmful while acting on various substrates on the other side, a regulation of SPRTN activity is necessary. We proposed that the C-terminus of SPRTN plays a role in this regulation, as a complete lack of the C-terminal SPRTN regions allows for cell viability, providing the essential DPC-cleavage, but shows a hypomorphic function⁹. This is demonstrated in Ruijs-Aalfs syndrome patients carrying a homozygous Δ^C SPRTN-mutation^{9,10}. This finding implies an important role of the C-terminus, which we wanted to further study.

4.5.1 SPRTN-level is replication-coupled

Several observations imply that SPRTN's function is related to DNA damage response (DDR). Our findings of chromatin binding (Fig. 14) in addition to previously demonstrated increased DNA-damage as a result of SPRTN-malfunction support this theory⁹. Accordingly, it was already shown that SPRTN shows a cell cycle dependent expression with an increased level when cells enter S phase¹⁷. We further proposed that the demonstrated DPC-cleaving function is connected to DNA replication, as DPCs primarily represent an obstacle for replication fork progression¹⁰⁶⁻¹⁰⁹. As SPRTN's high expression during S/G2 phase suggests a replication-dependent function, we wanted to provide evidence of this notion by analyzing SPRTN upon interrupted DNA replication, reasoning that a replication-blockage should result in a lower SPRTN level¹⁷. We made use of the fact that SPRTN is degraded mostly during mitosis, and wanted to determine if SPRTN degradation is increased upon replication-blockage¹⁷. Therefore a cycloheximide (CHX) chase assay was utilized, whereby the protein synthesis is blocked enabling to study its degradation. We used L-mimosine, which inhibits DNA-replication/elongation, to compare SPRTN-degradation in cells with intact and blocked replication. As presumed, SPRTN is degraded in a time-dependent

manner after replication is blocked demonstrating that SPRTN is primarily needed during active replication (Fig. 24B).

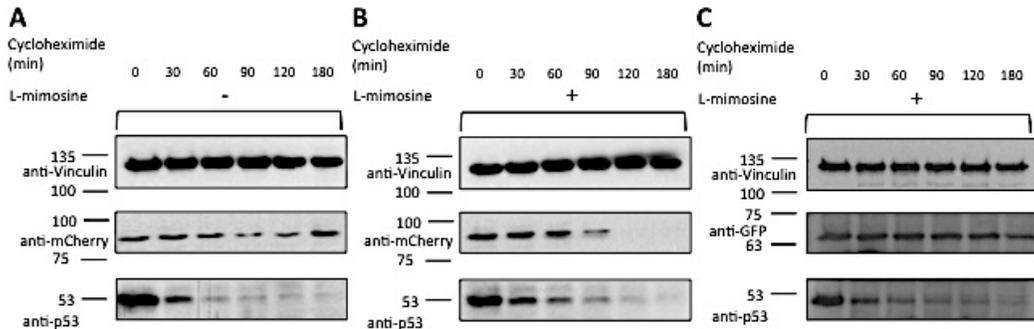


Figure 24 SPRTN-abundance depends on replication. HEK293T-cells transiently transfected with mCherry-SPRTN were subjected to translation inhibitor Cycloheximide and lysed after indicated time points to detect SPRTN-degradation. **(A)** No measurable degradation of mCherry-WT-SPRTN compared to p53 degradation. **(B)** Replication-blockage by L-mimosine induces SPRTN-degradation indicating a function of SPRTN related to DNA-replication. **(C)** Δ^C -SPRTN exhibits slower degradation compared to WT upon replication blockage.

However, the same experimental setting only including the Δ^C SPRTN-variant instead of WT SPRTN showed an impaired SPRTN-cleavage, thereby excluding the possibility that SPRTN is degraded due to L-mimosine-induced chelation (Fig. 24C). The finding that Δ^C SPRTN shows a slower degradation upon replication-blockage implies a role of the C-terminus in degradation of SPRTN. This is in accordance to the observed higher total cell level of Δ^C SPRTN in various experimental settings. The observed replication-dependent level of SPRTN strongly suggests a replication-coupled function and further leaves the interesting possibility that SPRTN functions together with the replication machinery, as supposed for the previously unknown DPC-cleaving protease found by Duxin et al.¹⁷⁴.

4.5.2 The C-terminus determines the nuclear localization of SPRTN

The fact that the C-terminal part of SPRTN is essential for correct nuclear localization of SPRTN has already been demonstrated by the defective cytosolic localization of the Δ^C SPRTN-variant⁹. This finding was confirmed in our visualization of cells stably expressing different variants of SPRTN used in our

studies (WT SPRTN, E112A SPRTN and $^{\Delta C}$ SPRTN respectively). Consistent with previous data the catalytic mutation has no impact on the localization, while a truncation of SPRTN with a C-terminal deletion diminishes nuclear localization. However, a weak but still present nuclear signal could be observed in $^{\Delta C}$ SPRTN (a finding in accordance to the low but still present $^{\Delta C}$ SPRTN -fraction bound to chromatin in immune-blotting assays) (Fig. 25C). We explain these observations by a possible entrance of a specific portion of the protein into the nucleus during a breakage of the nuclear envelope in mitosis.

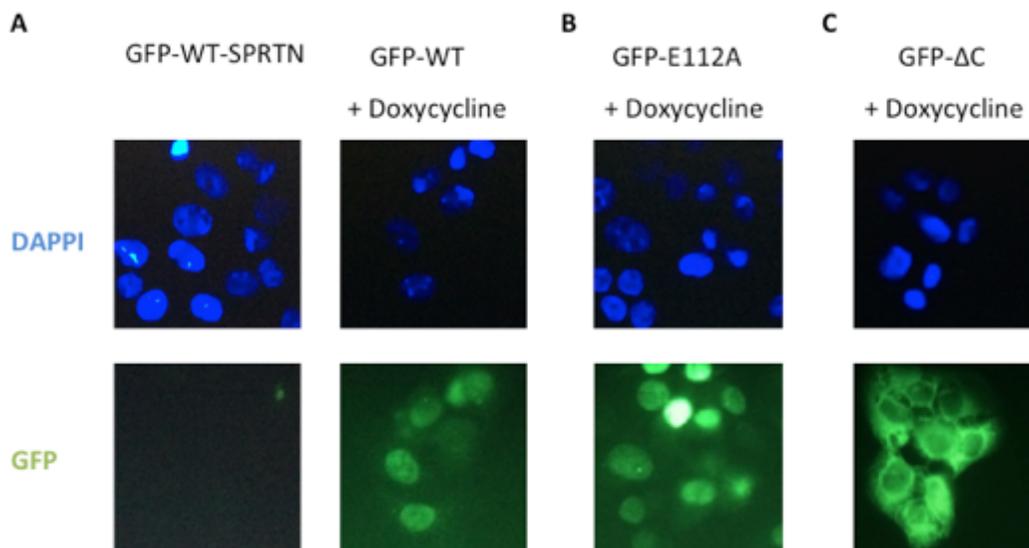


Figure 25 Cellular distribution of SPRTN depends on its C-terminus. Microscopic view of U2O3 cells with an inducible SPRTN-expression. Different GFP-tagged SPRTN-variants under a Flp-in TREX promoter are expressed upon Doxycycline-treatment. **(A)** WT-SPRTN exhibits a regular intranuclear expression by green GFP-fluorescence similar to blue Dapi-fluorescence of DNA after induction by Doxycycline. **(B)** The mutation of the catalytic domain is not affecting the intranuclear localization as displayed by the proper localization of the E112A-SPRTN-mutant. **(C)** The truncated ΔC -SPRTN missing the C-terminus exhibits a cytoplasmic localization.

However, the main question is why the C-terminal part of SPRTN is a prerequisite for a correct nuclear localization. As the nuclear envelope acts as a clear barrier allowing only small molecules to pass through the nuclear pores, all other molecules depend on two specific nuclear transport systems enabling either import or export¹⁸⁰. Proteins transported by these systems need special sequences, which are recognized either by importins or exportins (proteins responsible for the transport through nuclear pores). We reasoned that SPRTN

has to contain such a signal, a nuclear localization signal (NLS) in its C-terminal part, in order to be properly transported via importins into the nucleus. Such NLS requires a lysine residue being followed by two additional basic residues (K-K/R-X-K/R)¹⁸¹. Consistent with our presumption a bioinformatic analysis of the amino-acid sequence of SPRTN showed three potential NLS-sequences, fulfilling these criteria (Fig. 26A). Of these candidates, the NLS sequence at aa407-411 (KRPRL) seemed as the most promising one, since it is located in the C-terminal part and additionally highly conserved among vertebrates (Fig. 26B).

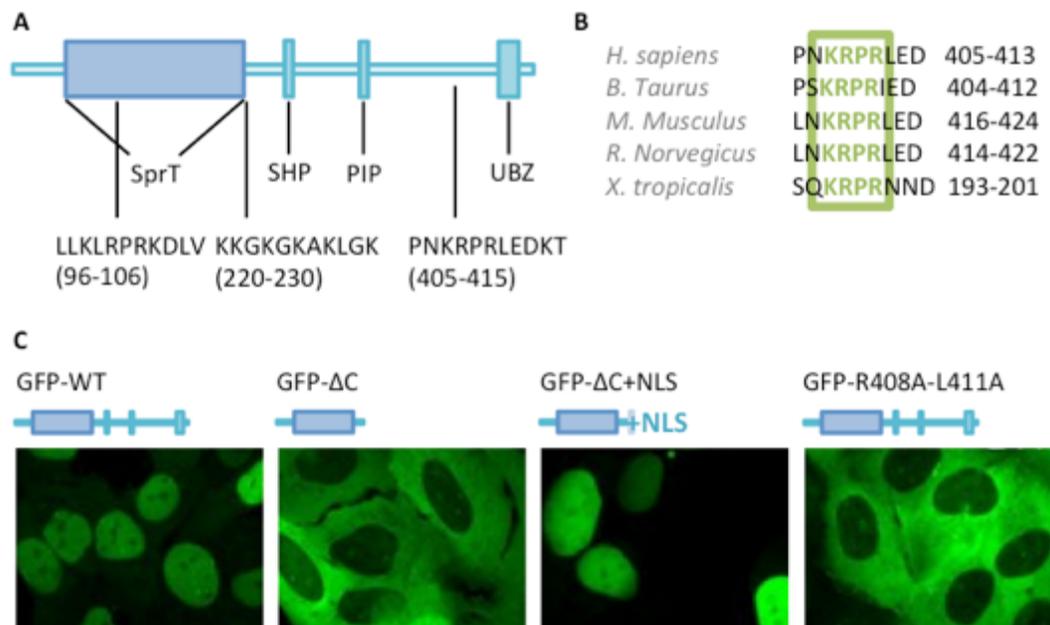


Figure 26 The nuclear localization of SPRTN results from a C-terminally located NLS (nuclear localization signal). **(A)** SPRTN encompasses three potential K-K/R-X-K/R motifs typical for NLS along its amino-acid-sequence. **(B)** One of three potential NLS-sequences is located at the C-terminus (KRPRL-sequence at aa407-411) and demonstrates a strong sequence-homology in SPRTN genes from various vertebrate species. **(C)** Microscopic demonstration of a cytosolic mislocalization upon mutation of the NLS sequence compared to intranuclear WT-SPRTN (with thanks to J. Lopez-Mosqueda, who performed the microscope localization experiment).

Indeed our hypothesis of a NLS at aa407-411 being necessary and sufficient for proper nuclear localization could be validated by the fact that upon mutation of this sequence (^{R408A-L411A}SPRTN) SPRTN is located in the cytosol, similar to ^{ΔC}SPRTN (Fig. 26C, data from experiments performed by J. Lopez-Mosqueda). Taken together our investigation serves as the first explanation of the mechanism underlying SPRTN nuclear localization.

4.5.3 DPC-cleavage is contained in the SprT-domain

Since the importance of SPRTN's C-terminal region for its proper localization was demonstrated and the underlying mechanism explained by the identification of the NLS, we further aimed to determine whether the proper localization is the main role of the C-terminus in DPC-cleavage. In order to do so, we wanted to see if a restoration of the proper nuclear localization is enough to establish an adequate DPC-cleavage. Therefore we established a hybrid SPRTN-variant containing the N-terminal SprT-domain with an additional NLS-sequence but missing the remaining parts of the C-terminus and named it $\Delta C+NLS$ SPRTN. Using this model, we tested the DPC-cleaving ability by measuring H3-cleavage in cells expressing either $\Delta C+NLS$ SPRTN, WT SPRTN or known mutant-variants in comparison to each other. We measured the DNA-recruitment by comparing chromatin fractions with total cell lysates (Fig. 27). Consistent with our previous finding we could detect a proper H3-cleavage upon overexpression of WT SPRTN, while $E112A$ SPRTN and ΔC SPRTN were not able to produce measurable cleaved fragments of H3. The addition of the NLS sequence to the functionally impaired ΔC -variant ($\Delta C+NLS$ SPRTN) restores a H3-cleavage comparable or even higher than in WT SPRTN.

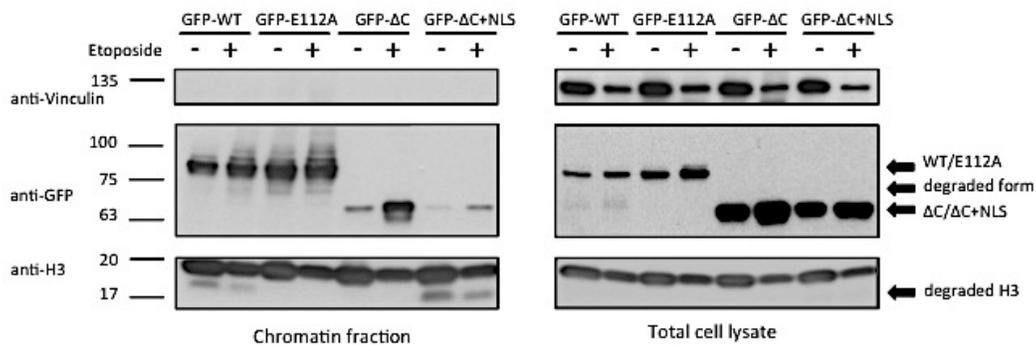


Figure 27 SPRTN's cleavage-function is determined by a proper nuclear localization. Simultaneous analysis of DNA-bound SPRTN-level and histone-cleavage. HEK293T-cells expressing GFP-tagged SPRTN-variants (WT, E112A, ΔC or the hybrid ΔC with added NLS respectively) were subjected to parallel chromatin fractionation and total cell lysis. Cleaved histone fragments are detected in chromatin fractions in cells expressing full-length WT and the ΔC -mutant with a rescued nuclear localization ($\Delta C+NLS$) but not in cells expressing the cytosolic ΔC -mutant or the catalytic inactive E112A. The $\Delta C+NLS$ shows a lower abundance on chromatin despite high cleavage activity.

This finding demonstrates that SPRTN's function is contained within its N-terminal region containing the SprT-domain, since this part alone (represented by the Δ^{C+NLS} SPRTN-variant) upon proper localization is as active as the WT SPRTN. Δ^C SPRTN showed impaired H3-cleavage and decreased DNA-recruitment, despite its total expression-level being higher than in WT SPRTN or E112A SPRTN, further suggesting its mislocalization as the explanation for its weak DPC-cleavage. However, opposite to our expectation, Δ^{C+NLS} SPRTN did not show a higher, but instead a lower level in the chromatin-bound fraction. Since its total expression level was higher than of WT SPRTN or E112A SPRTN and a proper nuclear localization was demonstrated separately (by microscopic detection of intranuclear localization of the Δ^{C+NLS} SPRTN-variant, data not shown) the only explanations are that Δ^{C+NLS} SPRTN is recruited less to chromatin than WT SPRTN (staying intranuclear but not chromatin bound) or that it is faster removed from chromatin. As a high level of H3-cleaved fragments (as a marker for DPC-cleavage) in Δ^{C+NLS} SPRTN was clearly detected and this function requires a previous recruitment to chromatin, a weaker chromatin-recruitment of Δ^{C+NLS} SPRTN is excluded. Thereby it is possible that a DNA-binding region is located in the C-terminal region of SPRTN, in addition to one located within the SprT-domain. The fact that the catalytic inactive E112A SPRTN shows the exact opposite effect - a higher chromatin-bound level with a lower cleavage activity - further supports this explanation. Taken together, our observations show that the C-terminus of SPRTN is not essential for the DPC-cleaving function and the deletion of the C-terminus (as in Δ^C SPRTN) does not alter the DPC-cleaving ability itself but compromises its execution by mislocalizing SPRTN to the cytosol and perhaps abrogates its ability to properly bind to DNA.

4.5.4 The C-terminus is important for SPRTN-SPRTN interaction

As we found that SPRTN cleaves itself *in trans*, this supports an interaction of two SPRTN molecules. We wanted to further analyze this interaction wondering which part of SPRTN is included in it. Therefore we performed a co-immunoprecipitation using a mCherry-labeled inactive form (mCherry-

^{E112A}SPRTN) with different GFP-tagged variants of SPRTN, full-length or missing either the C-terminal (Δ^C SPRTN) or the N-terminal half (Δ^N SPRTN). As expected, we could observe a binding of two full-length SPRTN-molecules. Interestingly, we found that this binding is preserved only in the Δ^N SPRTN-variant and missing in the Δ^C SPRTN, meaning that the interaction of two SPRTN molecules depends on the C-terminus (Fig. 28).

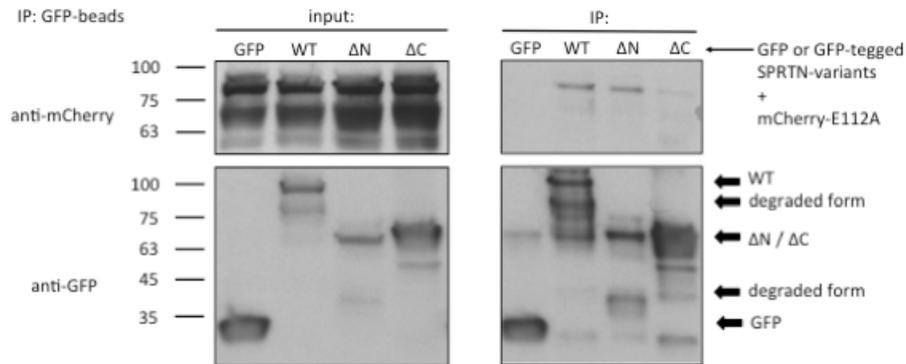


Figure 28 SPRTN-dimerization depends on its C-terminus. Co-immunoprecipitation with GFP-binding beads from cells expressing inactive mCherry-tagged-E112A and either GFP, GFP-tagged-WT or the C-terminal (Δ^N) or N-terminal half (Δ^C) of SPRTN. The mCherry-E112A is pulled down by WT and the C-terminal part (Δ^N) demonstrating an endogenous direct or indirect dimerization or interaction enabled by the C-terminus.

However, the N-terminal part is unable to perform the observed *in trans* auto-cleavage as it misses a protease domain. Furthermore, the C-terminus was already shown to be dispensable for substrate cleavage. Therefore the most reasonable explanation for this finding is that SPRTN recognizes another SPRTN-molecule or an interactor via its C-terminus enabling a direct or indirect dimerization and although this binding is not necessary for the substrate- or autocleavage function, it may promote it by proper SPRTN localization to its potential substrate, thereby demonstrating an additional function of the C-terminus of SPRTN. This theory is in accordance to the clinical observation that the patient with homozygous expression of Δ^C SPRTN was alive but symptomatic, meaning that SPRTNs essential function is still preserved but obviously disturbed, implying that the C-terminus plays a regulatory role, maybe via binding with interacting molecules or recruitment of SPRTN before acting on DPCs⁹.

4.6 RJALS-syndrome shows defective DPC-cleavage

After gaining new insight into SPRTN's function and providing support that SPRTN acts as a DPC-cleaving protease, we further wanted to confirm that this function of SPRTN is impaired and responsible for the clinical manifestations observed in RJALS-syndrome associated with SPRTN mutations. Increased DNA replication stress with slower replication, replication fork stalling and DSBs has already been shown in cells from patients with SPRTN-mutations⁹. These findings as mechanisms underlying the genetic instability with progeroid phenotypes and tumorigenesis could potentially be caused by an increased DPC-level, since DPCs represent bulky DNA lesions blocking DNA replication and leading to DSBs¹⁰⁶⁻¹⁰⁹.

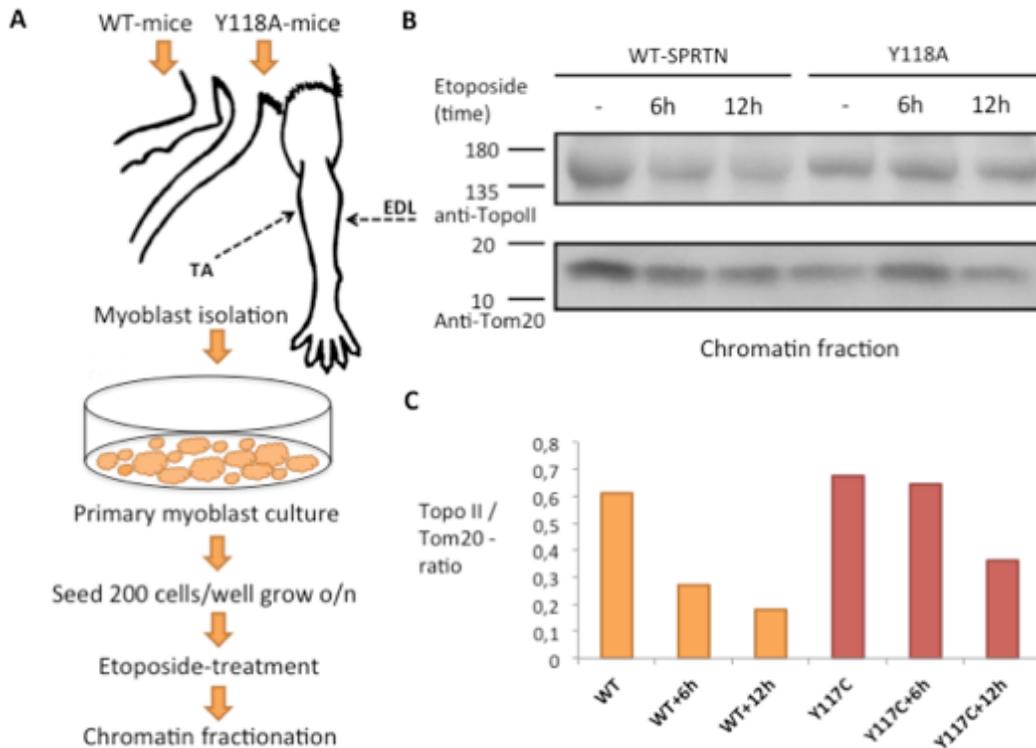


Figure 29 DPC-cleavage is impaired in mice with SPRTN-mutation. **(A)** Mice homozygously expressing WT-SPRTN or Y118A-SPRTN were used to isolate muscle fibers followed by cultivation of the isolated myoblasts. Cells were treated with Etoposide for the indicated time and harvested. **(B)** Immunoblotting of the chromatin fraction of myoblasts with Topoisomerase II-antibodies and Tom20-antibodies as a loading control. **(C)** Diagram comparing Topoisomerase II-DPCs after correction to the loading control demonstrates a higher abundance of Topo II-DPCs in myoblast from mice carrying the Y118A-SPRTN-mutation, especially upon Etoposide treatment.

