



A University of Sussex DPhil thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Cellular and biochemical analyses of TDP1 mediated chromosomal break repair

A thesis submitted to the University of Sussex for the degree of Doctor
of Philosophy

By

Owen Spencer Wells



Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signed

Owen S. Wells

Acknowledgements

First of all I would like to express my gratitude to Dr Sherif El-Khamisy for giving me the opportunity to undertake a PhD and for his continued guidance and support throughout. A special thank you to Jessica Hudson for looking after me, spending time listening to my half-baked ideas and sharing her expertise within the laboratory. I would like to thank Aaron, Mariella and Kyle for their encouragement and input in my work, and also to all the members in the El-Khamisy lab, Jude and Conny, my office and the GDSC for putting up with me for the last 3.5 years and making my time here enjoyable. Finally, a special mention to Diana, Chris, Sophie, Carol and Marcel for helping me through this last year.

Outside of the centre I would like to thank my Mother for whom I am extremely grateful and proud of, for raising my brother and myself and helping us accomplish the goals in our life. My brother, Robert, who I look up too and whose drive and success in anything he puts his mind to has inspired me to do the same, or at least try. The rest of my family especially my Nan, Grandma and Grandad who have all had a pivotal role in my upbringing and enabled me to be the person I am today. Last but not least my uncle Mark whose modest persona has rubbed off on me and has always kept me grounded.

Of my friends back home I would like to thank Huw Wilkinson, Jonathan James, Jon Gouran, Gareth Sherrington, James Edwards, Richard Martin and Chris Jones for helping me maintain a social life, for always showing an interest in what I do and listening when I talk about my research. Everyone at Littlemore rugby club and my old house mates Bunk and Tim for helping take my mind off work.

The last person I would like to thank is Dominika Wlazly, for giving me the drive to undertake this PhD, for supporting me throughout and without whom I would never have accomplished these achievements.

University of Sussex

Owen S. Wells

Doctor of Philosophy Biochemistry

**Cellular and biochemical analyses of TDP1 mediated
chromosomal break repair**

Summary

Tyrosyl DNA phosphodiesterase 1 (TDP1) is an end-processing enzyme involved in the repair of abortive topoisomerase I (Top1) complexes. Although not essential for survival, a hypomorphic mutation in TDP1 is linked to the autosomal recessive ataxia, spinocerebellar ataxia with axonal neuropathy 1 (SCAN1). SCAN1 is a rare human condition linked with neurodegeneration and ataxic gait and patients are usually wheel chair bound by their early teens. TDP1 primarily cleaves lesions at the 3'-end of DNA breaks and its most prominent substrate is stalled Top1 linked to the 3'-terminus of DNA. The enzymatic mechanism by which TDP1 functions are well understood and inhibitors are now being investigated for treatment of cancer. In contrast, the processes involved in TDP1 recruitment, localisation and regulation during the DNA damage response remain unclear. This thesis investigates how the evolutionarily driven N-terminus of TDP1, not conserved in lower Eukaryotes, is required for optimal cellular protection against genotoxic stress. I also characterise how post-translational modifications of TDP1 allow for efficient repair of transcriptionally associated, chromosomal single-strand breaks and uncover new protein interacting partners of TDP1 and their role in TDP1 mediated repair.

Contents

Table of Contents

Chapter I.....	1
Introduction	1
1.1 General introduction	2
1.2 Sources of DNA damage	2
1.2.1 Endogenous sources of DNA damage.....	3
1.2.1.1 Reactive oxygen species	3
1.2.1.2 Methylation	3
1.2.1.3 Hydrolysis	4
1.2.1.4 Mismatch of bases.....	4
1.2.2 Exogenous sources of DNA damage.....	4
1.2.2.1 UV	5
1.2.2.2 Ionising radiation	5
1.3 DNA repair	6
1.3.1 DNA double-strand break repair.....	6
1.3.1.1 Homologous recombination.....	6
1.3.1.2 Non-homologous end-joining	7
1.3.2 Nucleotide excision repair.....	9
1.3.3 Mismatch repair	10
1.3.4 Base excision repair	11
1.3.5 Single-strand break repair	12
1.3.5.1 Detection: PARP	15
1.3.5.2 End-processing.....	15
1.3.5.3 Gap filling	15
1.3.5.4 Ligation	16
1.3.6 Ligase family.....	19
1.3.6.1 Ligase I.....	19
1.3.6.2 Ligase III	20
1.3.6.3 Ligase IV	21
1.3.7 Protein-linked DNA single-strand breaks	23
1.4 Tyrosyl DNA phosphodiesterase 1.....	23
1.5 Diseases associated with defects in single-strand break repair.....	26
1.5.1 Neural development	26
1.5.2 Autosomal recessive ataxias	29
1.5.3 Spinocerebellar ataxia with axonal neuropathy 1 (SCAN1).....	30
1.5.4 Microcephaly, infantile-onset seizures, and developmental delay (MCSZ) ..	31
1.5.5 X-linked mental retardation (XLMR)	32
1.5.6 Ataxia oculomotor apraxia 1 (AOA1).....	33
1.5.7 Ataxia oculomotor apraxia 2 (AOA2).....	34
1.6 Neurodegeneration: a consequence of repair defect in the nucleus, mitochondria or both?	35
1.7 Aims and objectives.....	37
Chapter II	38