However, an accumulation of DPCs in patient cells needed to be demonstrated. Therefore we made use of a newly established mouse model carrying the mouse variant of the clinical mutation ^{Y117C}SPRTN, namely a Y118A-mutation (this model was ordered by Dikic laboratory after the discovery of the SPRTN-mutations underlying RJALS-syndrome). Muscle cells from Y118A/Y118A mice as well as WT/WT mice were isolated and chromatin-bound Topo II-level as a marker of DPCs was measured to see whether DPC-accumulation is increased spontaneously in cells carrying patient mutations *in vivo* (Fig. 29A)¹⁸². As an additional support of impaired DPC-cleavage, cells were exposed to Etoposide treatment in order to determine the ability of the clinical mutant to accurately respond to high DPC-levels (as a constant exposure to DPC-toxins is known). Indeed, cells harboring the ^{Y118C}SPRTN-mutation showed a raised Topo II-DPC-accumulation compared to cells with ^{WT}SPRTN, additionally pronounced upon Etoposide treatment (Fig. 29B+C).

Since our previous results based on cell models either overexpressing or silencing (by siRNA/knockout) SPRTN, this finding demonstrates that DPC-cleavage is compromised as a result of SPRTN-mutation under normal expression levels as well. Moreover, it demonstrated for the first time that DPCs accumulate *in vivo* in an organismal model - an animal model carrying a SPRTN-mutation analogous to that found in Ruijs-Aalfs syndrome patients. Further evidence of the impaired DPC-repair in the clinical mutant was demonstrated in lymphoblastoid cells derived from patients with RJALS-syndrome (ICE assay showing defective removal of Topo II-DPCs performed by Jaime Lopez-Mosqueda, data not shown).

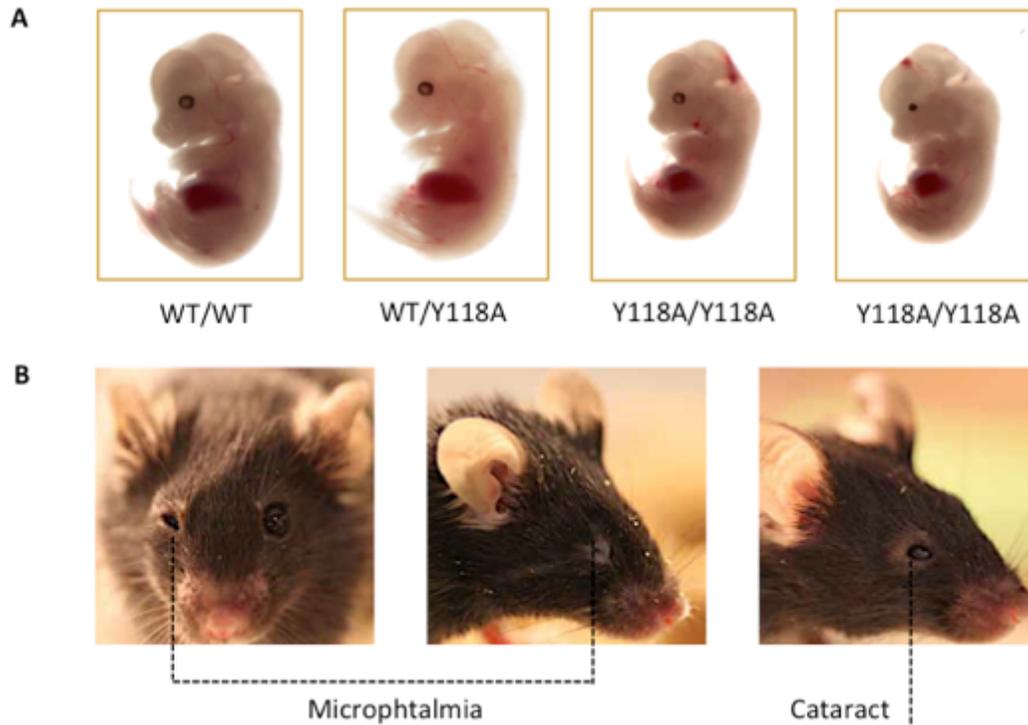


Figure 30 Anomalism found in mice carrying a Y118A-SPRTN-mutation. **(A)** Embryos sacrificed at week 3 demonstrate a dwarfism among Y118A/Y118A mice compared to heterozygous or WT mice. **(B)** Representative images of eyes from mice with a homozygous Y118A/Y118A-SPRTN-mutation showing microphthalmia or cataracts respectively.

The Y118A/Y118A-mice exhibited phenotypes similar to the patients as well as to mice with hypomorphic SPRTN expression described by Maskey et al. 2014^{9,20}. We found a decreased ratio of Y118A/Y118A-offsprings born alive in the litters. An analysis of embryos sacrificed in week 3 revealed a dwarfism among the Y118A/Y118A-mice compared to their control littermates (Fig. 30A). An according growth difference was found in adult mice as well with smaller body sizes of the mice carrying a SPRTN-mutation. Additionally, anomalies in eye development including microphthalmia and cataracts were observed at a higher frequency among Y118A/Y118A-mice compared to control WT mice (Fig. 30B). The observed phenotypes recapitulate the short stature and ophthalmological abnormalities observed in patients with RJALS, thereby further supporting this model.

As previously mentioned, the ^{Y118A}SPRTN-mutation in mice is equal to the human ^{Y117C}SPRTN-mutation lying close to the protease core and impairs SPRTN's DPC-cleavage activity (as recapitulated by the ^{E112A}SPRTN mutation).

The other clinical mutation found in RJALS-patients, the ΔC SPRTN-mutation was also already shown to be defective in DPC-cleavage, and the underlying mechanism explained by the mislocalization of this SPRTN-variant. Taken together, these molecular explanations of the impact of both clinical mutations on SPRTN's DPC-repair activity, the demonstrated raised sensitivity upon SPRTN-malfunction and the proof of DPC-accumulation upon clinical SPRTN-mutation *in vivo* collectively provide convincing evidence for a causal relation between defective DPC-repair and RJALS-syndrome.

5 Discussion

As this study identified SPRTN as the first protease that cleaves DNA-protein-crosslinks (DPCs) in metazoans and demonstrated the functional relevance of this activity, these novel findings of an essential previously unknown function of SPRTN initiated further validation in different experimental set-ups. This was done by our group as well as by several other groups simultaneously or subsequently to our observations. Therefore, a further aim of this work additionally to demonstrating our findings is to collectively analyze the published findings describing SPRTN in comparison with the results of this study as well as to give an overview of the existing data about SPRTN and thereby establishing a model of SPRTN's function.

We proposed that SPRTN has an essential function and that this function lies in its ability to cleave DPCs. Accordingly, a genome wide CRISPR screen revealed that SPRTN indeed is an essential protein¹⁸³. This is in accordance to the previously published data that a complete SPRTN knockout leads to embryonic lethality²⁰, as well as to our own observation of the lethal effect of SPRTN-malfunction on cellular survival. Furthermore, the proteolytic repair of DPCs was demonstrated to be crucial for DNA-replication upon DPC-induction and subsequently for genomic integrity and cellular survival^{37,184,185}. Noticeably, an unbiased screen for proteases involved in DPC-cleavage pinpointed SPRTN together with the proteasome as main players¹⁸⁶. As a result of this important role of SPRTN, its function demands a detailed analysis of its activity and its regulation.

5.1 Structure and protein-interaction of SPRTN

Solving the protein structure as well as its functional domains is important in unraveling the function of a protein. In the case of SPRTN, several functional domains have already been identified and partially characterized prior to this study. Accordingly, a N-terminal zinc-metalloprotease domain as well as further C-terminal domains including a p97-binding domain, an ubiquitin-binding domain and a PCNA-binding domain were already identified in SPRTN's amino acid sequence¹²⁻¹⁴. Our work contributed to the understanding of SPRTN's structure by revealing an additional nuclear localization signal (NLS) at aa407-411 and thereby explaining SPRTN's cellular localization and the mislocalization of ^{ΔC}SPRTN allele that is missing that NLS (Fig. 31).

In addition to the nuclear localization, the DNA-binding of SPRTN also needed to be explained. Our efforts provided evidence that SPRTN binds to DNA *in vitro* demonstrated by an electromobility shift assay (EMSA) a slower movement of DNA when incubated with purified SPRTN¹⁸⁷. Similar observations of SPRTN-DNA-binding were demonstrated by several other groups, additionally confirming the importance of the C-terminal part of SPRTN for DNA-binding, as shown in this study. Initially 5 DNA-binding regions in the C-terminus have been proposed¹⁸⁸. As the truncated ¹⁻³¹⁰SPRTN binds DNA similar to ^{WT}SPRTN and ¹⁻²¹⁶SPRTN in our work clearly less, the region following the SprT-domain has to be crucial for DNA-binding, more precisely the aa200-250, as later revealed by mass spectrometry analysis^{183,187,189}. Further work identified a DNA-binding box (KKGK) at aa220-223, named basic region (BR). This region is evolutionary conserved and a mutation of it shows a highly decreased affinity to any type of DNA¹⁸⁹. As the DNA-binding is not completely abolished upon mutation of the BR (^{KKGK}SPRTN to ^{AAAA}SPRTN) an additional DNA-binding region has to be contained in the N-terminal part. Although some experiments demonstrated a missing DNA-binding in truncated SPRTN-variants as ¹⁻²²⁰SPRTN or ¹⁻²¹⁶SPRTN, a conserved binding of these forms was shown *in vitro* and by NMR analysis demonstrating a direct binding of the SprT-domain, which was confirmed by the SPRTN crystal structure^{187,188,190}. The differences can be explained by different experimental set-

ups and by the fact that the BR binds dsDNA and ssDNA, while the SprT-domain can bind only ssDNA (Fig. 31C)¹⁹⁰.

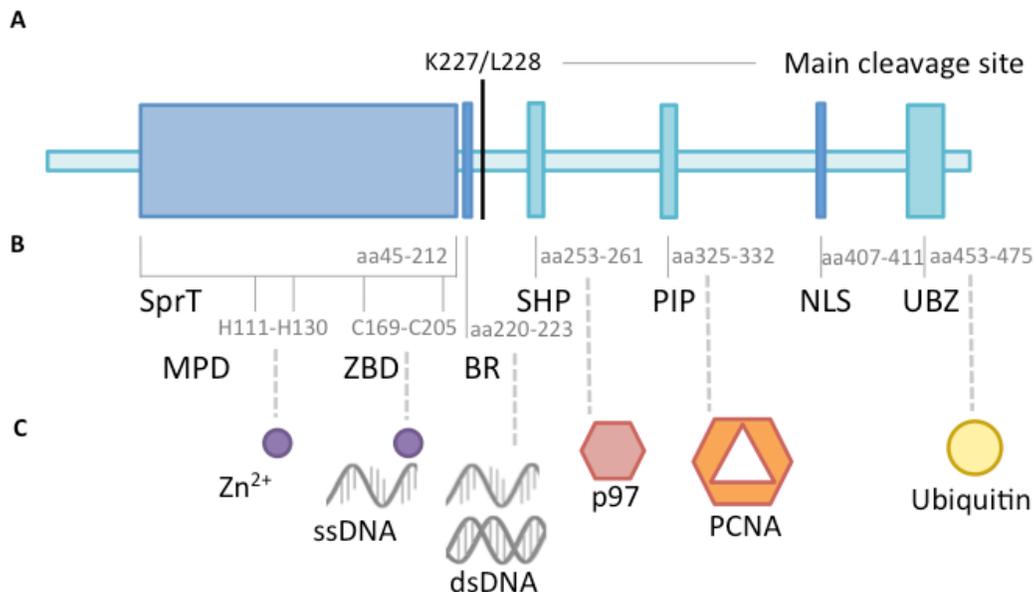


Figure 31 SPRTN and its clinical mutations. **(A)** The full length WT-SPRTN is a 489-amino-acid-long protein Zn-binding metalloprotease consisting of multiple domains. Among various different sites where SPRTN is getting cleaved the most prominent one is at K227/L228. **(B)** 6 identified SPRTN-domains including the previously known SHP, PIP and UBZ interacting motifs binding to p97, PCNA or Ubiquitin respectively and the nuclear localization signal (NLS) discovered in this study and the later identified basic region (BR). The previously sparsely characterized SprT-domain was shown to consist of 2 subdomains, a N-terminal metalloprotease domain (MPD) and a following zinc binding domain (ZBD). **(C)** Both SprT-subdomains bind one Zn²⁺-ion respectively. SPRTN's DNA-binding is enabled by the BR binding ssDNA and dsDNA and additionally by the ZBD binding solely ssDNA. Further interacting proteins are demonstrated as well.

SPRTN's main function as a DPC-protease, as demonstrated in this work and confirmed *in vitro* by our and other groups, was additionally resolved by the crystal structure of the SprT-domain (aa26-214) in complex with a 5nt-oligo dC-DNA¹⁹⁰. The SprT-structure revealed two subdomains important for proteolysis linked by a highly conserved loop. The N-terminal subdomain, named metalloprotease-domain (MPD), binds a Zn²⁺-ion (with H111, E112, H115 and H130) essential for substrate cleavage, as previously predicted by Phyre server¹⁹. Another Zn²⁺-ion is additionally bound by the more C-terminally positioned zinc-binding domain (ZBD). A binding of two Zn²⁺-ions was previously shown¹⁸⁸. Both these subdomains contain 3 alpha-helices and a three-banded beta-sheet respectively and are linked

by a loop, whereby the active site lies in a groove formed by these subdomains and shielded by the ZBD. This enables a substrate selection allowing only proteins with flexible regions able to fit in the active site to be cleaved by SPRTN (Fig. 32B)¹⁸³.

Insights into the structure of the SprT-domain enable an explanation of the effects of SPRTN mutations. The inactive ^{E112A}SPRTN used in this study and most SPRTN-experiments is impaired as the catalytic core in the SprT-family (HExxH) coordinating the Zn²⁺-ion depends on the glutamic acid that polarizes a H₂O molecule in order to enable a nucleophilic attack on the substrate protein (Fig. 10 and 32C)^{19,176,177,183}. Thereby a mutation to alanine abolishes the cleavage ability. Noticeably, a structural alteration by this mutation was excluded, leading to the conclusion that the defects seen in ^{E112A}SPRTN are solely a result of its catalytic impairment. However, the Y117C-mutation found in the patient is not a part of the HExxH-motif. Although proposed to be involved in substrate binding based on the Wss1-structure, the SprT-structure revealed that the tyrosine is part of the alpha-helix and important for its orientation through hydrophobic bounds with side chains^{183,190}. The mutation to a cysteine therefore causes an alteration of the alpha-helix structure impairing the catalytic activity of the HExxH core (Fig. 32C).

While SPRTN's proteolytic function is performed by the MPD-subdomain containing the catalytic core, the protease activity depends additionally on the ZBD-subdomain. Additionally to its function in substrate selection by shielding of the active side and restricting it to flexible protein-substrates, the ZBD is important for selective DNA-binding as well. It forms an aromatic pocket (with Y179 and W197), which binds DNA nucleotides¹⁹⁰. Since it is only able to bind nucleotides and not base pairs, it enables binding to ssDNA, restricting dsDNA, different to the BR region additionally explaining SPRTN's preferential binding to ssDNA¹⁸⁹. The importance of this ZBD-DNA-binding is demonstrated by the deficient auto- and substrate-cleavage ability upon mutation of the aromatic pocket. As a result SPRTN cleaves DPCs in the vicinity of ssDNA, as demonstrated in several other experiments^{183,191}. Another important part of the SprT-domain is a positively

charged patch which plays a determining role in DNA-binding and consequently in the proteolysis as well.

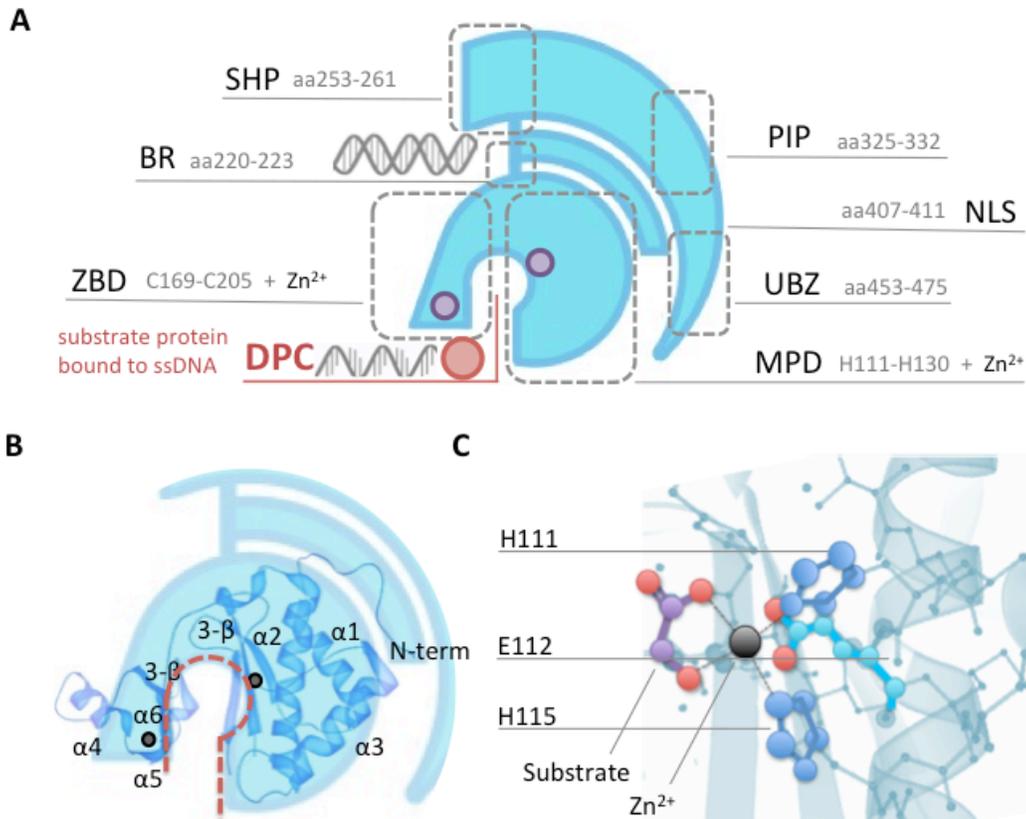


Figure 32 Scheme of SPRTN as a Zn-binding metalloprotease with additional C-terminal interacting motifs. **(A)** The domains are encircled and the respective amino-acid sequences listed for each domain. The SprT-domain consists of 2 subdomains (namely MPD and ZBD), each binding a Zn²⁺-ion marked by purple circles. The 2 subdomains form a pocket, wherein the DPC-substrates are cleaved. **(B)** Structure of the SprT-domain consisting of the bigger N-terminally located MPD and the smaller ZBD. Each subdomain is build up of 3 alpha-helices and a three-stranded beta-sheet and binds a Zn²⁺-ion marked by circles. A pocket is formed between the MPD and ZBD-subdomain and contains the catalytic active HEXxH motif in its center. This pocket represents the proteolytic core of SPRTN (marked with red). **(C)** More detailed view of the HEXxH-motif in the center of the active site pocket of the SprT-domain. The Zn²⁺-ion in the center is coordinated by 2 Histamines and together with the Glutamic acid E112A it interacts and cleaves the substrate DPC-protein.

Although most studies focused primarily on the essential N-terminal part of SPRTN, the C-terminal PCNA-binding domain, the PIP-box, was also further investigated¹⁹². Thereby PCNA-binding was shown in a model and the affinity of SPRTN's PIP-box was shown to be low, in accordance to observations on cellular

level, and lower than the affinity of PIP-boxes of TLS polymerases, thereby proposing that SPRTN is recruited to PCNA but in vicinity on TLS polymerases loses its PCNA-binding in favor of TLS¹².

An autocleavage activity of SPRTN was demonstrated in this work and further validated. The cleavage-site was also of interest, especially in terms of further functionality of the cleaved fragments. A mass spectrometry analysis revealed at least 5 autocleaved fragments and further analysis determined the major cutting-site at K227/L228 with other additional cleavage sites in accordance to our finding of the main cleaved fragment being approximately half the size of full-length SPRTN and accompanied with several other smaller fragments as further cleavage products (Fig. 31A)^{183,188}. The importance of the autocleavage will be discussed later.

5.2 Relation to the supposed yeast orthologue Wss1

A relation between SPRTN and Wss1 was been proposed relatively early after the discovery of SPRTN, primarily due to their domain-organization^{17,183,184}. Our finding of a similar DPC-cleaving activity of these two proteins further supports this idea. Additionally, a phylogenetic analysis using a reciprocal BLAST search for proteins homologous to Wss1 revealed a common ancestry and a conservation of eukaryotic Wss1/SPRTN proteases with three main groups, a Wss1 branch and a UBL-Wss1 branch in plants and fungi as well as a Spartan branch in metazoans, whereby a higher divergence was found in metazoans^{183,193}.

However, a closer analysis of the amino acid sequence by our group revealed a poor conservation of the sequence between SPRTN and Wss1¹⁸⁷. The protease domain, as the main connecting part thereby showed especially weak preservation as additionally proven by the comparison of their structures^{187,190}. Although the catalytic core with the HExxH-motif shows an undoubted similarity, the remaining sequences show limited analogy. On structural level the protein folding of the MPD-subdomain with 3 alpha-helices and a three-banded beta-sheet is comparable to the WLM-domain in Wss1 with 4 tightly packed alpha-helices and a four-stranded beta-sheet¹⁸³. Despite the coinciding active side organization a major difference between the protease domains results from the exposure of the active site of Wss1 compared to the rather closed active site of SPRTN, at the base of a groove shielded by the additional ZBD-domain, which is completely missing in Wss1^{183,190}. With this substrate binding cleft SPRTN can cleave only proteins with a flexible loop and subsequently shows a reduced substrate-promiscuity compared to Wss1 with an exposed catalytic core. Furthermore the shielded active site of SPRTN is opened only upon binding to ssDNA, whereby SPRTN displays an additional regulation of a potentially promiscuous proteolysis, missing in Wss1. The higher conservation of the SprT-domain among metazoans further argues against a similarity of SPRTN/Wss1 proteases.