Materials and Methods	38
2.1 Materials	39
2.2 Cloning and Molecular methods.....	41
2.2.1 PCR, restriction digests and ligations	41
2.2.2 Site-directed mutagenesis.....	41
2.2.3 Competent cells and transformations	43
2.2.3.1 <i>E.coli</i> DH5 α cells	43
2.2.3.2 Transformations	43
2.3 Electrophoresis and western blot analysis	44
2.3.1 Electrophoresis of DNA.....	44
2.3.2 Electrophoresis of proteins.....	44
2.3.2.1 Coomassie blue staining.....	45
2.3.2.2 Silver staining.....	45
2.4 Western blots	45
2.5 Protein expression and purification	46
2.5.1 Protein expression	46
2.5.2 Protein purification.....	47
2.5.2.1 Cell lysis.....	47
2.5.2.2 Immobilised metal affinity chromatography (IMAC) purification.....	47
2.6 Biochemical assays	48
2.6.1 SUMOylation assay	48
2.6.2 SUMO and TDP1 non-covalent binding studies.....	48
2.6.2.1 Recombinant TDP1 binding to SUMO immobilised resin	48
2.6.3 Casein Kinase 2 phosphorylation.....	49
2.7 Protein structure and folding.....	49
2.7.1 Circular dichroism.....	49
2.7.1.1 Quantitative analysis of circular dichroism by Dichroweb.....	50
2.7.2 Protein denaturation curves.....	50
2.8 Mammalian cell culture.....	51
2.8.1 Maintenance of cell lines	51
2.8.2 Transfection.....	51
2.8.2.1 Calcium phosphate transfection	51
2.8.2.2 Transfection by electroporation	52
2.9 Cell harvest and lysis	52
2.10 Co-immunoprecipitation	53
2.11 Creating stable cell lines in DT40 Tdp1-/- cells.....	53
2.12 Viability assays	54
2.13 FACS analysis.....	54
2.14 DNA single-strand break assays	54
2.14.1 Alkaline comet assay.....	54
2.14.1.1 Preparation	55
2.14.2 Gyrosyl assay	55
2.14.2.1 Preparation of oligonucleotide substrates:	55
2.14.2.2 TDP1 fluorescence assay:	55
2.15 Mass spectroscopy	56
Chapter III.....	58
Results 1: The N-terminal evolutionarily driven domain of TDP1 is required for optimal cellular protection against genotoxic stress	58
3.1 Introduction	59

3.1.1 TDP1; discovery, substrate and mechanism of action	59
3.1.2 TDP1's role in SCAN1	63
3.1.3 Chapter objectives	64
3.2 Results	67
3.2.1 Full-length TDP1 and the catalytic domain show no difference in activity, structure or stability.....	67
3.2.1.1 The C-terminal domain does not demonstrate a reduction in enzymatic activity <i>in vitro</i>	67
3.2.1.2 TDP1 and TDP1 ¹⁵¹⁻⁶⁰⁸ have no difference in thermal stability, or structure.....	70
3.2.2 Full-length hTDP1 is required for optimal protection against DNA damaging agents in DT40 cells, whilst hTDP1 ¹⁵¹⁻⁶⁰⁸ gives partial protection.....	75
3.2.2.1 Tdp1 ^{-/-} DT40 cells, complemented with Myc, Myc-hTDP1 or Myc-hTDP1 ¹⁵¹⁻⁶⁰⁸ maintain similar proliferation rates and cell cycle progression.....	77
3.2.2.2 TDP1 ¹⁵¹⁻⁶⁰⁸ does not exhibit reduced activity in cell extracts.....	77
3.2.2.3 Protection of cell lines treated with different genotoxic agents shows that the TDP1 ¹⁵¹⁻⁶⁰⁸ gives partial resistance when compared to the full-length protein	82
3.2.3 Accumulation of DNA single-strand breaks in hTDP1 ¹⁵¹⁻⁶⁰⁸ compared to hTDP1, is in part transcription dependent.....	85
3.2.3.1 DT40 Tdp1 ^{-/-} cells, and cells complemented with hTDP1 ¹⁵¹⁻⁶⁰⁸ accumulate more single-strand breaks than full-length hTDP1	85
3.2.3.2 DT40 Tdp1 ^{-/-} cells complemented with hTDP1 ¹⁵¹⁻⁶⁰⁸ accumulate more single-strand breaks than full-length hTDP1, and these appear to be transcription dependent	86
3.3 Discussion.....	88
Chapter IV	91
Results 2: SUMOylation facilitates TDP1 driven chromosomal break repair	91
4.1 Introduction	92
4.1.2 Post-translational modification by SUMOylation.....	92
4.1.3 SUMO	92
4.1.4 SUMOylation	93
4.1.5 SUMO family.....	93
4.1.6 Phospho-dependent SUMO motif.....	95
4.1.7 SUMO interacting motifs – SIMS.....	95
4.1.8 Molecular consequences of SUMOylation and SUMO-SIM interactions.....	96
4.1.9 Chapter objectives	99
4.2 Results	100
4.2.1 TDP1 is SUMOylated at lysine 111 by SUMO1 <i>in vitro and vivo</i>	100
4.2.1.1 TDP1 is SUMOylated <i>in vitro</i>	100
4.2.1.2 TDP1 SUMOylation occurs within the N-terminus <i>in vitro</i>	104
4.2.1.3 TDP1 SUMOylation occurs within the N-terminus at K111 <i>in vitro</i> ...	106
4.2.1.4 Mutation of lysine to arginine at 111 abrogates SUMOylation <i>in vitro</i> and <i>in vivo</i>	108
4.2.2 Mutation of lysine 111 to arginine does not alter protein structure, stability or catalytic activity	110
4.2.2.1 TDP1 and TDP1 ^{K11R} give similar melting temperatures and unfolding patterns	110

4.2.2.2 TDP1 ^{K111R} does not result in a reduced activity compared to wild-type	110
4.2.2.3 TDP1 ^{K111R} and TDP1 exhibit no significant structural differences	111
4.2.3 SUMOylation-deficient mutant accrues more DNA single-strand breaks that are in part transcription dependent	114
4.2.3.1 TDP1 K111R mutant shows reduced survival in the presence of CPT	114
4.2.3.2 TDP1 K111R accumulates more DNA single-strand breaks in the presence of CPT	116
4.2.3.3 Accumulation of DNA single-strand breaks is in part transcription dependant	116
4.2.4 TDP1 interacts non-covalently with SUMO2/3	119
4.2.4.1 TDP1 interacts with SUMO2 <i>in vitro</i>	119
4.2.4.2 TDP1 interacts with SUMO2 via the catalytic domain <i>in vitro</i>	119
4.2.4.3 Identification of SIM/s in TDP1	121
4.2.4.4 Mutation of the first two amino acids in potential SIM sites results in instability and co-expression with GroEL	121
4.2.4.5 TDP1 fragments within the catalytic domain causes loss of expression <i>in vivo</i>	122
4.3 Discussion	128
Chapter V	132
Results 3: The role of DNA ligases and CK2 during TDP1 mediated repair	132
5.1 Introduction	133
5.1.2 Ligase family and TDP1	133
5.1.3 Casein kinase II	134
5.1.4 Chapter objectives	135
5.2 Results	136
5.2.1 Ligase I and not ligase III α is implicated in TDP1-mediated nuclear repair	136
5.2.1.1 Ligase I deficient cells results in hypersensitivity to camptothecin	136
5.2.1.2 The conserved catalytic domain of TDP1 is sufficient to maintain interaction with Lig I	137
5.2.2 TDP1 is a novel substrate for Casein Kinase II	141
5.2.2.1 TDP1 is phosphorylated <i>in vivo</i>	141
5.2.2.2 Phosphorylation of TDP1 occurs in the N-terminal domain	141
5.2.3 Phosphorylation of TDP1 by CK2 inhibits SUMOylation, <i>in vitro</i>	144
5.2.4 TDP1 is phosphorylated by CK2 in whole cell extracts	150
5.3 Discussion	154
Chapter VI	157
Discussion: Cellular and biochemical analyses of TDP1 mediated chromosomal break repair	157
6.1 Discussion	158
Publications	167
References	168

Abbreviations

A-T	ataxia-telangiectasia
ATM	ataxia-telangiectasia mutated
ATR	ataxia-telangiectasia mutated and Rad3-related
ATRIP	ATR interacting protein
BER	base excision repair
BSA	bovine serum albumin
CDK	cyclin-dependent kinase
CHX	cycloheximide
CK2	casein kinase 2
CS	Chicken serum
CPT	camptothecin
DAPI	4'-diamino-2-phenylindole
DNA-PK	DNA-dependent protein kinase
DSB	double-strand DNA break
DRB	5,6-Dichloro-1- β -D-ribofuranosylbenzimidazole
FCS	foetal calf serum
HR	homologous recombination
hrs	hours
HU	hydroxyurea
IF	immunofluorescence
IR	ionising radiation
LCL	lymphoblastoid cell line
MCPH	primary microcephaly
MEF	mouse embryonic fibroblast

MMS	methyl methanesulfonate
MRN	Mre11-Rad50-Nbs1
NER	nucleotide excision repair
NHEJ	non-homologous end-joining
NBS	Nijmegen Breakage syndrome
PARP	Poly (ADP-ribose) polymerase
PARPi	PARP inhibitor
PBS	phosphate buffered saline
PIKK	phosphatidylinositol 3-kinase like kinase
RNAi	RNA interference
ROS	reactive oxygen species
RPA	replication protein A
SCAN1	spinocerebellar ataxia with axonal neuropathy 1
SD	standard deviation
S.E.M	standard error of the mean
SSB	single-strand DNA break
ssDNA	single-stranded DNA
TBS	tris-buffered saline
TBB	4,5,6,7-Tetrabromobenzotriazole
TDP1	tyrosyl DNA phosphodiesterase 1
TOP1	topoisomerase I
Ub	ubiquitin
UNT	untreated
UV	ultra-violet
WB	western blot
WCE	whole cell extract
WT	wild-type
XRCC1	X-ray repair cross-complementing protein 1