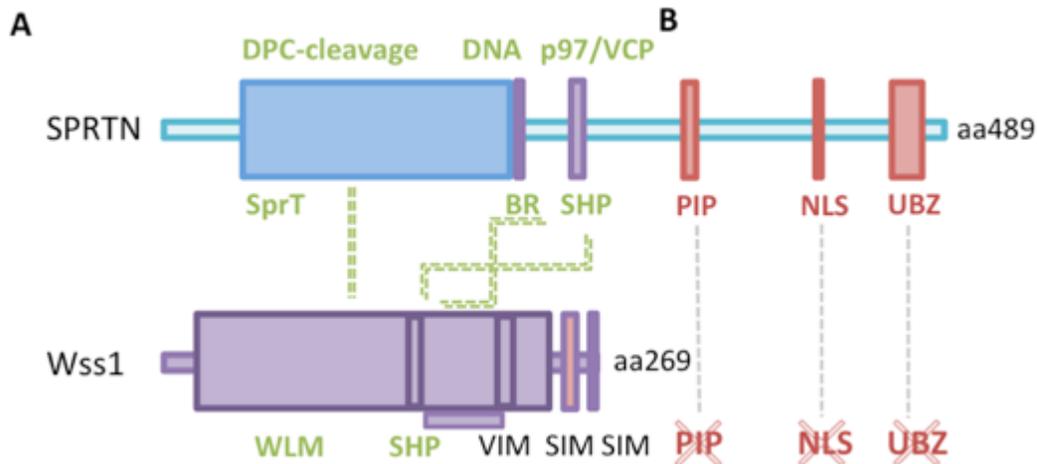


Figure 33 Comparison of SPRTN and Wss1 regarding their amino-acid-sequences and domain organization. **(A)** Similarities including DNA-binding, DPC-cleavage and interaction with p97 or Cdc48 are marked in green. **(B)** Differences (missing of a ubiquitin- or PCNA-interaction and an NLS sequence in Wss1) are marked in red. While SPRTN interacts with ubiquitin, Wss1 interacts with SUMO. Additional to the domain differences, Wss1 is about half the size of SPRTN.

A correlation in domain organization between SPRTN and Wss1 is exhibited by the N-terminal protease domain and a segregase-binding domain (p97 or the yeast homologue Cdc48 respectively)^{183,194}. Nonetheless there are two major discrepancies in the domain organization: first, SPRTN possesses an additional PCNA-binding motif (PIP-box) lacking in Wss1, which was shown to be relevant for SPRTN's activity and second, the interactor-recognizing motifs differ, whereby SPRTN contains a ubiquitin-binding zinc finger (UBZ) while Wss1 harbors two SUMO-interacting motifs (SIMs)^{12,14,16-18}. Each of these two interactors is important for proper functioning of SPRTN or Wss1 respectively (Fig. 33)^{175,183,195}.

Functionally a relation between SPRTN and Wss1 seems very obvious since they both act as DPC-cleaving proteases and show a DNA-dependent function^{183,196,197}. Furthermore their levels exhibit a tight regulation by a self-cleaving activity¹⁹³. However, despite the common DPC-cleaving activity there are indisputable differences regarding this function. Accordingly, our group demonstrated that SPRTN is not able to compensate Wss1 function in yeast, as it cannot promote cell growth in yeast lacking Wss1/Tdp1¹⁸⁷. In addition, while SPRTN is clearly essential in higher eukaryotes, Wss1 is dispensable in yeast, demonstrating the distinct roles these two proteins play^{20,175,187,18,199}.

In conclusion, SPRTN and Wss1 show a homologous relation with a common phylogenetic ancestry, a partially comparable domain organization and a common function in DPC-cleavage^{196,200}. Nonetheless, they show differences in amino acid sequence, protein-interaction and the functional performances greater than expected and consequently SPRTN is not a complete functional orthologue of Wss1.

5.3 Activities of SPRTN and their cellular functions

5.3.1 SPRTN exhibits an autocleavage activity

The results presented in this study demonstrate a new function for SPRTN as a protease. As the first evidence of proteolytic activity, we were able to show SPRTN self-cleavage, which happens *in trans* (Fig. 11). Our data on cellular level was confirmed by additional *in vitro* experiments. As demonstrated by our group and in several other studies, purified SPRTN is rapidly degraded by itself showing different cleaved fragments, whereby the cleavage happens from the C-terminal part and the main product is a 227aa-remnant, which is similar to the Δ^C SPRTN as found in our work (Fig. 11B) and is still proteolytically active^{183,188,190,191}. A further cleavage to a 218aa-remnant leads to a loss of activity¹⁸⁸. As expected, E^{112A} SPRTN exhibits no, and Y^{117C} SPRTN a diminished or reduced *in vitro* self-cleavage of 20%^{183,188}. An *in trans* autocleavage was confirmed *in vitro* correlating to the results shown here as the inactive E^{112A} SPRTN is cleaved by WT SPRTN^{183,187,188,190}. Importantly, the *in vitro* self-cleavage is dependent on the presence of DNA, regardless if ssDNA or dsDNA, as verified in several settings, but with ssDNA being the stronger inducer^{183,187,188,190,191}. Additionally, the autocleavage is increased by DPC-inducing agents as shown for Etoposide in this study (Fig. 21) or for Formaldehyde in another demonstrating a connection of SPRTN's substrate and^{183,188}. A causal relation of this observation still needs to be revealed.

5.3.2 SPRTN's main function is the proteolytic cleavage of DPCs

A potential proteolytic cleavage of DNA-protein-crosslinks was proposed to exist before this study as DPCs were removed from DNA without cutting DNA, leaving the protein as the only remaining part for repair¹⁷⁴. A proteolysis was demonstrated for the proteasome previously as well¹⁶⁷. This study identified for the first time a protein in metazoans able to cleave proteins crosslinked to DNA (Fig. 16 and 21). Even more, an unbiased screen for factors in DPC-cleavage identified SPRTN in addition to the proteasome as the prime DPC-cleaving agents¹⁸⁶. The requirement of SPRTN for DPC-repair was shown by an increase of the total level

of proteins on isolated DNA in SPRTN-deficient cells (up to 5-fold) or cell from RJALS-patients (up to 2-fold) with a higher DPC-accumulation upon known DPC-inducing agents such as Formaldehyde, Camptothecin or Etoposide¹⁸⁸. The findings of this study show an increase in Etoposide-induced Topo II-DPCs upon SPRTN-impairment (Fig. 19). The importance of SPRTN for response to Etoposide were further verified by our group in a SPRTN-KO model showing a higher Topo II-DPC abundance, which is corrected upon overexpression of SPRTN¹⁸⁷. Similar findings were observed upon Formaldehyde or Camptothecin-treatment^{188,191}. An increase in Topo I-DPCs was detected in SPRTN-KO cells and in livers of 4 months old SPRTN-deficient mice^{20,201}.

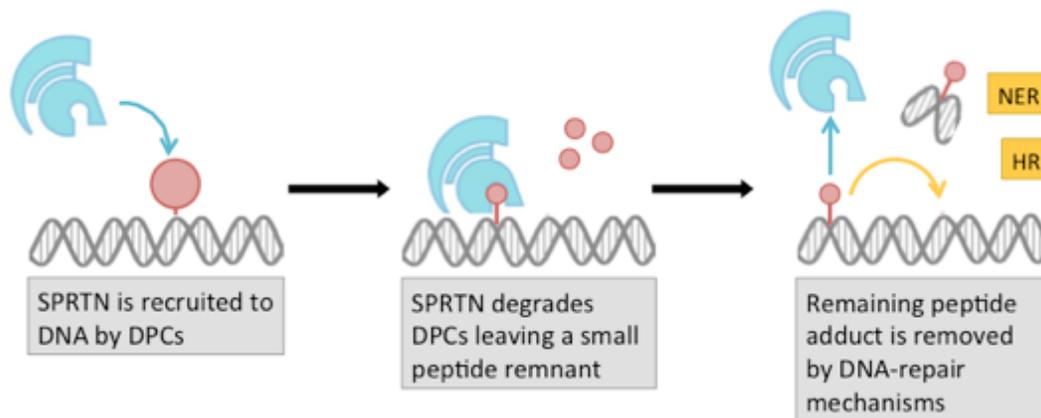


Figure 34 SPRTN's main function is the repair of DNA-protein crosslinks by proteolytic cleavage. Different endogenously or exogenously induced DPCs recruit SPRTN to DNA. SPRTN resolves these DPCs by proteolytic cleavage leaving a small peptide remnant on DNA, which is either further repaired by canonical DNA-damage repair mechanisms (e.g. NER, HR) or passed by translesion synthesis (TLS).

This study focused on Topo II and histone H3 as direct SPRTN substrates. The histone cleavage was later validated *in vitro* by independent studies^{183,188}. Further investigation revealed that SPRTN is able to cleave a variety of different proteins. Although a mass spectrometry analysis of SPRTN-depleted cells showed an increase predominantly in non-DNA binding proteins, only DNA-bound proteins could be cleaved by SPRTN *in vitro* as shown in several studies^{183,188,191}. However, proteins involved in DNA damage repair like PCNA (ubiquitylated or native), RFC or RPA are not cleaved by SPRTN¹⁹¹. A specific cleavage site in SPRTN substrates has not been identified but SPRTN primarily cleaves regions

enriched with serine, lysine or arginine and requires a distorted free region in accordance to its rather closed active site^{188,190}. Collectively the published data demonstrates a promiscuous activity in SPRTN - cleaving a variety of proteins crosslinked to DNA.

The relevance of SPRTN's uninterrupted function for DPC-repair is best illustrated by the impaired DPC-cleavage upon different mutations of SPRTN. The results of this study show an increase in DPCs in SPRTN-deficient cells and cells overexpressing SPRTN-mutant variants compared to ^{WT}SPRTN (Fig. 17, 20 and 29). This finding was further validated in a cell model devoid of endogenous SPRTN and reconstituted with different SPRTN-mutations¹⁸⁷. This analysis showed impaired DPC-cleavage in all conditions except where cells were reconstituted with ^{WT}SPRTN^{187,188,191}. Complementary findings were shown *in vitro* for histone- and Topo-cleavage, providing further evidence of the effect of SPRTN-mutations^{183,188}. ^{E112A}SPRTN does not elicit cleavage, ^{Y117C}SPRTN elicits a strongly (8%) and ^{ΔC}SPRTN only a partially reduced activity or even an activity similar to WT^{183,188,191}. Interestingly, on cellular level ^{ΔC}SPRTN shows a markedly reduced histone-cleavage as demonstrated in this study (Fig. 21). This difference can be explained by the mislocalization of ^{ΔC}SPRTN in cells, in accordance to a simultaneous intact activity on intermolecular level and an evidently deficient function in the patient cells (Fig. 25)^{183,187,188}.

5.3.3 SPRTN's proteolytic function is important for DNA-replication

SPRTN-deficiency was shown to have deleterious effects on cells like impaired cellular survival, genomic instability and severe replication deficits prior to this study^{9,13,16-19}. An increase in sensitivity in SPRTN-deficient cells to different DNA-damage-inducing agents such as Formaldehyde proposed a connection to SPRTN in the DNA damage response. In this study, a relevance of SPRTN for resistance towards Formaldehyde was displayed by an increased cell death upon treatment with Formaldehyde or other DPC-inducing agents like Etoposide or Camptothecin found in SPRTN-depleted cells (Fig. 22 and 23). As further evidence that this sensitivity is a result of an impaired DPC-repair, another study showed that the

high Formaldehyde-sensitivity in SPRTN-deficient cells is suppressed by concomitant removal of Topo I, thereby preventing Topo I-DPC-formation¹⁸⁸. MEFs hypomorphic for SPRTN showed a pronounced Camptothecin-sensitivity, and different SPRTN-mutations were shown to increase UV-sensitivity, corresponding to the findings of this study²⁰¹. The observed sensitivity required an explanation of the underlying mechanism. With this study we hypothesized that SPRTN's proteolytic function is necessary for providing unperturbed DNA-replication by cleaving DPCs and thereby preventing replication fork collision and stalling.

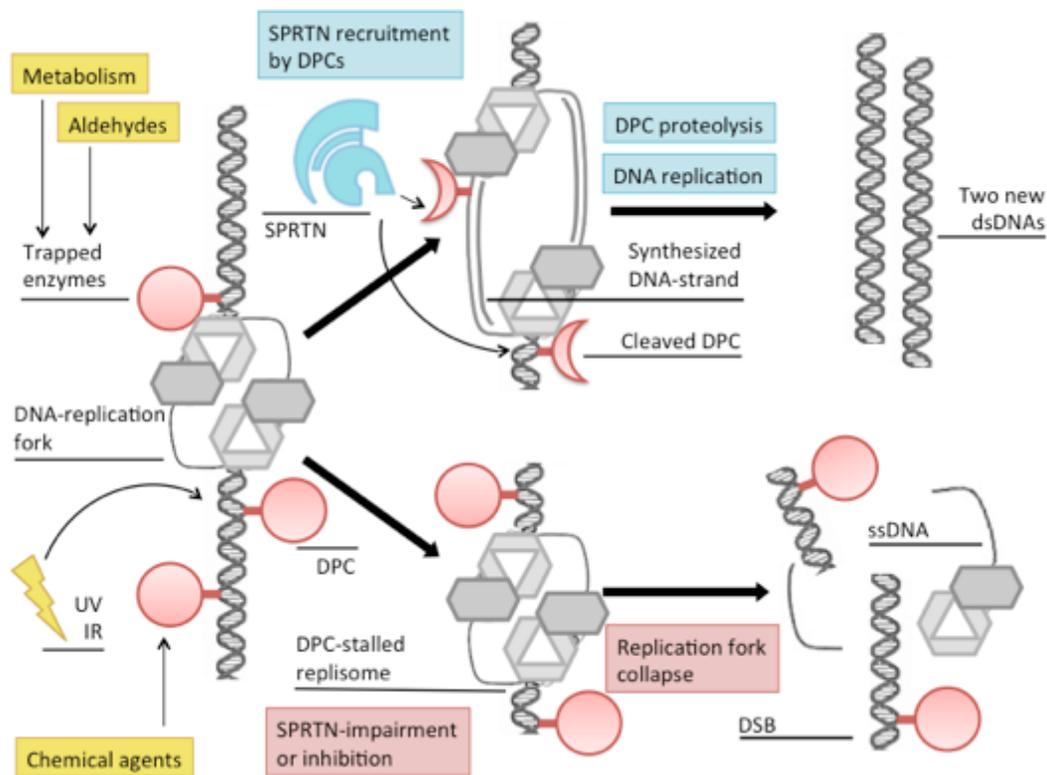


Figure 35 SPRTN's function in resolving DNA-protein-crosslinks is important for DNA replication and genomic integrity. Different endogenous (metabolites, aldehydes, enzymatic intermediates) or exogenous sources (chemotherapeutic agents, ionizing radiation) can cause DPCs, which block progression of the replication fork. **(Upper part)** DPCs induce SPRTN recruitment, which further resolves DPCs by proteolysis, enabling replication to continue and sustaining genomic integrity for cell proliferation. **(Lower part)** In case of SPRTN-depletion, mutation or inhibition, DPC proteolysis is impaired resulting in a stalling and collapsing of the replication fork with deleterious DNA-damage ultimately leading to cell death or tumorigenesis.

The relation of SPRTN to DNA-replication is implied by the finding that SPRTN's protein levels peak during S phase and that SPRTN-deficiency increases Formaldehyde-sensitivity only in proliferating cells, not in G0-arrested cells^{17,188}. Accordingly, we found that SPRTN is degraded when replication is inhibited by L-mimosine. Furthermore, BrdU and DNA fiber assays exhibited slower DNA-replication dynamics in SPRTN depleted cells^{188,191}. Interestingly the role of SPRTN during DNA replication was shown to be more important under Formaldehyde or UV-treatment¹⁹¹. This decrease in DNA replication could be a result of an impaired DPC-proteolysis, as DPC accumulation leads to replication fork stalling^{174,106-109}. Later studies revealed that SPRTN is indeed accumulating on replicating DNA and even travelling with the replication machinery as shown by iPOND analysis^{188,201,202}. Moreover SPRTN cleaves DPCs during replication, similarly to the proteasome¹⁸⁶. However, SPRTN is known to play an important role in translesion synthesis and subsequently to the negative effect on SPRTN-impairment and DNA-replication could potentially result from an impaired TLS^{14,16-18}. This is an unlikely explanation, as the proteolytic SprT-domain alone is able to rescue replication defects in SPRTN-deficient cells, but SPRTN-mutants missing either the UBZ, PIP or SHP-domain not^{20,191}. Additionally the SprT-domain alone travels with the replication fork equally as the full-length protein²⁰¹. These findings prove that SPRTN ensures efficient DNA-replication by its proteolytic function and independently of its TLS function. This is in accordance to our results showing that only mutations in SPRTN's proteolytic core lead to impaired cell proliferation upon exposure to DPC-inducing agents, while a missing of C-terminal interacting motifs does not affect DPC-cleavage (if the nuclear localization is provided) (Fig. 27). An analysis of cell proliferation in MEFs completely lacking SPRTN demonstrated an impaired TLS and an accumulation in S phase, whereby an addition of the proteolytic SprT-domain relieves the cells from S phase accumulation, but not from the TLS-deficit²⁰¹. Impaired DNA replication is known to produce unstable replication intermediates prone to DSBs and increase in genetic instability. SPRTN-deficient cells exhibit the same characteristic chromosomal deficits with 4C DNA content, measured by flow cytometry of SPRTN KO MEFs and increased

γ H2AX, as a marker of DSBs in mice cells with reduced SPRTN-level²⁰. Furthermore, a higher level of 53BP1, also a marker of DSBs, was observed in SPRTN-deficient cells upon DPC-induction (by Formaldehyde or Camptothecin), but only during S phase, thereby showing that the DSBs are produced as replication intermediates during an impaired DNA-synthesis¹⁸⁸.

In conclusion, all the findings discussed above collectively support that SPRTN, as a DPC repair protease, prevents replication fork stalling upon collision with DPCs and thereby avoids replication defects, DSBs and genomic instability independently of its TLS-related function. This explains the observed slower replication as well as the chromosome defects in cells with a deficient or mutated SPRTN-proteolysis.

5.3.4 SPRTN has an important role in cell cycle regulation

A connection of SPRTN and cell cycle regulation is implied by SPRTN's important role in DNA-replication^{9,188,201}. Interestingly SPRTN was shown to have a higher impact on cell cycle progression, beyond prolonged DNA-replication, as SPRTN-deficiency leads additionally to a lack in G2/M checkpoint activation upon replication stress with resulting chromosomal instability^{9,183}. Cells with impaired SPRTN activity actually exhibit faster initial S phase progression, although with more chromosomal aberrations, and mostly accumulate in late S or G2 phase^{20,201}. A similar effect of slower replication and fork stalling was found upon Checkpoint kinase 1(CHK1)-deficiency, with no additive effect during simultaneous SPRTN-deficiency implying a common pathway for cell cycle regulation²⁰². Indeed SPRTN acts on CHK1 as on a DPC-substrate, releases it from DNA and additionally cleaves its C-terminal part, thereby activating it and subsequently enabling a cell cycle arrest. This explains the observed G2/M-checkpoint deficit in SPRTN-deficient cells and proposes a model, in which upon replication stress SPRTN acts as a CHK1 activator thereby delaying mitotic entry, stabilizing replication forks and preventing chromosomal aberrations.

5.4 Review of regulation of SPRTN activity

The proteolytic function in SPRTN is on one side important for DNA damage repair, as it cleaves DPCs and thereby enables normal DNA replication and cell cycle progression¹⁰⁶⁻¹⁰⁹. Yet on the other side, the ability to cleave proteins bound to DNA simultaneously presents a potential danger, since many essential functions are performed by DNA-binding proteins and must not be interrupted. Therefore SPRTN's promiscuous proteolysis must be tightly regulated.

5.4.1 SPRTN is activated upon DNA-binding

Since SPRTN is a protease cleaving specifically proteins crosslinked to DNA, the results of this study demonstrate a role for DNA in SPRTN's function. As already discussed, SPRTN binds to DNA via its DNA-binding basic region (BR), and additionally via its ZBD-region of its protease domain^{189,190}. The type of DNA, which SPRTN binds to, differs among different experimental data, whereby SPRTN shows binding to ssDNA and dsDNA *in vitro* while other *in vitro* experiments and supershift assays showed a clear preference to ssDNA^{189,190}. The difference can be explained by a possible spontaneous formation of ssDNA parts at the end of dsDNA. Subsequently a causal relation between DNA-binding and SPRTN activity was supported by our group and others, showing that *in vitro* autocleavage is only possible upon addition of DNA^{183,188,190,191}. Similarly, a requirement of DNA was confirmed for substrate cleavage as well, whereby autocleavage was detected upon ssDNA and dsDNA while substrate cleavage was primarily or exclusively activated by ssDNA, but in other settings by dsDNA as well^{183,188,190,202}. The importance of DNA-binding for SPRTN's function is undoubtedly displayed by the impairment of substrate cleavage upon decreased DNA-binding. Accordingly, a mutation of SPRTN's BR-domain, which highly decreases DNA-binding, leads to an impaired repair of UV-induced lesions *in vivo* and a diminished cleavage activity *in vitro*¹⁸⁹⁻¹⁹¹. Furthermore, a mutation of the DNA-binding ZBD-domain reduces histone-cleavage markedly¹⁹⁰. In concordance, the addition of proteins known to have a high ssDNA-affinity as RPA or RFC suppresses SPRTN's cleavage activity, as these proteins cover ssDNA making it

inaccessible for SPRTN. Despite little differences, the existing data collectively confirms SPRTN-activity as DNA-dependent and primarily being induced by ssDNA.