Figures

Figure 1.1 DNA double-strand break repair	8
Figure 1.2 Outcome of DNA single-strand breaks if left unrepaired.....	13
Figure 1.3 Enzymes involved in the repair of DNA single-strand breaks.....	14
Figure 1.4 DNA structure and damage modifying 3' and 5' phosphate backbone of DNA	17
Figure 1.5 Typical modifications at DNA breaks, cause and enzymes involved in repair	18
Figure 1.6 Ligase Family	22
Figure 1.7 Role of TDP1 in cycling and non-cycling cells.....	25
Figure 1.8 Defects in end-processing enzymes and links to disease	28
Figure 3.1.1 Cleavage of abortive Top1-DNA complexes by TDP1.....	62
Figure 3.1.2 Crystal structure of TDP1.....	63
Figure 3.1.3 Schematic of TDP1	65
Figure 3.1.4 Alignment of Human and Chicken Tdp1 amino acid sequence ..	66
Figure 3.2.1 Purification of recombinant full-length human TDP1 and TDP1 ¹⁵¹⁻⁶⁰⁸	69
Figure 3.2.2 Removal of the N-terminal domain does not impact upon activity <i>in vitro</i>	69
Figure 3.2.3 Loss of the N-terminal domain does not impact on stability <i>in</i> <i>vitro</i>	72
Figure 3.2.4 Comparison of hTDP1 and hTDP1 ¹⁵¹⁻⁶⁰⁸ using CD shows no gross structural difference	73
Figure 3.2.5 Phyre2: N-terminus of TDP1 is mainly unstructured.....	74
Figure 3.2.6 Complementation of DT40 Tdp1 ^{-/-} with Myc-hTDP1, Myc- hTDP1 ¹⁵¹⁻⁶⁰⁸ or Vector	76
Figure 3.2.7 Characterisation of DT40 cells for cell cycle and proliferation .	79
Figure 3.2.8 Removal of the N-terminal domain does not impact upon activity <i>in vivo</i>	80
Figure 3.2.9 Cell viability assay: healthy cells metabolise Resazurin to Resorufin.....	81
Figure 3.2.10 hTDP1 protects cells against a variety of DNA damaging agents, whilst the catalytic domain TDP1 ¹⁵¹⁻⁶⁰⁸ provides partial protection	84

Figure 3.2.11 hTDP1 ¹⁵¹⁻⁶⁰⁸ accumulates a higher amount of DNA single-strand breaks than full-length TDP1	87
Figure 3.2.12 The increase in accumulation of DNA single-strand breaks in hTDP1 ¹⁵¹⁻⁶⁰⁸ compared to the full-length protein is, in part, transcription dependent.....	87
Figure 3.3.1 The N-terminal domain of TDP1 is required for optimal DNA single-strand break TDP1 mediated transcription repair	90
Figure 4.1.1 SUMO conjugation and de-conjugation of target substrates	97
Figure 4.1.2 Molecular consequences of SUMOylation and SUMO-SIM interactions	98
Figure 4.2.1 <i>In vitro</i> components of the SUMOylation assay.....	102
Figure 4.2.2 TDP1 is primarily SUMOylated by SUMO1 <i>in vitro</i>	103
Figure 4.2.3 <i>In silico</i> analysis highlights 5 potential sites for SUMOylation.	105
Figure 4.2.4 TDP1 is SUMOylated in the N-terminus by SUMO1 <i>in vitro</i>	105
Figure 4.2.5 N-terminal TDP1 mutants show varying amounts of SUMOylation	107
Figure 4.2.6 TDP1 is SUMOylated at lysine 111 <i>in vitro</i> and <i>in vivo</i>	109
Figure 4.2.7 Lysine at 111 is conserved in several higher eukaryotes	109
Figure 4.2.8 Removal of the N-terminal domain does not impact upon activity <i>in vivo</i>	112
Figure 4.2.9 Comparison of hTDP1 and hTDP1 ¹⁵¹⁻⁶⁰⁸ using CD shows no gross structural difference between the proteins.....	113
Figure 4.2.10 TDP1 K111R mutant shows reduced survival in the presence of CPT	115
Figure 4.2.11 TDP1 K111R accumulates more DNA single-strand breaks in the presence of CPT	117
Figure 4.2.12 Accumulation of DNA single-strand breaks is in part transcription dependent.....	118
Figure 4.2.13 TDP1 binds to SUMO2 non-covalently.....	120
Figure 4.2.14 SUMO2 binds non-covalently to the catalytic domain of TDP1	120
Figure 4.2.15 Sequence analysis predicts 6 potential SIM sites in TDP1	123
Figure 4.2.16 Mutation of SIM sites results in truncation and expression of GroEL	123
Figure 4.2.17 Determination of truncation points to map SIM motifs	124
Figure 4.2.18 Truncations within the catalytic domain results in loss of protein expression.....	125
Figure 4.2.19 Fragmentation of TDP1 ¹⁵¹⁻⁶⁰⁸ by incubation with trypsin.....	127

Figure 4.3.1 Model for the repair of abortive Top1 complexes at sites of transcription.....	130
Figure 5.2.1 Interaction with hTDP1 and Ligase III occurs within the N-terminal domain of TDP1 in DT40 cells	138
Figure 5.2.2 Ligase I interacts with hTDP1 and ligase I deficient cells are hypersensitive to camptothecin.....	139
Figure 5.2.3 Ligase I interacts with hTDP1¹⁵¹⁻⁶⁰⁸.....	140
Figure 5.2.4 TDP1 is constitutively phosphorylated as treatment of TDP1 with Lambda phosphatase results in a faster migrating band, indicative of dephosphorylation	142
Figure 5.2.5 <i>In vivo</i> phosphorylation of TDP1 occurs within the N-terminal domain	143
Figure 5.2.6 Mutation of TDP1 S81A reduces SUMOylation <i>in vitro</i>.....	147
Figure 5.2.7 TDP1 is phosphorylated by CK2 <i>in vitro</i>, reducing levels of SUMOylation.....	148
Figure 5.2.8 Phosphorylation of TDP1 and not components of the SUMO pathway ablate SUMOylation of TDP1.....	149
Figure 5.2.9 Purified His-TDP1 and TDP1 S81A are phosphorylated in the presence of cell extract, in an ATP dependent manner.....	152
Figure 5.2.10 Purified His-TDP1 is phosphorylated by CK2	153
Figure 6.1 Schematic of TDP1 mediated nuclear repair	166

Tables

Table 2.2.1 DNA plasmids created and oligonucleotide sequences	42
Table 2.2.2 Antibodies	46
Table 2.2.3 Proteins expressed and purified	48
Table 3.2.1 Bioinformatics analyses of core secondary structure using CDSSTR algorithm	74
Table 3.2.2 Comparison between the crystal structure and predicted structure from CD.....	74
Table 4.2.1 Bioinformatics analyses of core secondary structure using CDSSTR algorithm	113

Chapter I

Introduction

1.1 General introduction

Organism development, metabolism, function, and reproduction are dependent on maintenance and faithful transmission of the genetic code. For duplication of the genetic material and translation into functional proteins DNA must be unravelled and then replicated or transcribed by the appropriate polymerases. Damage to DNA and/or failure to faithfully transmit the genetic code has been implicated in a myriad of diseases.

Thankfully, a network of complex and intertwined repair systems have evolved corresponding to the type of damage that has occurred, leading to efficient and timely repair or cell death. There are several types of DNA repair pathways, which depend on the type of damage, the proteins involved, and the state of the cell (stage of cell cycle, or post-mitotic cells).

This thesis focuses on tyrosyl DNA phosphodiesterase 1 (TDP1): a neuroprotective protein that if mutated in humans leads to spinocerebellar ataxia with axonal neuropathy 1 (SCAN1). Although I will touch on other types of repair pathways in this introduction, I will mainly focus on TDP1 mediated repair and the single-strand DNA repair pathways that are associated with neurological disorders.

1.2 Sources of DNA damage

Whilst maintenance of the genetic code is imperative to cell and organism survival its chemical make-up and requirement for modification to allow cellular activities, such as replication and transcription makes it susceptible to unwanted chemical attack that can come from endogenous and exogenous sources (R. C. Gupta and Lutz, 1999; Lindahl, 1993; Marnett, 2003).

1.2.1 Endogenous sources of DNA damage

Endogenous damage occurs through by-products of reactions, primarily through reactive oxygen species, cellular enzymes and errors in DNA synthesis that result from within the cellular environment.

1.2.1.1 Reactive oxygen species

Existence within an aerobic environment means that reactive oxygen species (ROS) are an unavoidable consequence and occur as a by-product of normal cellular metabolism. Production of energy through reduction of oxygen within the mitochondria generates reactive oxygen intermediates such as hydroxyl radicals, superoxide radicals, and hydrogen peroxide (Chance et al., 1979; Floyd et al., 2001). Singlet oxygen atoms formed can also be reduced by neutrophils and macrophages as part of the immune response (Jackson and Loeb, 2001). These unstable oxygen species can interact with other cellular components to create secondary reactive intermediates. Ferric ion and nitric oxide are examples of these which lead to generation of more potent oxidants (Marnett, 2003). Due to the many reactive species formed (Calcerrada et al., 2011; J. M. Lee et al., 2002) and their ability to attack nucleic acid bases, deoxyribose residues, or the phosphodiester backbone, numerous types of DNA lesions can occur and levels of 3-3000 adducts per cell have been observed (Marnett, 2003). If unrepaired, these lesions can lead to DNA base transitions and transcription stalling resulting in single and double-stranded breaks, all of which can result in apoptosis and disease.

1.2.1.2 Methylation

Whilst damage from reactive oxygen and nitrogen species are the most prevalent endogenous sources of damage to DNA it is not the only threat that DNA repair pathways have to cope with. Various cellular enzymes in metabolic pathways can also impact upon the fragile state of our genetic code. One particular example is that of S-Adenosyl methionine (SAM). SAM is a small molecular co-substrate made up of ATP and methionine and is used in methylation reactions of DNA; an epigenetic mechanism for control of gene expression (Holliday and Ho, 1998a; 1998b). However, SAM can result in non-enzymatic methylation of DNA, which can be mutagenic (Rydberg and

Lindahl, 1982) and can lead to cytotoxic lesions blocking replication (Tudek et al., 1992).

1.2.1.3 Hydrolysis

The N-glycosidic bond linking the base to the sugar phosphate backbone is labile under certain conditions, such as heating, alkylation of bases or cleavage by glycosylases as part of base excision repair (Hegde et al., 2008; Lindahl, 1993). Cleavage of the glycosidic bond leads to generation of an abasic site (AP). With an estimated 10 000 lesions/human cell/day (Lindahl, 1993) AP sites are one of the most frequent lesions found in DNA with depurination occurring ~ 20 times more frequently than depyrimidination (Lindahl and Karlström, 1973).

Due to preferential incorporation of adenine by polymerases, next to the AP site, during replication (Jackson and Loeb, 2001) if left un-repaired AP sites can induce base pair substitutions or frame-shift mutations leading to problematic DNA synthesis and ultimately mutagenesis (Jackson et al., 1998; Lawrence et al., 1990).

1.2.1.4 Mismatch of bases

Throughout replication and DNA repair, DNA polymerases are required to accurately copy the correct base opposite the template strand of DNA. In addition to the insertion of the matching base some *E.coli* and eukaryotic polymerases also contain intrinsic endo- and exonuclease activity to excise mismatch bases (Boyer et al., 2013). Although uncommon, DNA polymerases can introduce incorrect bases that if unrepaired will form permanent mutations that are passed down through replication. Proof-reading loss of function through mutation has been achieved in multiple organisms, including mice and unsurprisingly leads to an increase in mutagenesis and susceptibility to cancer (Kunkel, 2003; Lange et al., 2011).

1.2.2 Exogenous sources of DNA damage

There are a multitude of environmental and man-made chemicals that can either directly or indirectly affect the stability of DNA.