5.4.2 SPRTN's function depends on DNA-synthesis

The fact that SPRTN's main function is DPC-cleavage and that DPCs primarily represent a problem during DNA-replication implicate the intriguing concept that SPRTN's function is dependent on DNA-replication. As a replication-blockade causes a persistence of DPCs and the results of this study show a degradation of SPRTN upon replication-blockage, the idea of a replication-dependent SPRTN-activity becomes even more reasonable (Fig. 24). Indeed, later studies showed that SPRTN co-immunoprecipitates major DNA-replication proteins as PCNA or Pol δ . In addition, an iPOND assay revealed a presence of SPRTN on nascent DNA and further supports a movement of it together with the replisome^{188,201}. Further analysis uncovered that SPRTN accumulates on DNA upon replication and that it cleaves DPCs during replication¹⁸⁶. However, the presence of the replisome is not strictly required for SPRTN's DPC-cleavage as SPRTN performs its function upon inhibition of replication likewise, as long as ssDNA is present, consistent with the *in vitro* data showing substrate cleavage upon ssDNA^{183,190,191}. On a cellular level, DNA-synthesis detected on ssDNA was shown to be the determining factor required for DPC-cleavage, as solely ssDNA with an inhibition of DNA-synthesis cannot trigger DPC-repair *in vivo*¹⁸⁶. More precisely, SPRTN's DPC-cleavage cannot be activated by the helicase colliding with the DPC but only by the polymerases synthesizing DNA during DNA-replication¹⁸⁶. This data is consistent with the finding that polymerases can approach up to the DPC-binding DNA-base and are blocked after insertion of a paired nucleotide at this position¹⁸⁶. The proof of DPC-cleavage induction by DNA-polymerases together with our and further findings that SPRTN-level and recruitment are replication-reliant demonstrates that SPRTN's DPC-cleavage is dependent on DNA-synthesis (distinct to the Proteasome) and reserved to the context of DNA-replication (Fig. 24)¹⁸⁶. However, SPRTN is not a part of the replication machinery as proposed,

since it is recruited to the replication fork separately than the replisome, during its resolution^{188,189}.

5.4.3 The influence of Interacting proteins on SPRTN function

The several different domains in SPRTN's amino acid sequence for interaction with other proteins prompt for a possible regulation of SPRTN's proteolytic function by some of these interacting proteins (Fig. 32). The most promising candidate is ubiquitin. As already mentioned SPRTN encompasses a ubiquitin-binding-zinc-finger (UBZ). A ubiquitylation of SPRTN has been demonstrated by the endogenous presence of a monoubiquitylated form of SPRTN (Ub-SPRTN)¹⁹¹. A site analysis revealed that distinct lysine residues in SPRTN can be ubiquitylated with no preference among them¹⁸³. Furthermore the monoubiquitylation of SPRTN depends on its UBZ-domain, as a mutation of this domain diminishes Ub-SPRTN^{183,191}. As a UBZ-dependent monoubiquitylation of a protein was shown for other proteins like Polymerase η , thereby affecting the activity of the protein, a similar regulation could be present in the case of SPRTN^{187,203,204}. The fact that Ub-SPRTN does not bind to DNA *in vitro* and that Formaldehyde-treatment induces de-ubiquitylation of SPRTN and simultaneously recruits SPRTN to DNA supports this idea¹⁸³. However, SPRTN recruitment to DNA is observed by other conditions like UV, which has no de-ubiquitylating effect, although this process depends on the UBZ and PIP-domains, as mutations of either of these motifs impair recruitment^{12-14,17,183}. Additionally our group showed an increased cleavage activity of SPRTN *in vitro* upon addition of ubiquitin, which is in contrast with the theory of a promoting effect of a SPRTN de-ubiquitylation on its activity¹⁸⁷. These findings make a regulation of SPRTN by ubiquitylated substrates or interactors more feasible than by ubiquitylation of SPRTN itself.

DPCs can be ubiquitylated similarly as other proteins and their ubiquitylation was shown to recruit the proteasome for subsequent cleavage¹⁸⁶. The DPC-repair by the proteasome was demonstrated to be dependent on this process. In comparison, an inhibition of DPC-ubiquitylation does not completely abolish SPRTN-activity, as in the case with the proteasome, but still reduces it.

Subsequently, non-ubiquitylated DPCs are cleaved by SPRTN, yet only if the UBZ-domain remains intact¹⁸⁶. This finding indicates that SPRTN's cleavage function does not necessary require an ubiquitylation of the DPC-substrate itself, but rather depends on an ubiquitylated interacting protein. Importantly, this interaction is promoting SPRTN's function but it is not indispensable for it, as DPC-cleavage is found in Δ^C SPR TN, lacking the UBZ-domain in our study (only upon proper nuclear localization demonstrated by Δ^C+NLS SPR TN) and *in vitro* by other studies as well (Fig. 27)^{183,188,191}. Additionally, the UBZ-domain was shown to be dispensable for progression through S phase²⁰. The influence of ubiquitin on SPR TN activity is illustrated by the fact that a reduction of the ubiquitin-level causes a decrease of DPC-proteolysis in cell extracts¹⁷⁴. In parallel *in vitro* data from our group showed that the addition of ubiquitin to SPR TN increases self- and histone-cleavage¹⁸⁷. The most probable underlying mechanism is an *in trans* binding to a ubiquitylated interacting protein either recruiting SPR TN to the DPC or activating its proteolytic activity.

A search for the potential ubiquitylated interactor of SPR TN proposes PCNA as a probable candidate, primarily because SPR TN contains a PCNA-interacting motif and shows a relevant PCNA-binding *in vitro* and *in vivo*¹⁶⁻¹⁸. Additionally, mono-ubiquitylation of PCNA is known to play a pivotal role in DNA-replication stress response^{15,103}. Correspondingly, the importance of SPR TN's interaction with PCNA and Ubiquitin was demonstrated by the finding that SPR TN with a mutated PIP or UBZ-domain shows an impaired recruitment to UV-induced DNA-lesions and a defective lesion bypass^{12-15,17,20,183}. Additionally, Rad18, a known PCNA-ubiquitylating enzyme, was shown to act with SPR TN in the same pathway in restoring replication fork as a simultaneous mutation of both proteins shows no additive defect in fork replication progression or sensitivity to Formaldehyde^{12,14,191}. However, only a UBZ-mutated SPR TN-variant exhibits an impaired DPC-cleavage, not a PIP-mutated¹⁷⁴. A closer analysis of the distinct findings shows that PCNA-binding is important for the recruitment of SPR TN and replication upon DNA lesions, functions previously already attributed to SPR TN's role in translesion synthesis (TLS), while a clear connection of PCNA-binding and

DPC-cleavage is not demonstrated until now. A positive effect of PCNA-binding on SPRTN's cleavage function seems feasible via an advance of SPRTN recruitment to DPCs or stalled replication forks, which still needs to be elaborated.

Collectively the existing data about interactions of SPRTN with other proteins shows an obvious promoting influence of ubiquitin-binding on DPC-cleavage, although not strictly required, which is most probably achieved via a binding of SPRTN to a yet unknown ubiquitylated interacting protein instead of an ubiquitylation of the substrate or SPRTN itself. The PCNA-binding of SPRTN seems less important for its DPC-cleavage activity and is primarily relevant for SPRTN's function in lesion bypass.

5.4.4 SPRTN dimerization

Additionally to the interaction to other proteins, the interaction of SPRTN with another SPRTN molecule also presents a possible regulation mechanism for its function. The assumption of a SPRTN-SPRTN interaction results from our finding that SPRTN shows a self-cleavage function *in trans* as shown in this study, proving an interaction of this protein with itself (Fig. 11). This study additionally demonstrated that two SPRTN-molecules dimerize (directly or indirectly) *in vivo* (Fig. 28). Interestingly, a structural analysis revealed that DNA binds two SPRTN molecules and that these two molecules interact with each other leaving the possibility that SPRTN also functions primarily as a dimer^{189,190}. Of even more importance, the sequence of SPRTN forming the interface between two SPRTN molecules shows a high evolutionary conservation as an indication for its relevance. The importance of SPRTN-dimerization for its activity was further shown by the fact that a mutation in the dimer-interface region of SPRTN leads to a decreased DPC-cleavage activity¹⁹⁰. However, while our findings support a co-immunoprecipitation between two SPRTN molecules depending on its C-terminal half, a structural analysis of DNA-bound SPRTN demonstrated a dimerization in the N-terminal SprT-domain. This difference could be explained by a potential dimerization at the N-terminus after DNA-binding of SPRTN, whereby a C-terminal interaction can happen independent of DNA, as in our experiment one SPRTN

molecule was lacking the DNA-binding N-terminal half. The fact that the *in trans* autocleavage of SPRTN happens primarily at the C-terminus further supports the hypothesis of two distinct mechanisms of SPRTN-SPRTN-interaction, especially since the N-terminally bound SPRTN-dimer is not capable of autocleavage *in trans* as the binding interface is right next to the catalytic core thereby sterically hindering the C-terminal part of one SPRTN molecule to come close enough to the protease core in order to be cleaved. Further investigation is needed to elaborate the proposed effect of SPRTN-dimerization for SPRTN's proteolytic activity.

5.4.5 The impact of autocleavage on SPRTN activity

As SPRTN-autocleavage was demonstrated to appear endogenously *in vivo* by this study (Fig. 12) and *in vitro* by other studies, an effect of the autocleavage on SPRTN-activity is a reasonable proposal^{183,187,188}. The autocleavage could thereby promote DPC-cleavage, if the cleaved SPRTN is its active form, or inhibit DPC-cleavage. A positive relation between autocleavage and substrate cleavage is suggested by a demonstrated increased autocleavage of SPRTN upon DPC-induction by Etoposide in our study (Fig. 21), or by Formaldehyde in another¹⁸³. The higher affinity of SPRTN to itself than to substrates further implies that the autocleavage may precede the substrate cleavage thereby promoting or even enabling it¹⁸³. However, an analysis of the cleavage kinetics *in vitro* revealed that SPRTN is cleaving substrates (e.g. histones) faster than itself, which rejects this hypothesis¹⁸⁸. The higher binding- affinity to itself than to substrates could instead maybe cause a dimerization of SPRTN as a first step, followed by DPC-cleavage by a functionally active SPRTN dimer, although this theory needs to be elaborated further.

On the other side, a potential negative effect of SPRTN's autocleavage on DPC-cleavage is indicated by the finding that autocleaved fragments like ¹⁻²¹⁸SPRTN lose their proteolytic activity^{188,190}. Interestingly this study showed an intact *in vivo* H3-cleavage of the truncated ^{ΔC}SPRTN (resembling an autocleaved fragment) if a nuclear localization is enabled (Fig. 27). Accordingly, ^{ΔC}SPRTN showed an unaltered H3-cleavage and autocleavage *in vitro* as well¹⁸⁸. However, an impaired

substrate cleavage (e.g. of Topo I and Topo II and histones) was found by other studies^{183,188}. In any case, a reduced substrate cleavage of shorter SPRTN-variants is demonstrated, appearing either at the level of the Δ^C SPRTN or at shorter fragments, and subsequently provides evidence for a negative effect of autocleavage on SPRTN function. The simultaneous retained autocleavage ability of cleaved SPRTN fragments further supports a model where autocleavage is performed after substrate cleavage thereby ending a prolonged SPRTN-mediated proteolysis^{183,188}. Accordingly, although autocleavage is activated by DPC-induction, it displays a different activity than substrate-cleavage. As DPC levels increase, SPRTN's substrate-cleavage is primarily activated, while a following decrease in DPCs activates autocleavage equally to substrate-cleavage¹⁸³. Furthermore, as it was shown that inactive SPRTN-variants are more recruited to DNA, like in this study and stay longer recruited, the autocleavage seems to play a role in SPRTN-release from DNA once activated (Fig. 15)¹⁸³. In conclusion, SPRTN's autocleavage activity serves most probably as a negative regulation mechanism aimed to end SPRTN's proteolytic activity after activation (as it happens simultaneously but shows slower kinetics) and thereby preventing unwanted proteolysis of important proteins by SPRTN.

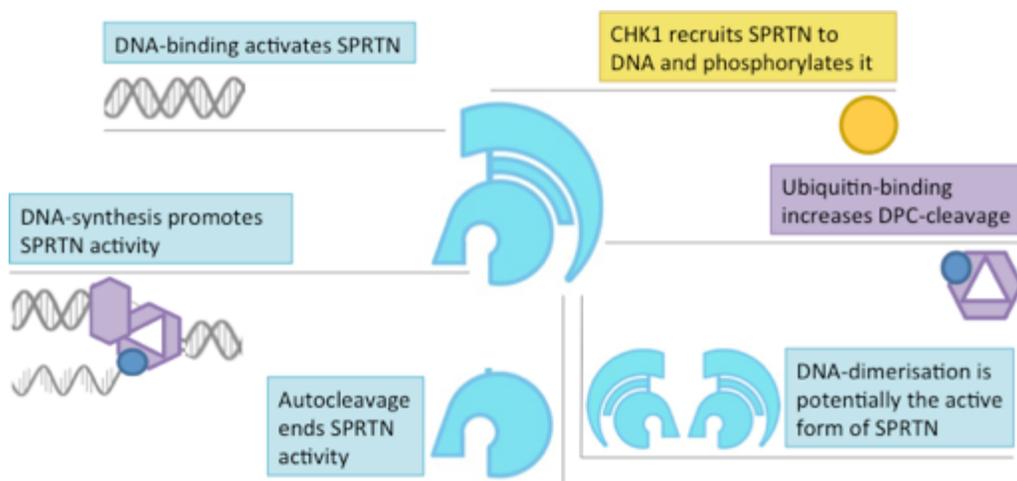


Figure 36 Regulation of SPRTN-activity. SPRTN-mediated DPC-cleavage shows a complex regulation with various involved factors. Distinct regulation mechanisms have been proposed with different supporting evidence. This figure serves as a short overview of factors with demonstrated positive or negative effects on SPRTN function. Further details can be found in the main text.

5.4.6 The cell cycle-dependent regulation of SPRTN

A connection of SPRTN to the cell cycle was demonstrated in several studies as already discussed^{17,202}. A cell cycle dependent SPRTN-level is supported and seems reasonable as SPRTN's activity is DNA-coupled and replication-dependent^{17,183,188,90}. An interesting question is how SPRTN-activity is regulated by the cell cycle. This study showed that SPRTN is degraded upon blockage of replication (Fig. 24). Previous work supports the notion that SPRTN is degraded by the APC/Cdh1-complex¹⁷. This study could further confirm this finding by the fact that an inhibition of APC/Cdh1-dependent proteasomal degradation stabilizes SPRTN even upon replication blockage (data not shown). Additionally, this study shows that the degradation of SPRTN depends on the C-terminus as Δ^C SPRTN shows an impaired degradation. These findings propose a regulation of SPRTN-level via the cell cycle dependent APC/Cdh1 complex as an underlying mechanism of SPRTN's cell cycle dependence.

Another interesting mechanism of cell cycle dependent SPRTN regulation is proposed by the finding that CHK1 is acting on SPRTN²⁰². It was shown that CHK1-overexpression restores fork progression and chromatin stability in SPRTN-deficient cells. An increased SPRTN-recruitment to chromatin upon CHK1-overexpression and a reduced DPC-level upon CHK1-overexpression in SPRTN-depleted cells was shown proposing that CHK1 promotes the activity of the remaining SPRTN. Furthermore, a phosphorylation of SPRTN by CHK1 was described. Thereby, phosphorylation sites in SPRTN were found at S373, S374 and S383 as typical target motifs for CHK1 with a strong reduction in SPRTN-phosphorylation upon their mutation. A regulation of SPRTN-activity by CHK1-mediated phosphorylation therefore seems promising.

Taken together, the existing data demonstrates a clear cell cycle dependence of SPRTN-level and activity with distinct regulation mechanisms being involved. Further investigation will be needed to elucidate the exact regulation.

5.5 Explanation of the clinical findings in RJALS-syndrome

The observed phenotypes in the three patients harboring mutations in the SPRTN gene were collectively described as RJALS-syndrome. The main characteristics, progeroid features and tumorigenesis, are commonly results of an underlying impaired DNA-replication or DNA damage repair²⁰⁵. Accordingly, severe replication defects were observed in SPRTN-deficient cells. These cells demonstrate slower replication speed with a higher rate of stalled replication forks and newly fired origins in fiber assay⁹. As a consequence of the defective DNA replication, genomic instability was shown in different models of SPRTN-deficiency^{9,20,188}. Replication intermediates and unrepaired breaks in the form of chromatin bridges and micronuclei were observed in SPRTN KO MEFs²⁰. Patient fibroblast exhibit variegated translocation mosaicism⁹. Additionally to structural abnormalities numerical chromosomal instability was detected as increased rates of aneuploidy in proliferative MEFs, lung and prostate mouse cells with reduced SPRTN activity and further confirmed in liver biopsies from patients as increased levels of DNA damage markers, primarily for DSBs^{9,20,201}.

However, the cause of the observed genomic instability upon SPRTN mutation remained unclear up to this study. With our work we were able to uncover the essential function of SPRTN, which is impaired in RJALS-patients and thereby explained the underlying mechanism of the observed phenotypes upon SPRTN mutation. As already discussed, SPRTN's TLS-related function was completely abolished in one patient, who was still alive, while the DPC-function was impaired in all three patients, either due to a homozygous mislocalization of Δ^C SPRTN, or due to a heterozygous partial mislocalization of one half and a simultaneous impaired proteolysis of the other half of SPRTN proteins in Y^{117C} SPRTN/ Δ^C SPRTN-cells. We provided evidence of an impaired DPC-cleavage in cells expressing mutated SPRTN-variants corresponding to the clinical mutations (Fig. 17 and 21). The impaired DPC-cleavage was further demonstrated in cells directly derived from RJALS-patients by our group, and by other groups with an additional increased DPC-sensitivity^{187,188}.

The defective DPC-repair and subsequent replication impairment lead to severe proliferation and survival impairments as shown by reduced proliferation rates in hypomorphic and KO MEFs and additional increases in apoptosis²⁰. Similar observations are shown in patient cells as well⁹. The severity of these impairments on a cellular level was confirmed by their consequences on an organismal level, namely in a zebrafish model and a mouse model with graded reduction in SPRTN expression²⁰. Mice with hypomorphic SPRTN function display phenotypes resembling the clinical features observed in patients with RJALS-syndrome, namely dwarfism, lordokyphosis, cataracts and cachexia, features typically found in progeroid syndromes^{9,10,20}. Similar findings were demonstrated in this study as well (Fig. 30). While hypomorphic (H/H) mice with less than 10% of SPRTN activity are born in an expected normal frequency, but exhibit dwarfism, H/- mice (heterozygous with one hypomorphic and another KO allele) as well as complete KO mice are not viable. A complete KO causes embryonic lethality in mice with no viable offsprings and embryos die before the implantation stage²⁰. This study demonstrated for the first time an increase in DPC-accumulation upon SPRTN-impairment on an organismal level (Fig. 29). Thereby a mouse model with a catalytic SPRTN-mutation exhibited a higher level of Topo II-DPCs, endogenously and even more pronounced upon Etoposide-treatment. This is in accordance to the hypothesized impaired DPC-cleavage as the underlying pathogenesis of the RJALS-syndrome. Similarly, a mouse model with reduced ^{WT}SPRTN-activity showed accumulated Topo I-DPCs²⁰¹.

Another important phenotype observed in RJALS-patients is tumorigenesis, primarily in the liver, which was also shown to be in common between mice and humans with deficient SPRTN-function^{9,201}. An analysis of 22-25 month-old SPRTN-deficient mice with a H/H genotype screened for neoplasia revealed a clear increased spontaneous tumor formation in 70%, as compared to control mice with only 25% tumor prevalence²⁰¹. The finding that the liver is the prime organ susceptible for tumorigenesis supports the hypothesized DPC cleavage as the main underlying mechanism in RJALS since the liver, as a metabolically highly active organ, produces high levels of potential DPC-inducing toxins. Furthermore,

due to its physiologic role as the prime detoxifying organ in vertebrates, the liver is constantly exposed to various toxins with possible DPC forming activity. For instance, Formaldehyde-accumulation in the liver was already demonstrated²⁰⁶.

The explained analysis of the existing data regarding RJALS from clinical observations, from mouse models as an organismal level and from cells derived from patients with the clinical mutations united give sufficient evidence for the hypothesized impaired DPC-cleavage as the underlying mechanism of the disease. SPRTN-malfunction thereby hinders DPC-cleavage and leads to impaired DNA-replication with subsequent genomic instability resulting in premature aging phenotypes and tumorigenesis.

5.6 Future perspectives

5.6.1 SPRTN as a DNA damage response regulator

The existing data demonstrates an involvement of SPRTN in several different processes related to DNA damage response. SPRTN shows an essential function in DPC-cleavage as demonstrated by this study; it is demonstrated to play an important role in TLS; and in parallel it is shown to be connected to the checkpoint regulation of cell cycle (Fig. 19 and 22)^{12-14,16-18,202}. These findings of SPRTN functioning in distinct mechanisms of DNA damage repair and control, together with the fact that SPRTN was proposed to be an essential protein by several methods prompts a hypothesis that SPRTN could act as a central actor in DNA damage response, for instance as a connecting point of different mechanisms, as a regulation/control point in pathway choice or as a factor in various DNA damage repair pathways. The fact that SPRTN was shown to act in the same pathway with TDP1, RAD18, CHK1 as well as the multiple interacting proteins of SPRTN connected to DNA like ubiquitylated PCNA, p97, Pol η , REV1 (DNA repair protein REV1) and REV7 (DNA polymerase zeta processivity subunit) further support the concept of SPRTN as a main actor in DNA damage response^{12,14,188,191,202}. All this data implies an important and complex role of SPRTN and pronounces the importance of further investigations of this protein.

5.6.2 The role of SPRTN in tumorigenesis and cancer therapy

The relevance and toxicity of DPCs was already explained in detail. It is noteworthy that the frequency of DPCs in native tissue correlates with proliferation rates and that DPC-toxicity is exaggerated during DNA-replication^{188,207,208}. For instance, DNA repair enzymes like MGMT naturally form transient DPC-intermediates and can potentially be crosslinked to DNA permanently similar to other proteins^{209,210,211,212}. Therefore it is not surprising that in tumors the level of SPRTN is elevated in comparison to non-proliferating normal tissue (Fig. 37)²¹³. These observations propose an important function of SPRTN in tumorigenesis on one side as well as in tumor treatment on the other side. Since SPRTN shows a protective function for cells in response to DPCs and DPCs are more present and

more toxic in proliferative cells like tumor cells, SPRTN-inhibition should sensitize tumor cells to DPCs subsequently enhancing tumor cell death. Therefore SPRTN-inhibitors represent a completely new potential chemotherapeutic approach.