1.2.2.1 UV

Ultraviolet (UV) light can be split into two types of damage: UV-A damage occurs indirectly and is a result of free radical production leading to reactive oxygen species that is discussed in 1.2.1.1. UV-B causes direct damage through the creation of pyrimidine dimers; cyclobutane pyrimidine dimers (CPDs) and pyrimidine-(6-4)-pyrimidone photoproducts (6-4PPs) (Otoshi et al., 2000) and can interfere with base-pairing during replication. Although translesion polymerases can by-pass this damage during replication, these polymerases also exhibit low fidelity and are prone to the insertion of adenine, thus causing G:C – A:T transitions during replication (Otoshi et al., 2000; Sale et al., 2012). Mutations in the nucleotide excision repair pathway that are responsible for the removal of UV induced damage can lead to Xeroderma pigmentosum; a rare disease resulting in sensitivity to sunlight and predisposition to skin cancer (Lehmann et al., 2011), typifying the mutational effect UV can have on DNA.

1.2.2.2 Ionising radiation

Ionising radiation (IR) involves the transfer of energy in the form of electromagnetic waves or particles that contain enough kinetic energy to excite and remove electrons from an atom. Natural sources of IR include cosmic rays and radioactive material from the earth as in the case of radon-derived α particles (Moore et al., 2014). Artificial (man-made) sources also occur from medical imaging, such as X-rays, radiation therapy and nuclear reactions. These sources can differ in the type of IR and are often referred to as high LET (linear energy transfer), such as alpha particles and low LET such as gamma rays or X-rays. High LET sources account for the majority of our exposure to IR and have a very high energy but short range and thus the DNA damage caused tends to be localized to a confined region, resulting in a myriad of chemical modifications at the DNA strand break (Hada and Georgakilas, 2008; World Health Organization, 2009). Low LET, such as X-rays have a lower energy but can travel much further distances. This causes DNA damage to be more sporadic, a stark contrast to high LET, and lesions are repaired with a greater speed (Eccles et al., 2011; Moore et al., 2014). DNA damage can also occur indirectly through IR as water molecules in the cell become ionized resulting in reactive oxygen species. Failure to repair these double-strand breaks can lead to cell death, mutations and has been linked to cancer (Heigener, 2011; S. Sun et al., 2007) .

1.3 DNA repair

Successful and efficient repair from the wide array of damage has led to the need for multiple repair pathways. When damage causes a break in both strands of the DNA such as lesions within close proximity on either strand of the DNA, a signalling cascade for repair via the DNA double-strand break repair pathway occurs and if it fails, this damage can lead to cell death.

1.3.1 DNA double-strand break repair

The repair of double-strand breaks (DSBs) is represented by two major pathways; homologous recombination (HR) and non-homologous end-joining (NHEJ).

1.3.1.1 Homologous recombination

Homologous recombination uses a DNA template from a sister chromatid and as a result occurs only in proliferating cells, and is restricted to late S and G2 phases of the cell cycle (Kasperek and Humphrey, 2011). Once the double-strand break has occurred the MRN complex (MRE11, RAD50 and NBS1) binds to both sides of the DNA break through MRE11 (**Figure 1.1**). RAD50 tethers the broken ends of the DNA and NBS1 causes activation of the ATM kinase through ATM autophosphorylation. Once active, ATM phosphorylates a multitude of other proteins such as BRCA1 and H2AX to form γ H2AX, the first step in the signalling process (Goodarzi and Jeggo, 2012; Uziel et al., 2003). This signals the recruitment and assembly of double-strand break repair complexes leading to chromatin restructuring/relaxation, resection of the damaged DNA termini by MRE11, C-terminal-binding protein interacting protein (CtIP) and exonuclease 1 (EXO1) (Eid et al., 2010; Sartori et al., 2007) and coating of the exposed DNA single-strand with Replication Protein A (RPA) (Jeppesen et al., 2011). Homology search is initiated and strand invasion occurs with the help of RAD54 and nucleoprotein filaments (made up of RAD51 monomers). Strand invasion leads to the formation of Holliday junctions. The polymerase then carries out the repair synthesis

before Ligase I seals the nicks creating a double Holliday junction. Once ligated, Holliday junctions can either be resolved by a resolvase such as GEN1 and SLX1/SLX4 (Mimitou and Symington, 2009), creating crossover products, or dissolved by TOPO3 α , BLAP75 and BLM RecQ helicase to generate non-crossover products (Kass and Jasin, 2010). Due to the error free repair HR appears to be the favoured mode of repair during the early stages of neuronal development. For a more comprehensive explanation of HR see reviews (Heyer et al., 2010; San Filippo et al., 2008).

1.3.1.2 Non-homologous end-joining

NHEJ occurs in both proliferating and terminally differentiated cells and is the only double-strand break repair pathway used in post-mitotic cells. Unlike HR, NHEJ does not require a sister chromatid as no template is required. Instead NHEJ modifies the two ends using various nucleases (Lees-Miller and Meek, 2003; Lieber et al., 2008) so that they are compatible (3'-hydroxyl and a 5'-phosphate), followed by ligation with DNA Ligase IV. Initially the DSB accrues the Ku70-Ku80 heterodimer to both ends of the break (Rivera-Calzada et al., 2007; Walker et al., 2001) (**Figure 1.1**). The binding of Ku70-Ku80 aligns the ends of the DNA through binding of the sugar phosphate backbone. These proteins form a complex with the catalytic sub-unit of DNA-PK (DNA-PK_{cs}), which activates DNA end-processing by artemis (Goodarzi et al., 2006) and XRCC4, DNA ligase IV and XLF to facilitate the ligation of the DNA ends (Ahnesorg et al., 2006; Riballo et al., 2009). In contrast to HR NHEJ is not error free and use of this pathway can lead to gene rearrangements, deletions and mutations, causing post-replicative cells to be more vulnerable to DSBs. DNA repair deficiencies in either of these pathways has been linked to multiple disorders (O'Driscoll and Jeggo, 2006).

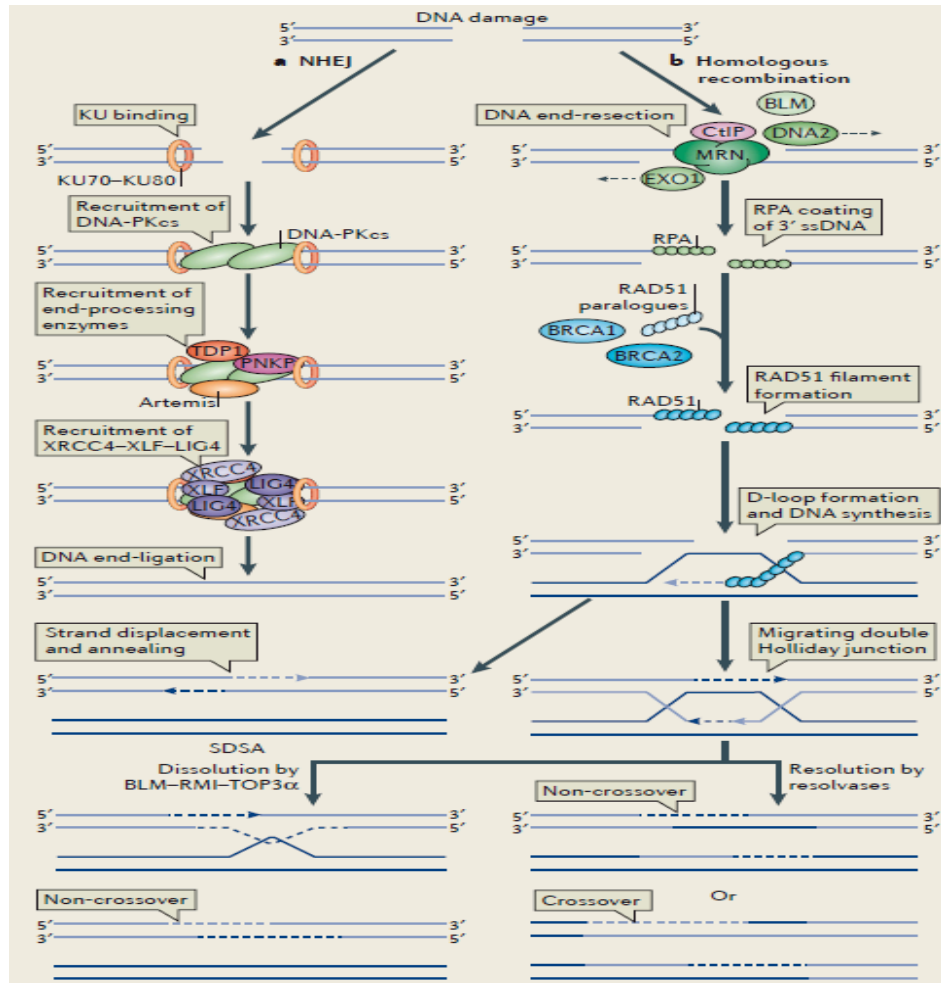


Figure 1.1 DNA double-strand break repair

The two main DNA double-strand break pathways are **(A)** non-homologous end-joining (NHEJ) and **(B)** homologous recombination (HR). NHEJ is initiated by binding of the KU70–KU80 heterodimer to both ends of the DNA break. This is followed by recruitment and activation of DNA-dependent protein kinase catalytic subunit (DNA-PKcs). Two key functions of DNA-PKcs are to tether the broken DNA ends and to recruit various end-processing factors, such as Artemis. These end-processing enzymes prepare the DNA ends for re-ligation by the XRCC4–XLF–DNA LIG4 complex. Unlike NHEJ, which can function in all stages of the cell cycle, HR-directed DSB repair is restricted to the late S and G2 phases of growing cells. It is detected by the MRN (MRE11–RAD50–NBS1) complex and followed by DNA end-resection. DNA end-resection is highly regulated and requires the activity of several nucleases, including CtBP-interacting protein (CtIP), the Resection generates long stretches of 3' single-stranded DNA that are coated by RPA. RAD51 then displaces RPA to form a RAD51–ssDNA nucleofilament. The nucleofilament searches for and binds with a homologous sequence elsewhere in the genome, to form a displacement D-loop, in which DNA synthesis is initiated to replace the DNA surrounding the former break site. Finally, depending on the type of HR, the D-loop is resolved either by dissociation of one of the invading strands (called synthesis-dependent strand annealing (SDSA)), or through migrating double Holliday junction intermediates that are cleaved by resolvases or dissolved by the BLM–RMI–TOP3α complex. Figure taken from (Panier et al., 2014)

1.3.2 Nucleotide excision repair

Most of the genome of an organism at any given time is transcriptionally silent, whilst a small proportion is found to be transcriptionally active utilising transcription machinery to produce mRNA and protein for cellular function. It is therefore not surprising that helical distortion of the DNA from UV light, resulting in pyrimidine dimers (e.g cyclobutane dimers or 6,4-photoproducts) in transcriptionally silent regions and DNA lesions blocking RNA polymerases, in transcriptionally active genes, are repaired at different rates through two sub-pathways of nucleotide excision repair (NER). Damage in transcriptionally silent regions requiring NER is termed global genome-NER (GG-NER), while removal of lesions blocking transcription machinery is called transcription coupled-NER (TC-NER). The most prominent difference between the sub-pathways is believed to be at the point of recognition; XPC-RAD23B identifies bulky distortions in the helix structure in GG-NER and initiation and recruitment of CSB, CDSA and XAB2 for repair occurs through arrested RNA polymerases in TC-NER.