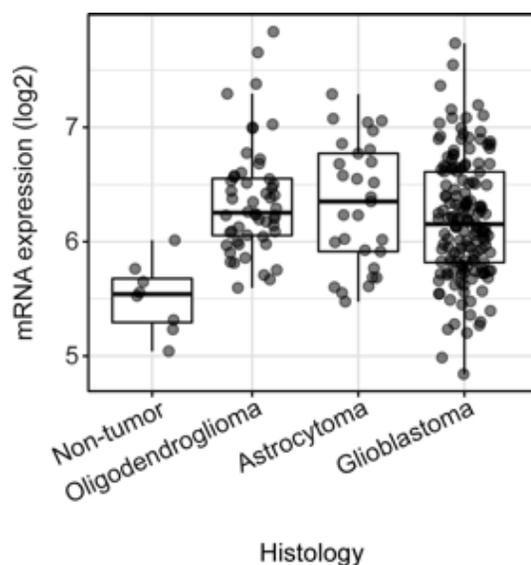


Figure 37 SPRTN gene expression across glioma histologies. Expression of SPRTN was generated using the Gravendeel dataset and the GlioVis algorithm.

The fact that many different commonly used chemotherapeutic agents induce DPCs makes SPRTN-inhibition even more interesting as a potential chemotherapeutic agent. The DPC-inducing agents employed in this study, like Etoposide, Camptothecin or Olaparib, are all established chemotherapeutic drugs²¹⁴. PARP-inhibitors like Olaparib were even shown to be primarily toxic in chemotherapy due to their DPC-inducing function, more than due to their catalytic inhibition of PARP1¹²¹. The effect of DPC-induction for other chemotherapeutics in comparison to their known inhibiting function would need to be investigated. According to our hypothesis of SPRTN as an important regulator of chemotherapeutic sensitivity, we already showed that SPRTN-deficiency sensitizes cells to different chemotherapeutic drugs (e.g. Etoposide) (Fig. 22 and 23). These results implicate a potential synergistic effect of SPRTN-inhibition in addition to DPC-causing agents in tumor therapy.

A further important connection to SPRTN with tumor therapy is the finding that ionic radiation (IR) used in radiotherapy induces DPCs in addition to double strand breaks (DSB), in particular under hypoxic conditions¹⁴². Hypoxic areas with necrosis are characteristic features of tumors and hypoxia per se has been implicated to reduce sensibility to cancer-treatments including radio- and chemotherapy²¹⁵. Generally, DPCs are less toxic than DSBs, therefore to exploit the full potential of radiotherapy an additional DPC-repair inhibition (e.g. by SPRTN inhibition) could be a synergistic co-treatment approach to target therapy resistant tumor regions³⁸. There is convincing data suggesting an important role of DPC resolving mechanisms and thereby of SPRTN in tumor progression and therapy resistance. Thus far however, this mechanism has not been exploited for therapeutic purposes. Analyzing the existing data a model where SPRTN mediates tumor resistance against chemo- and radiotherapy seems feasible. Subsequently, SPRTN-inhibition presents an interesting new strategy to improve tumor treatment.

In summary, SPRTN as the first mammalian protease shown to cleave DPCs represents a completely new and interesting DNA-repair mechanism with many potential implications in cellular physiology and pathology, especially in tumorigenesis. Therefore, this study together with the according SPRTN investigations opens a new and promising approach in the study of DNA damage response, tumorigenesis and tumor therapy.

6 Materials and Methods

Experimental settings, reagents and antibodies

Experiments were performed following standard procedures or new established variants of them as described. The utensils, wells and flasks used in the experiments were sterilized. Sterile deionized water was used in the experiments. Experimental materials, reagents and chemicals were purchased from Agilent, Applichem Calbiochem, Clontech, Enzo Life Sciences, Biolabs, Biorad, Invitrogen, LC Labs, Merck, Milipore, Promega, Roche, Roth, Sigma, Thermo Fisher Scientific. Restriction enzymes were ordered from Biolabs, NEB and Thermo Scientific. Antibodies used in this study are listed as follows: primary antibodies included anti-GFP (Santa Cruz, sc-9996), anti-mCherry (Eurogen, AB233), anti-Histone H3 (Abcam, ab1791), anti-p53 (Santa Cruz, sc-126), anti-PCNA (Santa Cruz, sc-56), anti-PARP1 (Cell Signaling, 9542), anti-SPRTN (Sigma, HPA025073), anti-TOM20 (Abcam, ab56783), anti-Topoisomerase I (Biomol, A302-590A), anti-Topoisomerase II alpha (Abcam, ab52934), anti-Vinculin (Sigma, V4505), secondary HRP-coupled goat antibody were anti-mouse (Bio-Rad, 170-6516) and anti-rabbit (Dako, P0448).

6.1 Cellular biology techniques

Buffers and solutions

Growth medium

- 450ml DMEM
- 50ml FBS to a final concentration of 10%
- 1mM Pyruvate

Mincing solution

- 14.85ml PBS
- 150µl Pen/Strep to a final concentration of 1%

Digesting medium

- 8.65ml DMEM
- 250µl HEPES (2.5%)
- 150µl Pen/Strep to a final concentration of 1%
- Collagenase II shortly before usage to a final concentration of 400U/ml

Isolating solution

- 450ml DMEM
- 50ml FBS to a final concentration of 10%
- 1mM Pyruvate
- 5ml Pen/Strep to a final concentration of 1%

Crystal violet

- 0.5g crystal violet powder
- 20ml Methanol
- 80ml H₂O
- 6% Glutaraldehyde

Cell lines

Cells lines (HEK293T, U2OS) were purchased from the American Type Culture Collectio (Manassas, VA) with identities verified by STR analysis.

Mouse embryonic fibroblasts (MEFs) with the conditional SPRTN-knock-out system were provided by Yuichi Machida (Mayo Clinic, Rochester Mn.).

^{Y118C/Y118C}SPRTN-primary myoblasts were isolated from skeletal muscle of adult ^{Y118C/Y118C}mice ordered for Dikic Laboratory as described. Cell lines used in the experiments were proven to be mycoplasma-free by PCR (Minerva Biolabs, Germany).

Culturing of cells

Cells were grown in Dulbecco's modified Eagles's minimal essential medium (DMEM) (Invitrogen, 41965-039) enriched with 10% fetal bovine serum (FBS) (Thermo Scientific, 10207). Cells were incubated at 37°C in a humidified atmosphere containing 5%CO₂ in 10cm dishes. Cell expansion was allowed up to approximately 80% confluence and cells were splitted by washing with PBS and detachment using 2.5ml Trypsin-EDTA-solution (Invitrogen, 25300-054) for 3min, which was blocked by addition of 2.5ml DMEM/10%FBS . Suspensions were collected in a Falcon tube and gently centrifuged followed by supernatant removal and resuspension of cell pellet in 10ml DMEM/10%FBS. Suspensions were diluted 1:20 in new 10cm dishes for further maintenance or in 6-well or 96-well plates for experiments.

Determination of cell number

Cell concentration was measured using the TC20 Automated Cell Counter (BioRad Laboratories). Samples were prepared by combining equal amounts (10µl) of cells suspended in PBS (Gibco, 14190-094) and T10 trypan blue dye and mixed by gentle pipetting ten times. 10µl of the mixture was added on a TC10 counting slide and analyzed by the cell counter. Samples were counted in duplicates and only the number of viable cells was calculated for further experimental steps.

Isolation of primary myoblasts

Myoblast isolation was performed based on a protocol by Hindi et al. 2017. Adult Y118C/Y118C mice were euthanized by cervical dislocation after previous anesthesia with isoflurane. The mice were pinned face up on a sterile plate and both posterior limbs were sprayed with 70% ethanol followed by skin removal with sterilized scissors and forceps. Afterwards the quadriceps, anterior tibialis, gastrocnemius, soleus and extensor digitorum longus muscles were resected and placed in a dish containing the mincing solution, each limb separately. Muscles were washed in the solution and replaced in another dish with an equal mincing solution. With a second pair of sterilized scissors and forceps muscles were minced followed by centrifugation (20000g at room temperature for 30s) and aspiration of supernatant. The pellets containing the minced muscles were digested by transfer to a tube containing digestion medium with Collagenase II and filtered by pipetting through a syringe of a sterile 1ml injection. After shaking and vortexing several times for 5s respectively, the digested muscle solutions were centrifuged (1400g at room temperature for 5min) followed by supernatant removal. Remaining muscle pellets were resuspended in isolation media with multiple pipetting to enable release from tissue filaments and filtered through a pre-wet (70µm)-filter. Subsequently, the solutions were centrifuged followed by supernatant removal and seeded in 10cm dishes containing growth media. Following an incubation at 37°C with 5%CO₂ to allow cell expansion, myoblasts were maintained by culturing in growth medium as described and left to expand to <70% confluence before splitting. Myoblast were additionally stored in freezing medium at -80°C.

Stable U2OS cell lines with inducible SPRTN-expression

The FLP-In TREx U2OS host cell line was previously established in Dikic laboratory using the FLP-In TREx system by Invitrogen. FLP-In TREx U2OS host cells were cotransfected with pcDNA5/FRT/TO expression vector containing SPRTN and the pOG44 Flp recombinase plasmid (in order to integrate the SPRTN-vector). SPRTN was cloned into the pcDNA5/FRT/TO expression vector (as described), which contained the inducible promoter and the target site for Flp

recombinase to enable an integration as well as a antibiotic resistance gene for selection. Following cotransfection with two plasmids, U2OS cells were subjected to selection by an addition of antibiotic and left for selection for 10 days with splitting at low confluence as well as an exchange of the antibiotic-containing growth medium every 3 days for selection of stable transfectants. Surviving colonies were picked and further cultured. Transfection-efficiency was tested by incubation with Doxycycline for 24h and subsequent analysis by microscope and immunoblotting. Expression of SPRTN was induced in further experiments by addition of 1µg/ml Doxycycline for 24h.

SPRTN-knock-out

H7 conditional SPRTN^{flox/-};Cre-ER^{T2}-MEFs were provided by Machida et al. MEFs were previously established with a floxed SPRTN-allele (SPRTN^{flox/-}) and transduced with retroviruses to express the Cre-ER^{T2} recombinant fusion protein (Cre-ER^{T2}-MEFs) with the ability to convert the floxed SPRTN-allele into a KO-allele. H7 conditional SPRTN^{flox/-};Cre-ER^{T2}-MEFs from Machida et al. were treated with 2µM 4-hydroxytamoxifen (4-OHT) to generate SPRTN-KO-MEFs. Media was replaced by growth medium after 96h and MEFs were used for further experiments.

SPRTN-depletion via siRNA

HEK293T cells were seeded in a 6-well dish to a confluence of 60-80%. For each well a 6-well dish two separate solutions containing either 2.5µl Lipofectamine RNAiMAX (Invitrogen, 11668019) or 2.5µl of 20µM siRNA were diluted in 250µl Opti-MEM Reduced Serum Medium (Invitrogen, 31985062) respectively. The two prepared solutions were mixed in equal amounts and incubated at room temperature for 30min. Afterwards the mixture was added to cells in growth medium (500µl per well) and left for incubation at 37°C with 5%CO₂ for 3 days. SPRTN-depletion was checked by immunoblotting.

Transient transfection

HEK293T cells were transfected with different pEGFP-C1 plasmid (Clontech, France) (GFP-SPRTN, GFP-E112A-SPRTN, GFP- Δ C-SPRTN, GFP- Δ C-SPRTN-NLS, GFP-R408A-L411A-SPRTN, mCherry-SPRTN, mCherry-E112A-SPRTN, GFP-SPRTN-mCherry, GFP-E112A-SPRTN-mCherry) using the Genejuice transfection reagent (EMD Millipore, 70967-4). For each well of a 6-well plate 100 μ l Opti-MEM was prepared in a tube and 3 μ l Genejuice was added and mixed by vortexing. After incubation at room temperature for 5min, 700ng of plasmid was added, mixed by gentle pipetting and left to incubate for 15min. The prepared solution was added to cells in growth medium and left to incubate at 37°C with 5% CO₂ for 48h. Transfection efficiency was checked daily under a fluorescent microscope.

Treatment with DPC-inducing toxins

DNA-protein-crosslinks (DPCs) were induced using following drugs: Camptothecin (Santa Cruz; CAS 7689-03-4) dissolved in DMSO at 25mM stock concentration; Etoposide (Santa Cruz; CAS 33419-42-0) dissolved in DMSO to a 50mM stock concentration; Formaldehyde (Santa Cruz, sc-281692) dissolved in PBS to a 100mM stock concentration and Olaparib (LC labs, 0-9201) dissolved in DMSO to a 10mM stock concentration. Previously to each experiment a Mastermix of each drug was prepared and added simultaneously to each cell culture. Cells were treated by exchange of growth medium with the prepared Mastermixes of DPC-inducing toxins at final concentrations as indicated and incubated for time points as indicated for each experiment.

Degradation assay

Degradation of SPRTN was determined by a cycloheximide chase assay (as described) of cells previously subjected to replication blockage by L-mimosine, proteasomal inhibition by MG132 or endogenous conditions. L-mimosine treatment was performed by an exchange of growth media with a Mastermix containing L-mimosine at a final concentration of 500 μ M in DMEM/10%FBS followed by

incubation for 24h. The Mastermix was generated using a stock solution of 10mM L-mimosine (Sigma, M0253) dissolved in DMSO. MG132 (Calbiochem, 474790) was applied in a final concentration of 10 μ M for 1h from a stock solution of 10mM MG132 dissolved in DMSO (Roth, A994.1)

Cycloheximide chase assay

Cells were seeded in a 6-well plate in growth media at equal concentrations and left to grow to a confluence of 80%. Cycloheximide (CHX) (Appllichem, A0879) was dissolved in DMSO to a stock with a 100mg/ml concentration and added to growth media to form a final concentration of 100 μ g/ml by gently pipetting. The cycloheximide addition was performed for each well subsequently at different time point with 30min intervals in order to establish 6 different incubation times ranging from 180min to 0min. Cells were harvested from all wells of a 6-well plate simultaneously to generate total cell lysates as described. The collected lysates from all time points were subjected to SDS-PAGE and analyzed by immunoblotting.

Cell viability assay

Cells were seeded in a 96-well plate, seeding 200 cells per well after cell counting (as described). Samples were prepared using the conditional KO-system as explained and MEFs were incubated at 37°C in 5%CO₂ for 96h followed by cell counting, dilution and seeding in a 96-well plate. MEFs were exposed to a treatment with Etoposide at 50 μ M, Formaldehyde at 500 μ M or Camptothecin at 1000 μ M respectively. After treatment for 20h, DPC-toxin containing media were replaced by DMEM/10%FBS. Cell survival was determined by ATP concentration. ATP level was measured via a luciferase reaction and measured by the relative endpoint luminescence signal. Samples were prepared by the CellTiter-Glo Luminescent Cell Viability Assay (Promega, G755A). Plates were left at room temperature for 30min. The CellTiter-Glo Buffer and Substrate were incubated at room temperature and mixed by gentle vortexing. The mixed CellTiter-Glo Reagent was added to the growth medium in a 1:1 ratio, mixed on a shaker for

2min and left to incubate at room temperature for 30min. Luminescence was recorded using BioTek Synergy H1 hybrid multi mode reader. Control wells without cells and only growth medium were used for a background luminescence determination.

Clonogenic assay

HEK293T cells were counted as described and a total of 500 cells was seeded in each well of a 6-well plate and incubated at 37°C in 5%CO₂ for 24h. Afterwards, cells were treated with either Formaldehyde at 100µM or 2mM for 3h, Etoposide at 100µM or 2mM for 24h or Camptothecin at 1µM for 24h and after the indicated time the toxin-containing solutions were washed out. Following treatment, cells were left to grow in DMEM/10%FBS at 37°C in 5%CO₂ for a total of 10 days. The grown colonies were washed from growth medium with PBS and afterwards fixed and stained with 0.5% crystal violet in 6% glutaraldehyde for 30min followed by careful washing with H₂O and drying at room temperature. Colonies of more than 50 cells were macroscopically counted. Counting was repeated in triplicates. The survival fraction was defined by the ratio of the number of colonies formed after treatment and the number of colonies formed without treatment.

Microscopy

Live cell imaging was performed using 6-well-plates on a Visitron confocal spinning disc microscope with Yokogawa CSU-X1 Scanning head CSU X1-A1-5000rpm, single filter sets for GFP, mCherry, EBFP2 and a HXP 120 Xenon Lamp for epi-fluorescence illumination. Images were processed using ImageJ (NIH).

6.2 Molecular biology techniques

Buffers and solutions

PCR reaction mix

- 36µl H₂O
- 5µl Buffer Ultra II 10x
- 1,5µl dNTP
- 1µl Plasmid (concentration 25ng/mL)
- 5µl Oligonucleotide Forward or Oligonucleotide Reverse
- 1,5µl Pfu Ultra II

LB plates

- 5g Trypton
- 2.5g Yeast extract
- 2.5g NaCl
- 4g Agar Agar
- fill up to 0.5l with H₂O and autoclave
- 0.5ml Amp/Kan added later

LB medium

- 50mM Tris base pH 8.5
- 100mM NaCl
- 10µM ZnCl₂
- 10% Glycerol
- 1% NP40

Molecular cloning

SPRTN was tagged with GFP by inserting SPRTN into pEGFP-C1 plasmid (Clontech). Double-tagged GFP-SPRTN-mCherry was created by insertion of mCherry from pMcherry-N1 (Clontech) downstream of the GFP-SPRTN (pEGFP-C1). SPRTN was additionally cloned into the pcDNA5/FRT/TO expression vector. The gene of interest and plasmid were checked for restriction sites by Serial

Cloner program and the according restriction enzymes used for digestion of both DNA-products separately. 1µg of DNA was digested by both restriction enzymes in a 20µl reaction at 37°C for 2h followed by separation of the restriction fragments by agarose gel electrophoresis on a 0.9% gel running for 35min under a voltage of 100V. The separated restriction fragments were extracted using the Gel Extraction Kit by GenJET (Thermo Scientific, K0692) following manufacturer's instructions. Gel fragment were diluted in Binding buffer, bound in a Purification column by centrifugation followed by 2 washing steps and finally eluted from the column by centrifugation of an Elution buffer through the column. The extracted DNA fragments were ligated in a mixture containing 1µl of the vector plasmid and 5µl of the gene of interest with T4DNA Ligase (Biolabs, B0202S) in a reaction volume of 10µl T4 Buffer incubated at 25°C for 2h. Ligation product was used for reproduced in bacteria, plasmids prepared and sequenced as described.

Mutagenesis

Mutagenesis was performed to add point mutations to SPRTN (E112A, ΔC-SPRTN, NLS, R408A-L411A), using site-directed mutagenesis (Agilent, Germany). Oligonucleotides containing the desired mutations were designed manually and ordered from Sygma. Amplification of two complementary oligonucleotides was carried out by PCR. First, PCR-mixes of the plasmid and the oligonucleotide were prepared for both primers separately in a final reaction volume of 50µl with an addition of Pfu Ultra II polymerase (Agilent, 600674-51) shortly before amplification. PCR-amplification was performed by thermocycling in a Biometra T1 Thermocycle (Biometra/Analytik Jena, TProfessional TRIO) by programs with parameters adjusted to melting temperatures of primers and elongation times for product length following manufacturers instructions. After 8 cycles the complementary forward and reverse plasmids were combined followed by additional amplification cycles (17 repeats in total). Amplification was followed by digestion of parental DNA with Dpn I (1.5µl) (Thermo Scientific, FD1704) incubated at 37°C for 90min. Cleaning of the PCR reaction was performed using the GenJET PCR Purification Kit (Thermo Scientific, K0701) according to

manufacturer's instructions. The PCR-mix was added to 150µl Binding Buffer and the DNA collected in a Kit-column by spinning and washing with 700µL Washing Buffer followed by an elution in 25µl H₂O. Plasmids were cloned by transformation and analyzed by sequencing as described.

Plasmid multiplication

12.5µl of plasmid-DNA were transformed into 50µL DH5α *E.coli* strains (Invitrogen, 18258012) and incubated at 4°C for 10min followed by a heat shock at 42°C for 45s. Bacteria were incubated in 500µl LB medium at 37°C for 30min. 100µl of the bacteria-solution were plated on LB agar plates containing an antibiotic needed for selecting. Cultures were incubated at 37°C overnight and 3 single colonies per plate were inoculated in 5 ml liquid LB medium with selection antibiotics at 37°C for 9h.

Plasmid preparation

Plasmid DNA was purified using a Mini Prep purification Kit (Thermo Scientific, K0503) following manufacturer's guidelines. Bacteria were centrifuged at 6800g at room temperature for 2min and after supernatant removal resuspended in 250µl Resuspension Buffer, followed by lysis in 250µl Lysis Buffer for 1-2min, which was stopped by 350µl Neutralization Buffer. Cell debris was removed by centrifugation (17900g at 4°C for 10min) and supernatant was collected in a column by spinning down, followed by 2 cycles of washing in 700µl Washing Buffer. Plasmids were washed from the column by spinning down 50µl H₂O and suspension analyzed by sequencing.

Plasmid sequencing

Purified plasmids were diluted to 100ng/µl and 10µl from each plasmid-solution diluted in 5µl H₂O. Sequence analysis was performed by Sanger Cycle Sequencing (SeqLab-Microsynth GmbH, Göttingen, Germany).

Nucleic acid quantification

The concentration of DNA was determined photometrically using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific) following manufacturer's instructions provided by the. Absorbance was measured and concentrations calculated with the Nucleic Acid application. Samples were diluted to required DNA-concentrations as indicated.

6.3 Biochemical techniques

Buffers and solutions

Lysis buffer

- 50mM Tris/HCl pH 7.5
- 150mM NaCl,
- 1% NP40
- 5mM BME
- Complete protease inhibitors (EDTA-free)

Buffer A

- 10mM HEPES pH 7.9
- 10mM KCl
- 1.5mM MgCl₂
- 0.34 M Sucrose
- 10% Glycerol
- 1mM DTT
- Protease inhibitors cocktail

Buffer B

- 3mM EDTA
- 0.2mM EGTA
- 1mM DTT
- Protease inhibitors cocktail

1% Sarkosyl/TE

- 10mM Tris base pH 8.0
- 1mM EDTA
- 1% Sarkosyl

Laemmli buffer 4x

- 277.8mM Tris base pH 6.8

- 44.4% Glycerol
- 4.4% SDS
- 0,02% Bromophenol blue
- fill up to 50ml with H₂O
- addition of BME 5% before use

TBS 20x

- 121.85g Tris base
- 350g NaCl
- fill up to 5l with H₂O pH 7.6-7.7

TBS-T 0,1%

- 40ml 20% Tween20
- 400ml 20x TBS
- fill up to 8l with H₂O

TBS/BSA

- 100ml 1M Tris base
- 60ml 5M NaCl
- 20ml 10% Na-Acid
- 100g BSA
- 10ml 0,5% Phenol red
- fill up to 2l with H₂O pH 7.5
- filter with 0,45um

TBE 5x

- 54g Tris base
- 27.5g Boric acid
- 4.6875g Disodium EDTA
- fill up to 750ml H₂O pH 8.3
- 100ml of solution + 900ml H₂O for final TBE 0.5x for gel running

Transfer buffer 10x

- 840g Glycine
- 178.58g Tris base
- fill up to 8l with H₂O pH 8.3
- 100ml of solution + 200ml Methanol pH 8.3 + 700ml H₂O for final Transfer Buffer 1x

Ponceau

- 0.5% Ponceau S
- 5% TCA

Gel 10%

- 3ml 40% Acrylamide
- 3ml Lower buffer 4x
- 6ml H₂O
- 60µl 10% APS (ammonium persulfate)
- 6µl Temed
- fill up with Upper buffer

Lower buffer

- 91g Tris base
- 300ml H₂O pH 8.8
- take 500ml with H₂O, add 2g SDS

Upper buffer

- 30.25g Tris base
- 700ml H₂O pH 6.8
- take 500ml, add 2g SDS

Total cell extract lysis

Plates containing cell cultures were put on ice and after aspiration of growth medium washed with 1ml ice-cold PBS per well 2 times. Cells were detached and lysed by addition of 150µl Lysis buffer and pipetting inside the well. After incubation at 4°C for 10min, cell suspensions were transferred to a cooled Falcon tubes and sonicated (Sonics, Vibracell) ten times at 50% for 2s with an interval of 1min. Afterwards the solutions were centrifuged in a microcentrifuge (1300g at 4°C for 15min) to remove cell debris and 120µl of the supernatant were collected. Concentrated Laemmli buffer was added to the solutions in a ratio to form a final 1x concentration, followed by vortexing and boiling at 95°C for 10min.

In case of parallel chromatin fractionation and total cell lysis cells were initially lysed in the wells with 200µl of Buffer A containing Triton X with an incubation on ice for 8min and 65µl of each suspension were used for total cell lysis performed by addition of 135µl Lysis buffer and incubation for 15min. The solutions were afterwards sonicated three times, centrifuged, mixed with Laemmli buffer and boiled as described.

Chromatin fractionation

Cells were harvested from a 6-well plate by exchange of growth medium with 750µl PBS at 4°C followed by scratching with a cold plastic cell lifter and additional rinsing with 750µl PBS. The cell suspensions were transferred to a cooled microcentrifuge tubes. Centrifugation was performed in a microcentrifuge (1300g at 4°C for 2min) followed by aspiration of supernatant. After addition of 1ml ice-cold PBS without resuspension an additional centrifugation with supernatant removal was performed. Cell pellets were resuspended in 200µl of Buffer A with TritonX-100 and left to incubate on ice for 8min. 65µl were used for total cell lysis as described. The remaining 135µl were microcentrifuged (1300g at 4°C for 5min) followed by separation of supernatant. The supernatant was further microcentrifuged (20000g at 4°C for 5min) and the supernatant collected as cytosolic fraction. 400µl Buffer A was added to the remaining cell pellets, microcentrifuged and removed afterwards. Pellets were incubated in 200µl Buffer

B on ice for 30min, followed by microcentrifugation (1700g at 4°C for 5min) and supernatant removal. An additional resuspension in 400µl Buffer B was followed by microcentrifugation and removal of 400µl. The remaining chromatin pellet was lysed by an addition of 120µl Laemmli buffer, sonication in 3 cycles of 5s at 50% with intervals of 15s and subsequent boiling at 95°C for 10min. The chromatin-associated proteins were analyzed by SDS-PAGE and immunoblotting.

Cesium Chloride (CsCl) density gradient assay

A separation of DNA-bound and free proteins was established by a CsCl density gradient based on a protocol described by Aedo et al 2012. Samples were prepared from cell cultures grown in 6-well plates. On ice, growth medium was aspirated from each well followed by resuspension of the cells in 500µl 1% Sarkosyl/TE. After incubation on ice for 10min, cell lysates were replaced into tubes and sheared slowly with a 25gauge needle ten times. A CsCl stock solution with a density of 1.86g/ml was diluted to final densities of 1.82g/ml, 1.72 g/ml, 1.50g/ml or 1.37g/ml respectively. The solutions were subsequently put into an OptiSeal tube, 1ml of each in layers slowly to prevent mixing, with the highest density being at the bottom. Afterwards, 1ml of the prepared samples containing cell lysates were added on top of each solution respectively using 16gauge needles slowly without mixing. Finally, four OptiSeal tubes analyzed in parallel were carefully balanced on a scale with 1% Sarkosyl/TE to equal their weights. The tubes were sealed with the caps and placed in the SW41 rotor in a Beckman Coulter, Avanti J-30I.. Ultracentrifugation was performed at 36000rpm at 4°C for 20h with slow speeding up and slowing down to prevent disturbances of the gradient. After completion of the ultracentrifugation, the tubes were removed; the top 1ml of each solution removed and the necks of the tubes were cut off with a scalpel. The CsCl gradient solutions containing the fractionated cell lysates were separated in 20 fractions by careful collection of 250µl from the top respectively. The fractions were separately analyzed by dot blot and immunoblotting as described.

Dot Blot

Cell lysate fractions from CsCl-gradient assay were analyzed using the Biometra Dot Blot 96 Blotting System (Biometra Analytik Jena) First the nitrocellulose membrane was equilibrated in Transfer Buffer for 5min at room temperature for 5min. The dot blotting apparatus was set up and the membrane was placed inside on a Whatman filter paper previously wet in Transfer Buffer and covered with the 96-hole cover. After establishing a vacuumization by the membrane vacuum pump, 100µl of the Transfer Buffer were soaked through each hole. Afterwards, 200µl of solution of each cell lysate fraction was added in each hole. Under constant vacuumization the membrane was washed with 100µl Transfer Buffer per hole and left to vacuumize until complete drying. After release from the apparatus, the dot blot membrane was cut in parallel slices and blocked by BSA followed by a simultaneous analysis of all slices by immunoblotting as described.

Immunoprecipitation and co-immunoprecipitation

Samples used for immunoprecipitation were prepared from cells containing GFP-fused proteins. Cells were harvested by scraping and lysing as described. After cell lysis, 10% of the lysate was analyzed by immunoblotting as an input and the rest was used for immunoprecipitation of GFP-tagged SPRTN. GFP-trap agarose beads (ChromoTek, gta-20) were equilibrated following manufacturers instructions by vortexing and washing with dilution buffer for three times. Cell lysates diluted in dilution buffer were added to the equilibrated beads and incubated at 4°C for 1h under continuous rotation. After three washes with dilution buffer including centrifugation (2500g at 4°C for 2min) and supernatant removal, the GFP-trap beads were resuspended in Lysis buffer and boiled at 95°C for 5min for dissociation of complexes from beads. Following centrifugation, the supernatant collecting the GFP-tagged proteins and covalently bound complexes was analyzed by SDS-PAGE and immunoblotting as described.

SDS-polyacrylamide gel electrophoresis SDS-PAGE

Electrophoresis was performed using self-cast 10%-gels or precast 4-12%-gradient gels (Biorad, 4561095) in a running buffer at a voltage of 90V for different times (60-90 minutes) in a Biorad running system. Samples analyzed by SDS-PAGE were prepared by addition of Laemmli buffer to a final concentration of 1x followed by boiling at 95°C for 10 minutes. Equal amounts of cell lysates were loaded in each well parallel to a molecular weight marker.

Transfer from gel to membrane

Proteins separated by electrophoresis were transferred from the SDS-polyacrylamide gel on a nitrocellulose membrane. Membranes were previously put in transfer buffer for 1 min. Transfer was performed in a Biorad tank system in Transfer Buffer cooled by ice cubes under a constant current of 200mA for 90 minutes with checking of voltage to remain over 65V. Effective protein-transfer to the nitrocellulose membrane was checked by detection of proteins using Ponceau staining for 5min upon shaking and washed five times afterwards.

Immunoblot analysis

Membranes containing transferred proteins were blocked in TBS-T containing bovine serum albumin solution (BSA) (Thermo Scientific, 23209) for 1h. After washing out the blocking buffer with TBS-T membranes were incubated in primary antibody solutions overnight and washed 3 times with TBS-T for 10min respectively. Afterwards, membranes were incubated with a dilution of the HRP-conjugated secondary antibody for 1h followed by 3 washes with TBS-T for 10min respectively. All incubations were performed at room temperature with continuous shaking. ECL was performed by mixing 1ml of peroxide solution and 1ml Lumigen enhancer solution (Lumigen, TMA-100) immediately before incubating membranes in the mixture. After incubation at room temperature for 1 minute, membranes were removed from the solution and put in plastic sheet protector followed by removal of excess liquid by pressure and fixing of the plastic sheet in a film cassette. In a dark room X-ray films were placed above the membranes and left

for light exposure. Several X-ray films were used subsequently with increasing exposure time. X-ray films were developed in an AGFA, Curix 60 developer in the dark room and scanned afterwards.

Mass spectrometry

Mass spectrometry analysis of SPRTN substrates was performed using ^{WT}SPRTN and ^{ΔC}SPRTN cells in parallel. Cells were first SILAC-labeled (stable isotope labeling with amino acids in cell culture) by growing them in lysine- and arginine-free DMEM with 10% FBS, arginine with the addition of either heavy- or light-labeled lysine for 24h in order to allow cells to incorporate the labeled lysines in their proteins. Cells were harvested, lysed and mixed together in an equal ratio. The proteins were extracted by immunoprecipitation using agarose-coupled beads for 1h and subjected to elution on a SDS-gel reduction, alkylation with chloroacetamide and in-gel digestion with Trypsin overnight and extracted from the gel. Cleaning of the digested peptides was performed using high performance liquid chromatography columns to desalt and decontaminate them. Afterwards, the desalted peptide mixtures were stage tipped, reconstituted by injection and elution in high performance liquid chromatography organic solvents and loaded on fused silica capillaries followed by detection on an Orbitrap Fusion LUMOS (Thermo Scientific). The mass spectrometry data was processed using default parameters of the MaxQuant software with the human UniProt database.

7 References

1. Carter CO. Monogenic disorders. *J Med Genet.* 1977;14(5):316-20.
2. Burtner CR, Kennedy BK. Progeria syndromes and ageing: what is the connection? *Nat. Rev. Mol. Cell Biol.* 2010;11, 567–578.
3. Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. *Nat. Rev. Cancer* 2010;10, 353–361.
4. Navarro CL, Cau P, Levy N. Molecular bases of progeroid syndromes. *Hum. Mol. Genet.* 2006;15,R151–R161.
5. Pereira S, Bourgeois P, Navarro C, Esteves-Vieira V, Cau P, De Sandre-Giovannoli A, Lévy N. HGPS and related premature aging disorders: from genomic identification to the first therapeutic approaches. *Mech Ageing Dev.* 2008;129(7-8):449-59.
6. Oshima J, Sidorova JM, Monnat RJ Jr. Werner syndrome: Clinical features, pathogenesis and potential therapeutic interventions. *Ageing Res Rev.* 2017;33:105-114.
7. Yu CE, Oshima J, Wijsman EM, Nakura J, Miki T, Piussan C, Matthews S, Fu YW, Mulligan J, Martin GM, Schellenberg GD. Mutations in the consensus helicase domains of the Werner's syndrome gene. *Am J Hum Genet* 1997;60:330-341.
8. Martin GM. Genetic syndromes in man with potential relevance to the pathobiology of aging. *Birth Defects Orig. Artic. Ser.* 1978;14:5–39

9. Lessel D, Vaz B, Halder S, Lockhart PJ, Marinovic-Terzic I, Lopez-Mosqueda J, Philipp M, Sim JC, Smith KR, Oehler J, Cabrera E, Freire R, Pope K, Nahid A, Norris F, Leventer RJ, Delatycki MB, Barbi G, von Ameln S, Högel J, Degoricija M, Fertig R, Burkhalter MD, Hofmann K, Thiele H, Altmüller J, Nürnberg G, Nürnberg P, Bahlo M, Martin GM, Aalfs CM, Oshima J, Terzic J, Amor DJ, Dikic I, Ramadan K, Kubisch C. Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet.* 2014;46(11):1239-44.
10. Ruijs MW, van Andel RN, Oshima J, Madan K, Nieuwint AW, Aalfs CM. Atypical progeroid syndrome: an unknown helicase gene defect? *Am J Med Genet A.* 2003;116A(3):295-9.
11. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. The complete genome sequence of Escherichia coli K-12. *Science.* 1997;277(5331):1453-62.
12. Centore RC, Yazinski SA, Tse A, Zou L. Spartan/C1orf124, a reader of PCNA ubiquitylation and a regulator of UV-induced DNA damage response. *Mol Cell.* 2012;46(5):625-35.
13. Machida Y, Kim MS, Machida YJ. Spartan/C1orf124 is important to prevent UV-induced mutagenesis. *Cell Cycle.* 2012;11(18):3395-402.
14. Ghosal G, Leung JW, Nair BC, Fong KW, Chen J. Proliferating cell nuclear antigen (PCNA)-binding protein C1orf124 is a regulator of translesion synthesis. *J Biol Chem.* 2012;287(41):34225-33.

15. Lehmann AR, Niimi A, Ogi T, Brown S, Sabbioneda S, Wing JF, Kannouche PL, Green CM. Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst)*. 2007;6(7):891-9.
16. Juhasz S, Balogh D, Hajdu I, Burkovics P, Villamil MA, Zhuang Z, Haracska L. Characterization of human Spartan/C1orf124, an ubiquitin-PCNA interacting regulator of DNA damage tolerance. *Nucleic Acids Res*. 2012;40(21):10795-808.
17. Mosbech A, Gibbs-Seymour I, Kagias K, Thorslund T, Beli P, Povlsen L, Nielsen SV, Smedegaard S, Sedgwick G, Lukas C, Hartmann-Petersen R, Lukas J, Choudhary C, Pocock R, Bekker-Jensen S, Mailand N. DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. *Nat Struct Mol Biol*. 2012;19(11):1084-92.
18. Davis EJ, Lachaud C, Appleton P, Macartney TJ, N athke I, Rouse J. DVC1 (C1orf124) recruits the p97 protein segregase to sites of DNA damage. *Nat Struct Mol Biol*. 2012;19(11):1093-100.
19. Kim MS, Machida Y, Vashisht AA, Wohlschlegel JA, Pang YP, Machida YJ. Regulation of error-prone translesion synthesis by Spartan/C1orf124. *Nucleic Acids Res*. 2013;41(3):1661-8.
20. Maskey RS, Kim MS, Baker DJ, Childs B, Malureanu LA, Jeganathan KB, Machida Y, van Deursen JM, Machida YJ. Spartan deficiency causes genomic instability and progeroid phenotypes. *Nat Commun*. 2014;5:5744.
21. Lindahl T. Instability and decay of the primary structure of DNA. *Nature* 1993;362, 709–715.

22. Kunkel TA. DNA replication fidelity. *J Biol Chem* 2004;279:16895–16898.
23. Kunkel TA. Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol* 2009;74:91–101.
24. Kunkel TA. Balancing eukaryotic replication asymmetry with replication fidelity. *Curr Opin Chem Biol* 2011;15:620–626.
25. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature*. 2001;411(6835):366-74.
26. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–674.
27. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. *Nature* 2009;461:1071–1078.
28. Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. *Transl Cancer Res* 2013;2:107–129.
29. Henle ES, Linn S. Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide. *J Biol Chem* 1997;272:19095–19098.
30. Bjelland S, Seeberg E. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutat Res* 2003;531:37–80
31. Lindahl T, Barnes DE. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet*. 2004;38:445-476.
32. De Bont R, van Larebeke N. Endogenous DNA damage in humans: A review of quantitative data. *Mutagenesis* 2004;19:169–185.

33. Davies RJ. Royal Irish Academy Medal Lecture. Ultraviolet radiation damage in DNA. *Biochem Soc Trans* 1995;23(2):407–418.
34. Henner WD, Grunberg SM, Haseltine WA. Sites and structure of gamma radiation-induced DNA strand breaks. *J Biol Chem* 1982;257:11750–11754.
35. Hutchinson F. Chemical changes induced in DNA by ionizing radiation. *Prog Nucleic Acid Res Mol Biol* 1985;32:115–15
36. Reed RL, Ahern KG, Pearson GD, Buhler DR. Crosslinking of DNA by dehydroretronecine, a metabolite of pyrrolizidine alkaloids. *Carcinogenesis*. 1988;9(8):1355-61.
37. Chvalova K, Brabec V, Kasparkova J. Mechanism of the formation of DNA-protein cross-links by antitumor cisplatin. *Nucleic Acids Res*. 2007;35:1812–1821.
38. Barker S, Weinfeld M, Murray D. DNA-protein crosslinks: their induction, repair, and biological consequences. *Mutat Res*. 2005;589(2):111-35.
39. Pegg AE. Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. *Chem Res Toxicol* 2011;24(5):618–639.
40. Sanchez-Pulido L, Andrade-Navarro MA. The FTO (fat mass and obesity associated) gene codes for a novel member of the non- heme dioxygenase superfamily. *BMC Biochem* 2007;8:23

41. Kurowski MA, Bhagwat AS, Papaj G, Bujnicki JM. Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. *BMC Genomics* 2003;4:48
42. Odell ID, Wallace SS, Pederson DS. Rules of engagement for base excision repair in chromatin. *J Cell Physiol* 2013;228:258–266.
43. Huffman JL, Sundheim O, Tainer JA. DNA base damage recognition and removal: new twists and grooves. *Mutat Res* 2005;577:55–76.
44. Krokan HE, Bjørås M. Base excision repair. *Cold Spring Harb Perspect Biol.* 2013;5(4):a012583.
45. Dianov GL, Hubscher U. Mammalian base excision repair: The forgotten archangel. *Nucleic Acids Res* 2013;41:3483–3490.
46. Almeida KH, Sobol RW.. A unified view of base excision repair:lesion-dependent protein complexes regulated by post-translational modification. *DNA Repair (Amst)* 2007;6:695–711.
47. Sobol RW, Horton JK, Kühn R, Gu H, Singhal RK, Prasad R, Rajewsky K, Wilson SH. Requirement of mammalian DNA polymerase-beta in base-excision repair. *Nature.* 1996;379(6561):183-6.
48. Zharkov DO. Base excision DNA repair. *Cell Mol Life Sci.* 2008;65(10):1544-65.
49. Akbari M, Pena-Diaz J, Andersen S, Liabakk NB, Otterlei M, Krokan HE. Extracts of proliferating and non-proliferating human cells display different base excision pathways and repair fidelity. *DNA Repair (Amst)* 2009;8:834–843.

50. Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr Biol.* 2000;10(15):919-22.
51. Prasad R, Dianov GL, Bohr VA, Wilson SH. FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. *J Biol Chem.* 2000;275(6):4460-6.
52. Stucki M, Pascucci B, Parlanti E, Fortini P, Wilson SH, Hübscher U, Dogliotti E. Mammalian base excision repair by DNA polymerases delta and epsilon. *Oncogene.* 1998;17(7):835-43.
53. Min JH, Pavletich NP. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature.* 2007;449(7162):570-5.
54. Nishi R, Okuda Y, Watanabe E, Mori T, Iwai S, Masutani C, Sugasawa K, Hanaoka F. Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol Cell Biol* 2005;25:5664–5674.
55. Nospikel T. DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility *Cell Mol Life Sci.* 2009;66(6):994-1009.
56. Fousteri M, Vermeulen W, van Zeeland AA, Mullenders LH. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol Cell* 2006;23:471–482.

57. Compe E, Egly JM. TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* 2012;13:343–354.
58. Schärer OD. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol*. 2013;5(10):a012609.
59. Fagbemi AF, Orelli B, Scharer OD. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair (Amst)* 2011;10:722–729.
60. Vermeulen W, Fousteri M. Mammalian transcription-coupled excision repair. *Cold Spring Harb Perspect Biol*. 2013;5(8):a012625.
61. Arana ME, Kunkel TA. Mutator phenotypes due to DNA replication infidelity. *Semin Cancer Biol*. 2010;20(5):304-11.
62. Pavlov YI, Mian IM, Kunkel TA. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol*. 2003;13(9):744-8.
63. Sachadyn P. Conservation and diversity of MutS proteins. *Mutat Res*. 2010;694:20–30.
64. Ghodgaonkar MM, Lazzaro F, Olivera-Pimentel M, Artola-Borán M, Cejka P, Reijns MA, Jackson AP, Plevani P, Muzi-Falconi M, Jiricny J. Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell* 2013;50(3):323-32.
65. Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L, Li GM. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* 2005;122:693–705

66. Tishkoff DX, Amin NS, Viars CS, Arden KC, Kolodner RD. Identification of a human gene encoding a homologue of *Saccharomyces cerevisiae* EXO1, an exonuclease implicated in mismatch repair and recombination. *Cancer Res.* 1998;58(22):5027-31.
67. Genschel J, Modrich P. Mechanism of 5'-directed excision in human mismatch repair. *Mol Cell* 2003;12:1077–1086.
68. Davidovic L, Vodenicharov M, Affar EB, Poirier GG. Importance of poly(ADP-ribose) glycohydrolase in the control of poly(ADP-ribose) metabolism. *Exp Cell Res.* 2001;268(1):7-13.
69. McKinnon PJ, Caldecott KW. DNA strand break repair and human genetic disease. *Annu Rev Genomics Hum Genet* 2007;8:37–55.
70. Lan L, Nakajima S, Oohata Y, Takao M, Okano S, Masutani M, Wilson SH, Yasui A. In situ analysis of repair processes for oxidative DNA damage in mammalian cells. *Proc Natl Acad Sci U S A* 2004;101:13738–13743.
71. Caldecott KW. Single-strand break repair and genetic disease. *Nat Rev Genet.* 2008;9(8):619-31.
72. Khanna KK, Jackson SP. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet.* 2001;27(3):247-54.
73. Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell.* 2012;47(4):497-510.
74. Li X, Heyer WD. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* 2008;18:99–113.

75. Stracker TH, Petrini JH. The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol* 2011;12:90–103.
76. Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, Kowalczykowski SC. BLM DNA2- RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* 2011;25:350–362.
77. Symington LS. End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol.* 2014;6(8). pii: a016436.
78. Ferretti LP, Lafranchi L, Sartori AA. Controlling DNA-end resection: a new task for CDKs. *Front Genet.* 2013;4:99.
79. Heyer WD, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet.* 2010;44:113-39.
80. Holloman WK. Unraveling the mechanism of BRCA2 in homologous recombination. *Nat Struct Mol Biol* 2011;18:748–754
81. Renkawitz J, Lademann CA, Kalocsay M, Jentsch S. Monitoring homology search during DNA double-strand break repair in vivo. *Mol Cell.* 2013;50(2):261-72.
82. Renkawitz J, Lademann CA2, Jentsch S3. Mechanisms and principles of homology search during recombination. *Nat Rev Mol Cell Biol.* 2014;15(6):369-83.
83. Barber LJ, Youds JL, Ward JD, McIlwraith MJ, O’Neil NJ, Petalcorin MI, Martin JS, Collis SJ, Cantor SB, Auclair M, et al. RTEL1 maintains

genomic stability by suppressing homologous recombination. *Cell*. 2008;135:261–271.

84. Rass U, Compton SA, Matos J, Singleton MR, Ip SC, Blanco MG, Griffith JD, West SC.. Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev* 2010;24:1559–1569.
85. Chiruvella KK, Liang Z, Wilson TE. Repair of double-strand breaks by end joining. *Cold Spring Harb Perspect Biol*. 2013;5(5):a012757.
86. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem*. 2010;79:181-211.
87. Yano K, Chen DJ. Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end- joining pathway. *Cell Cycle* 2008;7:1321–1325.
88. Hartley KO, Gell D, Smith GC, Zhang H, Divecha N, Connelly MA, Admon A, Lees-Miller SP, Anderson CW, Jackson SP. DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell*. 1995;82(5):849-56.
89. West RB, Yaneva M, Lieber MR. Productive and nonproductive complexes of Ku and DNA-dependent protein kinase at DNA termini. *Mol Cell Biol*. 1998;18(10):5908-20.
90. Li S, Kanno S, Watanabe R, Ogiwara H, Kohno T, Watanabe G, Yasui A, Lieber MR. Polynucleotide kinase and aprataxin-like forkhead-associated protein (PALF) acts as both a single-stranded DNA endonuclease and a

single-stranded DNA 3' exonuclease and can participate in DNA end joining in a biochemical system. *J Biol Chem* 2011;286:36368–36377.

91. Goodarzi AA, Yu Y, Riballo E, Douglas P, Walker SA, Ye R, Härer C, Marchetti C, Morrice N, Jeggo PA, Lees-Miller SP. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO J*. 2006;25(16):3880-9.
92. Grawunder U, Wilm M, Wu X, Kulesza P, Wilson TE, Mann M, Lieber MR. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature*. 1997;388(6641):492-5.
93. Ramadan K, Shevelev IV, Maga G, Hubscher U. De novo DNA synthesis by human DNA polymerase lambda, DNA polymerase mu and terminal deoxyribonucleotidyl transferase. *J Mol Biol* 2004;339:395–404.
94. Moldovan GL, D'Andrea AD. How the fanconi anemia pathway guards the genome. *Annu Rev Genet*. 2009;43:223-49.
95. Noll DM, Mason TM, Miller PS. Formation and repair of interstrand cross-links in DNA. *Chem Rev*. 2006;106(2):277-301.
96. Kim JM, Kee Y, Gurtan A, D'Andrea AD. 2008. Cell cycle-dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. *Blood* 2008;111(10):5215-22
97. Montes de Oca R, Andreassen PR, Margossian SP, Gregory RC, Taniguchi T, Wang X, Houghtaling S, Grompe M, D'Andrea AD. Regulated interaction of the Fanconi anemia protein, FANCD2, with chromatin. *Blood* 2005;105(3):1003-9.

98. Clauson C, Scharer OD, Niedernhofer L. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb Perspect Biol* 2013;5(10):a012732
99. Knipscheer P, Räschele M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, Elledge SJ, Walter JC. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science*. 2009;326(5960):1698-701.
100. Klug AR, Harbut MB, Lloyd RS, Minko IG. Replication bypass of N2-deoxyguanosine interstrand cross-links by human DNA polymerases eta and iota. *Chem Res Toxicol* 2012;25:755–762.
101. Long DT, Raschle M, Joukov V, Walter JC. 2011. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* 2011;333(6038):84-7.
102. Lu K, Ye W, Zhou L, Collins LB, Chen X, Gold A, Ball LM, Swenberg JA. Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *J Am Chem Soc*. 2010;132(10):3388-99.
103. Bertolin AP, Mansilla SF, Gottifredi V. The identification of translesion DNA synthesis regulators: Inhibitors in the spotlight. *DNA Repair (Amst)*. 2015;32:158-164.
104. Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb Perspect Biol*. 2013;5(3):a012708.

105. Trincao J, Johnson RE, Escalante CR, Prakash S, Prakash L, Aggarwal AK. Structure of the catalytic core of *S. cerevisiae* DNA polymerase ϵ : Implications for translesion DNA synthesis. *Mol Cell* 2001;8:417–426.
106. Fu YV, Yardimci H, Long DT, Ho TV, Guainazzi A, Bermudez VP, Hurwitz J, van Oijen A, Schärer OD, Walter JC. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* 2011;146:931–941.
107. Kuo HK, Griffith JD, Kreuzer KN. 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication in vivo. *Cancer Res.* 2007;67:8248–8254.
108. Nakano T, Ouchi R, Kawazoe J, Pack SP, Makino K, Ide H. T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription. *J. Biol. Chem.* 2012;287:6562–6572.
109. Nakano T, Miyamoto-Matsubara M, Shoukamy MI, Salem AM, Pack SP, Ishimi Y, Ide H. Translocation and stability of replicative DNA helicases upon encountering DNA-protein cross-links. *J. Biol. Chem.* 2013; 288:4649–4658.
110. Sczepanski JT, Wong RS, McKnight JN, Bowman GD, Greenberg MM. Rapid DNA–protein cross-linking and strand scission by an abasic site in a nucleosome core particle. *Proc. Natl Acad. Sci. USA* 2010;107:22475–22480.
111. DeMott MS, Beyret E, Wong D, Bales BC, Hwang JT, Greenberg MM, Demple B. Covalent trapping of human DNA polymerase β by the oxidative DNA lesion 2-deoxyribonolactone. *J. Biol. Chem.* 2002;277:7637–7640.

112. Pommier Y, Huang SY, Gao R, Das BB, Murai J, Marchand C. Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). *DNA Repair (Amst.)* 2014;19:114–129.
113. Carey JF, Schultz SJ, Sisson L, Fazio TG, Champoux JJ. 2003. DNA relaxation by human topoisomerase I occurs in the closed clamp conformation of the protein. *Proc Natl Acad Sci U S A.* 2003;100(10):5640-5.
114. Champoux JJ. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem.* 2001;70:369-413.
115. Pommier Y, Marchand C. Interfacial inhibitors: targeting macromolecular complexes. *Nat. Rev. Drug Discov.* 2011;11:25–36
116. Burgin AB Jr, Huizenga BN, Nash HA. A novel suicide substrate for DNA topoisomerases and site-specific recombinases. *Nucleic Acids Res* 1995;23:2973–2979
117. Pourquier P, Ueng LM, Kohlhagen G, Mazumder A, Gupta M, Kohn KW, Pommier Y. Effects of uracil incorporation, DNA mismatches, and abasic sites on cleavage and religation activities of mammalian topoisomerase I. *J Biol Chem.* 1997;272(12):7792-6
118. Pommier Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nat Rev Cancer.* 2006;6(10):789-802.
119. Bush NG, Evans-Roberts K, Maxwell A. DNA Topoisomerases. *EcoSal Plus.* 2015;6(2).

120. Nitiss JL. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer* 2009;9:338–350
121. Murai J, Huang SY, Das BB, Renaud A, Zhang Y, Doroshow JH, Ji J, Takeda S, Pommier Y. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res.* 2012;72(21):5588-99.
122. Prasad R, Horton JK, Chastain PD, Gassman NR, Freudenthal BD, Hou, EW, Wilson SH. Suicidal cross-linking of PARP-1 to AP site intermediates in cells undergoing base excision repair. *Nucleic Acids Res* 2014;42:6337–6351.
123. Faure V, Sapparbaev M, Dumy P, Constant JF. Action of multiple base excision repair enzymes on the 2'-deoxyribonolactone. *Biochem Biophys Res Commun* 2005;328:1188–1195.
124. Sung JS, DeMott MS, Demple B. Long-patch base excision DNA repair of 2-deoxyribonolactone prevents the formation of DNA-protein cross-links with DNA polymerase β . *J Biol Chem* 2005;280:39095–39103.
125. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004;119:941–953.
126. Tretyakova NY, Groehler A 4th, Ji S. DNA-Protein Cross-Links: Formation, Structural Identities, and Biological Outcomes. *Acc Chem Res.* 2015;48(6):1631-44.
127. Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, Starr T.B. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol Sci* 2011;120:130–145.

128. Bandaru V, Sunkara S, Wallace SS, Bond JP. A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair* 2002;1:517–529.
129. Kooistra SM, Helin K. Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol.* 2012;13(5):297-311.
130. Ma TH, Harris MM. Review of the genotoxicity of formaldehyde. *Mutat Res.* 1988;196(1):37-59.
131. Maslov AY, Lee M, Gundry M, Gravina S, Strogonova N, Tazearslan C, Bendebury A, Suh Y, Vijg J. 5-aza-2'-deoxycytidine-induced genome rearrangements are mediated by DNMT1. *Oncogene* 2012;31:5172–5179.
132. Loeber R, Michaelson E, Fang Q, Campbell C, Pegg AE, Tretyakova N. Cross-linking of the DNA repair protein O6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards. *Chem Res Toxicol* 2008;21:787–795.
133. Costa M, Zhitkovich A, Harris M, Paustenbach D, Gargas M. DNA-protein cross-links produced by various chemicals in cultured human lymphoma cells. *J Toxicol Environ Health* 1997;50:433–449.
134. Giloni L, Takeshita M, Johnson F, Iden C, Grollman AP. Bleomycin-induced strand-scission of DNA. Mechanism of deoxyribose cleavage. *J Biol Chem* 1981;256:8608–8615.
135. von Sonntag C. Carbohydrate radicals: from ethylene glycol to DNA strand breakage. *Int J Radiat Biol.* 2014;90(6):416-22.

136. VanderVeen LA, Hashim MF, Nechev LV, Harris TM, Harris CM, Marnett LJ. Evaluation of the mutagenic potential of the principal DNA adduct of acrolein. *J Biol Chem* 2001;276:9066–9070
137. Chodosh LA. UV crosslinking of proteins to nucleic acids. *Curr Protoc Mol Biol*. 2001;Chapter 12:Unit 12.5.
138. Peak MJ, Peak JG, Jones CA. Different (direct and indirect) mechanisms for the induction of DNA-protein crosslinks in human cells by far- and near-ultraviolet radiations (290 and 405 nm). *Photochem Photobiol*. 1985;42(2):141-6.
139. Omar Desoukya ND, Zhou G. Targeted and non-targeted effects of ionizing radiation. *J Radiat Res Appl Sci* 2015;8:247–254.
140. Ravanat JL, Breton J, Douki T, Gasparutto D, Grand A, Rachidi W, Sauvaigo S. Radiation-mediated formation of complex damage to DNA: a chemical aspect overview. *Br J Radiol*. 2014;87, 20130715.
141. Meyn RE, vanAnkeren SC, Jenkins WT. The induction of DNA-protein crosslinks in hypoxic cells and their possible contribution to cell lethality. *Radiat Res*. 1987;109(3):419-29.
142. Zhang H, Koch CJ, Wallen CA, Wheeler KT. Radiation-induced DNA damage in tumors and normal tissues. III. Oxygen dependence of the formation of strand breaks and DNA-protein crosslinks. *Radiat Res*. 1995;142(2):163-8.
143. Oleinick NL, Chiu SM, Ramakrishnan N, Xue LY. The formation, identification, and significance of DNA-protein cross-links in mammalian cells. *Br J Cancer Suppl*. 1987;8:135-40.

144. Ide H, Shoukamy MI, Nakano T, Miyamoto-Matsubara M, Salem AM. Repair and biochemical effects of DNA-protein crosslinks. *Mutat Res*. 2011;711(1-2):113-22.
145. de Graaf B, Clore A, McCullough AK. Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *DNA Repair (Amst)*. 2009;8(10):1207-14.
146. Vance JR, Wilson TE. Yeast Tdp1 and Rad1-Rad10 function as redundant pathways for repairing Top1 replicative damage. *Proc Natl Acad Sci U S A*. 2002;99(21):13669-74.
147. Minko IG, Kurtz AJ, Croteau DL, Van Houten B, Harris TM, and Lloyd RS. Initiation of repair of DNA-polypeptide cross-links by the UvrABC nuclease. *Biochemistry* 2005;44:3000–3009.
148. Nakano T, Morishita S, Katafuchi A, Matsubara M, Horikawa Y, Terato H, Salem AMH, Izumi S, Pack SP, Makino K, Ide H. Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Mol Cell* 2007;28:147–158.
149. Nakano T, Katafuchi A, Matsubara M, Terato H, Tsuboi T, Masuda T, Tatsumoto, T, Pack SP, Makino K, Croteau DL, Van Houten B, Iijima K, Tauchi H, Ide H. Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *J Biol Chem* 2009;284:27065–27076.
150. Fornace AJ Jr, Seres DS. Repair of trans-Pt(II)diamminedichloride DNA-protein crosslinks in normal and excision-deficient human cells. *Mutat Res* 1982;94:277–284.

151. Baker DJ, Wuenschell G, Xia L, Termini J, Bates SE, Riggs AD, O'Connor TR. Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J Biol Chem* 2007;282:22592–22604.
152. Ridpath JR, Nakamura A, Tano K, Luke AM, Sonoda E, Arakawa H, Buerstedde JM, Gillespie DA, Sale JE, Yamazoe M, Bishop DK, Takata M, Takeda S, Watanabe M, Swenberg JA, Nakamura J. Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* 2007;67:11117–11122.
153. Garaycochea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* 2012;489:571–575.
154. Langevin, F, Crossan GP, Rosado IV, Arends MJ, Patel KJ. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* 2011;475:53–58.
155. Rosado IV, Langevin F, Crossan GP, Takata M, Patel KJ. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat. Struct. Mol. Biol.* 2011;18:1432–1434.
156. Kottemann M, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* 2013;493:356–363.
157. Orta ML, Calderón-Montaña JM, Domínguez I, Pastor N, Burgos-Morón E, López-Lázaro M, Cortés F, Mateos S, Helleday T. 5-Aza-2'-deoxycytidine causes replication lesions that require Fanconi anemia-dependent homologous recombination for repair. *Nucleic Acids Res.* 2013;41:5827–5836.

158. Pommier Y, Barcelo JM, Rao VA, Sordet O, Jobson AG, Thibaut L, Miao ZH, Seiler JA, Zhang H, Marchand C, Agama K, Nitiss JL, Redon C. Repair of topoisomerase I-mediated DNA damage. *Prog Nucleic Acid Res Mol Biol* 2006;81:179–229.
159. Pouliot JJ, Yao KC, Robertson CA, Nash HA. Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science*. 1999;286(5439):552-5.
160. Desai SD, Liu LF, Vazquez-Abad D, D'Arpa P. Ubiquitin-dependent destruction of topoisomerase I is stimulated by the antitumor drug camptothecin. *J Biol Chem*. 1997;272(39):24159-64.
161. El-Khamisy SF. To live or to die: a matter of processing damaged DNA termini in neurons. *EMBO Mol. Med*. 2011;3:78–88.
162. Cortes Ledesma F, El Khamisy SF, Zuma MC, Osborn K, Caldecott KW. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. Identified TDP2 as the first-characterized human enzyme with 5'-tyrosyl DNA phosphodiesterase activity, which is important for releasing covalent TOP2 adducts. *Nature* 2009;461:674–678.
163. Gao R, Schellenberg MJ, Huang SY, Abdelmalak M, Marchand C, Nitiss KC, Nitiss JL, Williams RS, Pommier Y. Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2·DNA and Top2·RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2). *J. Biol. Chem*. 2014;289:17960–17969.
164. Gomez-Herreros F, Romero-Granados R, Zeng Z, Alvarez-Quilón A, Quintero C, Ju L, Umans L, Vermeire L, Huylebroeck D, Caldecott KW,

- Cortés-Ledesma F. TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *PLoS Genet.* 2013;9:e1003226.
165. Hartsuiker E, Neale MJ, Carr AM. Distinct requirements for the Rad32Mre11 nuclease and Ctp1CtIP in the removal of covalently bound topoisomerase I and II from DNA. *Mol. Cell* 2009;33:117–123.
166. Malik M, Nitiss JL. DNA repair functions that control sensitivity to topoisomerase-targeting drugs. *Eukaryot. Cell* 2004;3:82–90.
167. Lin C P, Ban Y, Lyu YL, Desai SD, Liu LF. A ubiquitin-proteasome pathway for the repair of topoisomerase I-DNA covalent complexes. *J. Biol. Chem.* 2008;283:21074–21083.
168. Interthal H, Champoux JJ. Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1. *Biochem. J.* 2011;436:559.
169. Debethune L, Kohlhagen G, Grandas A, Pommier Y. Processing of nucleopeptides mimicking the topoisomerase I-DNA covalent complex by tyrosyl-DNA phosphodiesterase. *Nucleic Acids Res.* 2002;30:1198–1204.
170. Mao Y, Desai SD, Ting CY, Hwang, J, Liu LF. 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes. *J. Biol. Chem.* 2001;276:40652–40658.
171. Gómez-Herreros F, Schuurs-Hoeijmakers JH, McCormack M, Greally MT, Rulten S, Romero-Granados R, Counihan TJ, Chaila E, Conroy J, Ennis S, Delanty N, Cortés-Ledesma F, de Brouwer AP, Cavalleri GL, El-Khamisy SF, de Vries BB, Caldecott KW. TDP2 protects transcription from abortive

topoisomerase activity and is required for normal neural function. *Nat. Genet.* 2014;46:516–521.

172. Quievryn G, Zhitkovich A. Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteasome function. *Carcinogenesis* 2000;21:1573–1580.
173. Zecevic, A, Hagan E, Reynolds M, Poage G, Johnston T, Zhitkovich A. XPA impacts formation but not proteasome- sensitive repair of DNA-protein cross-links induced by chromate. *Mutagenesis* 2010;25:381–388.
174. Duxin JP, Dewar JM, Yardimci H, Walter JC. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell.* 2014;159(2):346-57.
175. Stinglele J, Schwarz MS, Bloemeke N, Wolf PG, Jentsch S. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell.* 2014;158(2):327-338.
176. Jongeneel CV, Bouvier J, Bairoch A. A unique signature identifies a family of zinc-dependent metallopeptidases. *FEBS Lett.* 1989;242(2):211-4.
177. Rawlings ND, Barrett AJ. Evolutionary families of peptidases. *Biochem J.* 1993;290 (Pt 1):205-18.
178. Aedo S, Tse-Dinh YC. Isolation and quantitation of topoisomerase complexes accumulated on Escherichia coli chromosomal DNA. *Antimicrob Agents Chemother.* 2012;56(11):5458-64.

179. Olsson T, Sandstedt K, Holmberg O, Thore A. Extraction and determination of adenosine 5'-triphosphate in bovine milk by the firefly luciferase assay. *Biotechnol Appl Biochem*. 1986;8(5):361-9.
180. Ullman KS, Powers MA, Forbes DJ. Nuclear export receptors: from importin to exportin. *Cell*. 1997;90(6):967-70.
181. Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J Biol Chem*. 2001;276(2):1317-25.
182. Hindi L, McMillan JD, Afroze D, Hindi SM, Kumar A. Isolation, Culturing, and Differentiation of Primary Myoblasts from Skeletal Muscle of Adult Mice. *Bio Protoc*. 2017;7(9). pii: e2248.
183. Stingele J, Bellelli R, Alte F, Hewitt G, Sarek G, Maslen SL, Tsutakawa SE, Borg A, Kjær S, Tainer JA, Skehel JM, Groll M, Boulton SJ. Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. *Mol Cell*. 2016;64(4):688-703
184. Stingele J, Jentsch S. DNA-protein crosslink repair. *Nat Rev Mol Cell Biol*. 2015;16(8):455-60.
185. Yeo JE, Wickramaratne S, Khatwani S, Wang YC, Vervacke J, Distefano MD, Tretyakova NY. Synthesis of site-specific DNA-protein conjugates and their effects on DNA replication. *ACS Chem Biol*. 2014;9(8):1860-8.
186. Larsen NB, Gao AO, Sparks JL, Gallina I, Wu RA, Mann M, Räschle M, Walter JC, Duxin JP. Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in *Xenopus* Egg Extracts. *Mol Cell*. 2019;73(3):574-588.e7.

187. Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J, Dikic I. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *Elife*. 2016;5. pii: e21491.
188. Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S, Singh AN, Vendrell I, Fischer R, Torrecilla I, Drobnitzky N, Freire R, Amor DJ, Lockhart PJ, Kessler BM, McKenna GW, Gileadi O, Ramadan K. Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Mol Cell*. 2016;64(4):704-719.
189. Toth A, Hegedus L, Juhasz S, Haracska L, Burkovics P. The DNA-binding box of human SPARTAN contributes to the targeting of Pol η to DNA damage sites. *DNA Repair (Amst)*. 2017;49:33-42.
190. Li F, Raczynska JE, Chen Z, Yu H. Structural Insight into DNA-Dependent Activation of Human Metalloprotease Spartan. *Cell Rep*. 2019;26(12):3336-3346.e4.
191. Morocz M, Zsigmond E, Tóth R, Enyedi MZ, Pintér L, Haracska L. DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. *Nucleic Acids Res*. 2017;45(6):3172-3188.
192. Wang Y, Xu M, Jiang T. Crystal structure of human PCNA in complex with the PIP box of DVC1. *Biochem Biophys Res Commun*. 2016;474(2):264-270.
193. Stinglele J, Habermann B, Jentsch S. DNA-protein crosslink repair: proteases as DNA repair enzymes. *Trends Biochem Sci*. 2015;40(2):67-71.

194. Balakirev MY, Mullally JE, Favier A, Assard N, Sulpice E, Lindsey DF, Rulina AV, Gidrol X, Wilkinson KD. Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates. *Elife*. 2015;4.
195. Borgermann N, Ackermann L, Schwertman P, Hendriks IA, Thijssen K, Liu JC, Lans H, Nielsen ML, Mailand N. SUMOylation promotes protective responses to DNA-protein crosslinks. *EMBO J*. 2019;38(8). pii: e101496.
196. Stinglele J, Bellelli R, Boulton SJ. Mechanisms of DNA-protein crosslink repair. *Nat Rev Mol Cell Biol*. 2017;18(9):563-573.
197. Vaz B, Popovic M, Ramadan K. DNA-Protein Crosslink Proteolysis Repair. *Trends Biochem Sci*. 2017;42(6):483-495.
198. Biggins S, Bhalla N, Chang A, Smith DL, Murray AW. Genes involved in sister chromatid separation and segregation in the budding yeast *Saccharomyces cerevisiae*. *Genetics*. 2001;159(2):453-70.
199. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*. 1999;285(5429):901-6.

200. Klages-Mundt NL, Li L. Formation and repair of DNA-protein crosslink damage. *Sci China Life Sci.* 2017;60(10):1065-1076.
201. Maskey RS, Flatten KS, Sieben CJ, Peterson KL, Baker DJ, Nam HJ, Kim MS, Smyrk TC, Kojima Y, Machida Y, Santiago A, van Deursen JM, Kaufmann SH, Machida YJ. Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Res.* 2017;45(8):4564-4576.
202. Halder S, Torrecilla I, Burkhalter MD, Popović M, Fielden J, Vaz B, Oehler J, Pilger D, Lessel D, Wiseman K, Singh AN, Vendrell I, Fischer R, Philipp M, Ramadan K. SPRTN protease and checkpoint kinase 1 cross-activation loop safeguards DNA replication. *Nat Commun.* 2019;10(1):3142.
203. Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science.* 2005;310(5755):1821-4.
204. Bienko M, Green CM, Sabbioneda S, Crosetto N, Matic I, Hibbert RG, Begovic T, Niimi A, Mann M, Lehmann AR, Dikic I. Regulation of translesion synthesis DNA polymerase eta by monoubiquitination. *Mol Cell.* 2010;37(3):396-407.
205. Foo MXR, Ong PF, Dreesen O. Premature aging syndromes: From patients to mechanism. *J Dermatol Sci.* 2019;pii: S0923-1811(19)30327-5.
206. Eskild W, Kindberg GM, Smedsrod B, Blomhoff R, Norum KR, Berg T. Intracellular transport of formaldehyde-treated serum albumin in liver endothelial cells after uptake via scavenger receptors. *Biochem J.* 1989;258(2):511-20.

207. Barker S, Weinfeld M, Zheng J, Li L, Murray D. Identification of mammalian proteins cross-linked to DNA by ionizing radiation. *J Biol Chem.* 2005;280(40):33826-38.
208. Hoeijmakers JH. DNA damage, aging, and cancer. *N Engl J Med.* 2009;361(15):1475-85.
209. Cheng J, Ye F, Dan G, Zhao Y, Wang B, Zhao J, Sai Y, Zou Z. Bifunctional alkylating agent-mediated MGMT-DNA cross-linking and its proteolytic cleavage in 16HBE cells. *Toxicol Appl Pharmacol.* 2016;305:267-273.
210. Cheng J, Ye F, Dan G, Zhao Y, Zhao J, Zou Z. Formation and degradation of nitrogen mustard-induced MGMT-DNA crosslinking in 16HBE cells. *Toxicology.* 2017;389:67-73.
211. Ide H, Nakano T, Salem AMH, Shoukamy MI. DNA-protein cross-links: Formidable challenges to maintaining genome integrity. *DNA Repair (Amst).* 2018;71:190-197.
212. Tubbs JL, Pegg AE, Tainer JA. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst).* 2007;6(8):1100-15.
213. Bowman RL, Wang Q, Carro A, Verhaak RG, Squatrito M. GliOVis data portal for visualization and analysis of brain tumor expression datasets. *Neuro-Oncology* 2017;19:139-141.

214. Fielden J, Ruggiano A, Popović M, Ramadan K. DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. *DNA Repair (Amst)*. 2018;71:198-204.
215. Yasui H, Kawai T, Matsumoto S, Saito K, Devasahayam N, Mitchell JB, Camphausen K, Inanami O, Krishna MC. Quantitative imaging of pO₂ in orthotopic murine gliomas: hypoxia correlates with resistance to radiation. *Free Radic Res*. 2017;51(9-10):861-871.

8 Abbreviations

3'-ends	DNA end with phosphate at 3' carbon of ribose ring
4-OHT	4-hydroxytamoxifen
5-aza-dC	Decitabine / 5-aza-2'-deoxycytidine
5'-ends	DNA end with phosphate at 5' carbon of ribose ring
53BP1	p53-binding protein 1
A	Alanine
A	Amper
aa	Amino acid
ADH	Alcohol dehydrogenase
AGT	O ⁶ -alkylguanine-DNA alkyltransferase
ALKBH1-8	Alkylation repair homologues 1-8
Amp/Kan	Ampicillin/Kanamycin
AP	Apurinic/aprimidinic
APE1	Apurinic/aprimidinic AP endonuclease 1
APS	Ammonium persulfate
APTX	Aprataxin
ATM	Ataxia-teleangiectasia mutated
ATP	Adenosine triphosphate
BER	Base excision repair
BLM	Bloom syndrome protein
BME	Beta-mercaptoethanol
BRCA1	Breast cancer type 1 susceptibillity protein
BRCA2	Breast cancer type 2 susceptibillity protein
BSA	Bovine serum albumin
C	Cysteine
c	centi-
C=O	Carbonyl group
C-terminus	Carboxy-terminus
C1orf124	Chromosome 1 open reading frame 124 / Spartan

CF	Chromatin fraction
CHK1	Checkpoint kinase 1
CHX	Cycloheximide
CO ₂	Carbon dioxide
CPT	Camptothecin
Cre	Causes recombination protein
CsCl	Cesium Chloride
CSA	Cockayne syndrome A
CSB	Cockayne syndrome B
CtIP	CtBP(C-terminal binding protein)-interacting protein
Da	Dalton
DDR	DNA damage response
DH5α	D. Hanathan 5 α <i>Escherichia coli</i> strain
DMEM	Dulbecco's modified Eagles's minimal essentail medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PKcs	DNA-dependent protein kinase, catalytic subunit
DNMT	DNA methyltransferase
dNTP	Deoxynucleotide triphosphate
DPC	DNA-protein-crosslink
Dpn I	N6-methyladenine(m6A)-dependent restriction enzyme
dsDNA	Double-stranded DNA
DSB	Double-strand breaks
DSBR	Double strand break repair
DTT	Dithiothreitol
DVC1	DNA damage protein targeting VCP
E	Glutamic acid
<i>E. coli</i>	<i>Escherichia coli</i>
E112A	Glutamic acid at aa112 mutated to Alanine
ECL	Electrochemiluminescence
EDTA	Ethylenediaminetetraacetic acid

EGTA	Ectazic acid
EME1	Crossover junction endonuclease EME1
ERCC1	Excision repair cross-complementing protein 1
ER ^{T2}	Estrogen receptor activated by tamoxifen
et al.	et alii
ETO	Etoposide
EXO1	Exonuclease 1
FA	Formaldehyde
FANCI	Fanconi anemia group I protein
FANCD2	Fanconi anemia group D2 protein
FBS	fetal bovine serum
flox	Flanked by LoxP
Flp	Flippase
FEN1	Flap endonuclease (FEN1)
FTO	Fat mass and obesity associated
g	gram
g	Gravity as relative centrifugal force
G2 phase	Growth phase 2
GEN1	Flap endonuclease GEN homolog 1
GFP	Green fluorescent protein
GG-NER	Global genome NER
H	Histidine
h	Hour
H ⁺	Hydrogen ion
H/H	Hypomorphic / hypomorphic
H2AX	Histone H2A family member X
H ₂ O	Water
H3	Histone H3 family member
H7	Inducible SPRTN-KO mouse cell line
HCl	Hydrochloric acid
HEK293T	Human embryonic kidney cell mutated SV40 large T antigen

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HExxH	Histidine-Glutamic acid-variable aa-variable aa-Histidine
HGPS	Hutchinson-Gilford Progeria Syndrome
HMGB1	High mobility group box 1
HN2	Chlormethine / mustine
HR	Homologous recombination
HRP	Horseradish peroxydase
HU	Hydroxyurea
ICE	<i>In vivo</i> complex of enzyme assay
ICL	Interstrand crosslinks
ID	Complex of FANCI andFANCD
IR	Ionizing radiation
K	Lysine
k	kilo-
K-K/R-X-K/R	Lysin-Lysine/Arginine-variable aa-Lysine/Arginine
KCl	Potassium chloride
KO	Knockout
KRPRL	Lysine-Arginine-Proline-Arginine-Leucine
Ku70	ATP-dependent DNA helicase II 70kDa subunit
Ku80	ATP-dependent DNA helicase II 80kDa subunit
L	Leucine
l	liter
L411A	Leucine at aa411 mutated to Alanine
LB	Lysogeny broth
LC	Liquid chromatography mass spectrometry
LIG1	DNA ligase 1
LIG3	DNA ligase 3
LIG4	DNA ligase 4
LMNA	Lamin A
M	mol
m	mili-

mCherry	Red fluorescent protein mCherry
MEF	Mouse embryonic fibroblast
MG132	Protease inhibitor MG132
MgCl ₂	Magnesium chloride
MGMT	O ⁶ -methylguanine-DNA methyltransferase
min	minute
MMC	MitomycinC
MMR	Mismatch repair
MMS	Methyl methanesulfonate
MRE11	Meiotic recombination 11
MRN	Complex consisting of MRE11, RAD50 and NBS1
MUS81	Crossover junction endonuclease MUS81
MutL	Mutator L complex
MutS	Mutator S complex
N	Nitrogen
N-terminus	Amino-terminus
NaCl	Sodium chloride
NBS1	Nijmegen breakage syndrome protein 1
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NLS	Nuclear localization signal
NP40	Nonyl phenoxypolyethoxyethanol
O ⁻	Oxyanion
OH ⁻	Hydroxide
OLA	Olaparib
Opti-MEM	Opti-MEM Reduced Serum Medium
P	Proline
p53	Cellular tumor antigen p53 / p53 protein
p97	Protein p97 cilium adhesin
PARP1	Poly ADP-ribose polymerase 1
PARylation	Poly ADP-ribosylation

PBS	Phosphate-buffered saline
pcDNA	Protamine complementary DNA
pcDNA5/FRT/TO	pcDNA with 5/Flp recombination target/Tetracycline On
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
pEGFP-C1	Plasmid expressing GFP
Pen/Strep	Penicillin-streptomycin
Pfu Ultra II	<i>Pyrococcus furiosus</i> DNA Polymerase Ultra 2
pH	Acidity scale
PIP	PCNA-interacting protein
pMcherry-N1	Plasmid expressing mCherry
PNKP	Polynucleotide kinase 3'-phosphatase
pOG44	Flp-Recombinase Expression plasmid
Pol β	DNA Polymerase β (beta)
Pol δ	DNA Polymerase δ (delta)
Pol ϵ	DNA polymerase ϵ (epsilon)
Pol η	DNA polymerase η (eta)
Pol λ	DNA Polymerase λ (lambda)
Pol μ	DNA Polymerase μ (mu)
R	Arginine
R408L	Arginine at aa408 mutated to Leucine
RAD5	Restriction site associated DNA-marker 5
RAD6	Restriction site associated DNA-marker 6
RAD18	Restriction site associated DNA-marker 18
RAD23	Restriction site associated DNA-marker 23
RAD50	Restriction site associated DNA-marker 50
RAD51	Restriction site associated DNA-marker 51
RBBP8	RB(retinoblastoma-associated protein)-binding protein 8
REV1	DNA repair protein REV1
REV7	DNA repair protein REV7
RFC	Replication factor C

RJALS	Ruijs-Aalfs syndrome
RMI1	RecQ-mediated genome instability protein 1
RMI2	RecQ-mediated genome instability protein 2
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPA	Replication protein A
rpm	Rotations per minute
RTEL1	Regulator of telomerase elongation helicase 1
S	Serine
s	second
S phase	Synthesis phase
SDSA	Synthesis dependent strand annealing
SILAC	Stable isotope labeling with amino acids in cell culture
siRNA	Small interfering RNA
SLX1	Synthetic lethal of unknown function protein 1
SLX4	Synthetic lethal of unknown function protein 4
Spo11	SPO11 initiator of meiotic double stranded break
SprT	Protein SprT in <i>Escherichia coli</i>
SPRTN	Protein with SprT-like domain at the N terminus / Spartan
SSB	Single-strand breaks
ssDNA	Single-stranded DNA
SDS	Sodium dodecyl sulfate
SDSA	Synthesis-dependent strand annealing
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
STR	Short tandem repeat
T4	Bacteriophage T4
TBE	Tris/Borate/EDTA
TBS	Tris-buffered saline
TBS-T	TBS with 0.1% Tween20
TC-NER	Transcription coupled NER
TCA	Trichloro acetic acid

TCL	Total cell lysate
TDP1	Tyrosyl-DNA-phosphodiesterase 1
TE	Tris buffer with EDTA
Temed	Tetramethylethylenediamine
TFIIH	Transcription factor II human
TLS	Translesion synthesis
TOM20	Translocase of outer mitochondrial membrane 20 homolog
Topo I	Topoisomerase I
Topo II	Topoisomerase II
Topo III	Topoisomerase III
TRex	Tetracycline response element
Tris	Tris(hydroxymethyl)aminomethane
Tween20	Polysorbate 20
U	Unit
U2OS	Human bone steosarcoma epithelial cell
Ub	Ubiquitin
Ub-PCNA	Ubiquitylated PCNA
Ub-SPRTN	Ubiquitylated SPRTN
UBZ	Ubiquitin-binding zinc finger
UV	Ultraviolet
V	Volt
VCP	Valosin-containing protein
WRN	Werner syndrome ATP-dependent helicase
Wss1	Weak suppressor of smt3
WT	Wild type
X	times / concentration
X-ray	Röntgen radiation
XLF	XRCC4-like factor
XPB	Xeroderma pigmentosum group B
XPC	Xeroderma pigmentosum group C
XPF	Xeroderma pigmentosum group F

XPG	Xeroderma pigmentosum group G
XRCC1	X-ray repair cross-complementing protein 1
XRCC4	X-ray repair cross-complementing protein 4
Y	Tyrosine
Y117C	Tyrosine at aa117 mutated to Cytosine
Y118A	Tyrosine at aa118 mutated to Alanine
Zn	Zinc
ZnCl ₂	Zinc chloride
ΔC	SPRTN mutation missing the C-terminus
μ	micro-
γH2AX	Phosphorylated Histone H2AX

9 Acknowledgements

First of all I would like to thank my supervisor Ivan Dikic for all his help, for his belief in me and for giving me the great opportunity to become a scientist. I am really sincerely thankful for the trust you have in me and for your constant support in all aspects, whenever I needed or asked you.

I am really thankful to Jaime Lopez-Mosqueda as my mentor in the laboratory and the conceiver of our project, for his support in the experiments and for the many additional experiments and data instrumental for the publication of our study. The same way I want to thank you for the tutoring, the discussions and the introduction and guidance through science and problem solving as well as for being a great example of a dedicated scientist.

Thanks go to the collaborators and helpers of this project, to Khartik Maddi for his work on yeast strains, to Yuichi Machida for providing conditional SPRTN-KO MEFs and to the members of the Dikic laboratory for providing essential cell lines, plasmids and materials for experiments. I want to especially highlight Tihana Bionda, Mira Polajnar, Doris Popovic and Veronique Schaeffer for the technical support, discussions, advices and the great time in the laboratory.

I would additionally like to express my gratitude to Joachim Steinbach and Michael Ronellenfisch for the support in my clinical path of becoming a neurologist and for enabling a connection of clinical work and science.

A very deep and honest thank I want to express to my parents Ante and Majda Prgomet and my sister Angela Prgomet Secan for the immeasurable love, faith and help during my whole life enabling me to fulfill my dreams, of which this thesis is an important step. Finally I wish to truly thank Zrinka Matak as the person at my side through all time. Thank you for your love and sincere belief in me, for making a better person of me and for the happiness you bring into my life.

10 Curriculum vitae

Career

- 02/2019 - now **Department of Neurology (Prof. Dr. Helmut Steinmetz) and "Dr. Senkenbergisches Institut für Neuroonkologie" (Prof. Dr. Joachim Steinbach)**
University Hospital Frankfurt, Frankfurt, Germany
Neurology Resident
- 11/2016 – 01/2019 **Department of Neurosurgery (Prof. Dr. Volker Seifert)**
University Hospital Frankfurt, Frankfurt, Germany
Neurosurgery Resident
- 11/2014 – 10/2016 **Institute of Biochemistry II (Prof. Dr. Ivan Dikic)**
Goethe University Frankfurt, Frankfurt, Germany
Scientist
- 06/2014 – 10/2014 **Health Center Zagreb County, Velika Gorica, Croatia**
General practitioner and emergency physician
- 12/2013 – 05/2014 **Clinical Hospital Dubrava, Zagreb, Croatia**
Internship
- 07/2007 – 10/2014 **Laboratory for Mineralized Tissue and Department for Anatomy and Clinical Anatomy, University of Zagreb School of Medicine, Zagreb, Croatia**
Scientific assistant

Education

- 11/2014 – now **Promotion study (Prof. Dr. Ivan Dikic)**
Institute of Biochemistry II
Goethe University Frankfurt, Frankfurt, Germany
- 11/2014 **Medical approbation** in Hesse, Germany

10/2007 – 05/2014	<p>Approbation</p> <p>University</p> <p>MD (medical doctor)</p> <p>Thesis: <i>"BMP-1-3 antibody delayed bone fracture healing"</i> (prof.dr.sc. Slobodan Vukičević); 07/2013</p> <p><i>suma cum laude</i></p> <p>Zagreb County fellowship</p>
07/2003 – 06/2007	<p>Gymnasium Velika Gorica, Velika Gorica, Croatia</p> <p>High school</p> <p>Zagreb County fellowship</p>
09/1998 – 06/2003	<p>Primary School Eugen Kumičić, Velika Gorica, Croatia</p> <p>Primary school</p>
09/1995 – 07/1998	<p>Catholic Primary School Niederkassel, Düsseldorf, Germany</p> <p>Primary school</p>

Awards and Honors

- 2014 **"Drago Perović" Award** for the best student of the University of Zagreb Medical School 2013
- 2010 **University Rectors Award** for the best student scientific work of the University of Zagreb *"TSH and vitamin D3 synergistically increase bone volume in rats with removed thyroid and parathyroid glands"*
- 2009 **Deans Award** for the best student scientific work of the University of Zagreb Medical School *"BMP-1-3 antibody delayed bone fracture healing"*
- 2008 **Deans Award** for the best student of the first year of the University of Zagreb Medical School 2007/2008

Publications

Lopez-Mosqueda J, Maddi K*, Prgomet S*, Kalayil S, Marinovic-Terzic I, Terzic J, Dikic I. **equal contribution*. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *Elife*. 2016;5. pii: e21491.

Dumic-Cule I, Eljuga D, Izadpanah A, Erjavec I, Prgomet S, Hladnik A, Bicanic I, Rora M, Vinter I, Grgurevic L. Dynamics of optic canal and orbital cavity development revealed by microCT. *Surg Radiol Anat*. 2014;36(10):989-92.

Grahovac G, Rajappa P, Vilendecic M, Zic R, Lambasa S, Prgomet S. Sinus pericranii in the left frontal region involving the superior eyelid: a case report. *J Neurol Surg A Cent Eur Neurosurg*. 2013;74 Suppl 1:e166-9.

Grgurevic L, Erjavec I, Lipar M, Stejskal M, Dumic-Cule I, Prgomet S, Maticic D, Vukicevic S. BMP7 plasma levels increase during early stage of bone fracture. *Bone* 2011; 48 Suppl 2:167.

Grgurevic L, Macek B, Mercep M, Jelic M, Smoljanovic T, Erjavec I, Dumic-Cule I, Prgomet S, Durdevic D, Vnuk D, Vukicevic S. Bone morphogenetic protein (BMP)1-3 enhances bone repair. *Biochem Biophys Res Commun*. 2011;408(1):25-31.

Grgurevic I, Erjavec I, Grgurevic L, Prgomet S, Krizanac S, Heinzl R, Romcic Z, Unic A, Petroveckic M, Kujundzic M, Vukicevic S. BMP7 and BMP1-3 antibody prevent development and promote regression of CCl4-induced liver fibrosis in rats. *J Hepatol*. 2011; 54 Suppl 1:1037.

Grgurevic L, Brkljacic J, Erjavec I, Cipicic A, Lipar M,

Dumic-Cule I, Prgomet S, Vukicevic S. A new device for repair of bone defects. *Bone*. 2010; 47 Suppl 1:106-107.

Grgurevic L, Durdevic D, Erjavec I, Sanader B, Prgomet S, Sarajlic G, Pintaric L, Vukicevic S. BMP-1-3 antibody delayed bone fracture healing. *Bone*. 2009;44 Suppl 2:309.

Schriftliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die dem Fachbereich Medizin der Johann Wolfgang Goethe-Universität Frankfurt am Main zur Promotionsprüfung eingereichte Dissertation mit dem Titel

Characterization of SPRTN, the first mammalian metalloprotease that repairs DNA-protein-crosslinks

in dem Institut für Biochemie II unter Betreuung und Anleitung von Prof. Dr. Ivan Dikic mit Unterstützung durch Jaime Lopez-Mosqueda, PhD ohne sonstige Hilfe selbst durchgeführt und bei der Abfassung der Arbeit keine anderen als die in der Dissertation angeführten Hilfsmittel benutzt habe. Darüber hinaus versichere ich, nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Ich habe bisher an keiner in- oder ausländischen Universität ein Gesuch um Zulassung zur Promotion eingereicht*. Die vorliegende Arbeit wurde bisher nicht als Dissertation eingereicht.

Vorliegende Ergebnisse der Arbeit wurden (oder werden) in folgendem Publikationsorgan veröffentlicht:

Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J, Dikic I. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *Elife*. 2016;5. pii: e21491.

(Ort, Datum)

(Unterschrift)

*) im Falle des Nichtzutreffens entfernen