Once the site of damage is identified both pathways amalgamate to form a common repair process. TFIIH a complex consisting of a 7 subunit core (XPD, XPB, P62, P52, P44, P34 and TTDA) and a 3 subunit (CDK7, Cyclin H and Mat1) coiled cyclin activating kinase complex (CAK) is recruited to the site of damage. XPB and XPD, 5'-3' and 3'-5' DNA helicases, found in the TFIIH complex, are then able to unwind the helix in proximity to the lesion located in the DNA. Following this, recruitment of XPA and RPA, resulting in the dissociation of the CAK complex and protection of the single-stranded DNA allows the full opening of the DNA around the lesion to occur. This, in turn favours recruitment of XPG and XPF endonucleases to incise and remove a short stretch of single-stranded DNA of about 25-30 nucleotides containing the lesion. Polymerases are then recruited through PCNA, RPA and the clamp loader RFC to fill in the excised fragment. Finally, Lig I-FEN1, in S-phase, or Lig III α -XRCC1 throughout the cell cycle are recruited to seal the phosphate backbone and restore the integrity of the DNA.

1.3.3 Mismatch repair

Correct structure of the helical DNA and maintenance of genetic integrity is dependent on the fidelity of DNA replication, and must follow the Watson-Crick base pairing sequence. This means that Guanine must always pair with a Cytosine and Adenine must be paired to a Thymine. Mismatch repair is responsible for scanning and maintaining this sequence by correcting mismatch base substitutions and insertion-deletion mismatches.

Undamaged DNA is generally replicated with great efficiency and accuracy, due to high nucleotide selectivity of DNA polymerases and their ability to proof-read (a process that enables the polymerase to identify an incorrect base pair, reverse its direction by one base pair of DNA and excise the mismatched base) during replication. However, polymerases required for translesion synthesis increase the frequency of mismatch bases (Sale et al., 2012) and microsatellite lesions can cause polymerase slippage leading to insertion and deletion loops. It is not surprising then that defects in polymerase proof-reading activity and defects in mismatch repair proteins results in an increase in mutation rates, microsatellite instability and disease (Goellner et al., 1997; Kunkel, 2003; Lange et al., 2011; B. Liu et al., 1995; 1996).

Helical distortions from mismatched bases are detected by one of the two major heterodimeric complexes MutS (made up of MutS α and MutS β). Identification and binding of MutS drives the recruitment of multiple MutL complexes, consisting of MutL α , MutL β and MutL γ , to the mismatch (Kunkel and Erie, 2005). Upon activation by binding PCNA and replication factor C MutL's endonuclease activity generates a nick in the DNA backbone (Kunkel and Erie, 2005; Modrich, 1994). Exonuclease 1 can then remove a stretch of nucleotides containing the mismatch and is thought to stop, either on contact with an okazaki fragment (Genschel et al., 2002) or when it encounters a second nick, created by MutL (Hombauer et al., 2011). RPA coats the exposed single-strand of DNA until replicative polymerases and the corresponding ligases repair the DNA (Ramilo et al., 2002).

1.3.4 Base excision repair

Base excision repair removes small, non-helix-distorting base lesions; often caused by deamination, oxidation and alkylation (Lindahl and Wood, 1999). The base damage is recognized and removed by a specific mono- or bi- functional glycosylase, resulting in a single-strand DNA break which then converges into the same pathway as single-strand break repair as discussed in **1.3.5**.

Currently there are 11 known DNA glycosylases which fall into one of 6 structural superfamilies (Brooks et al., 2013). Whilst the DNA glycosylases are divided into structural superfamilies and have distinct substrate action, their mode of action is shared throughout. This encompasses a flipping action of the damaged base, resulting in a rotation of the substrate into the DNA glycosylases active pocket. The catalytic mechanism of cleavage is where DNA glycosylases are further divided into mono- and bi- functional enzymes. Once the damaged base has been identified the N-glycosidic bond is cleaved through a nucleophilic attack leaving an abasic site. This glycosylase activity is found in both mono- and bi-functional enzymes. The bi-functional DNA glycosylases also contain a lyase activity that can convert the base lesion into a DNA single-strand break and therefore does not require AP endonuclease activity. There are currently 6 mono-functional DNA glycosylases (UNG, SMUG1, MBD4, TDG, MYH and MPG) and 5 bi-functional (OGG1, NTLH1, NEIL1, NEIL2 and NEIL3) (Jacobs and Schär, 2012).

1.3.5 Single-strand break repair

Although not as toxic as DSBs, single-strand breaks (SSBs) are at least one order of magnitude more frequent than DSBs. SSBs occur as a result of direct damage with reactive oxygen species (degradation of the sugar-phosphate backbone), collision with transcription machinery (such as RNA polymerases), as an intermediate of base excision repair or due to stalled complexes of endogenous enzymes such as ligases and topoisomerase I (Top1). If left unrepaired these SSBs can block transcription, lead to double-strand breaks and can induce cell death (**Figure 1.2**) (Alano et al., 2010; Ljungman et al., 1999; Virág et al., 2013; Yu et al., 2002).

In case of base damage, once repair is initiated by a lesion specific glycosylase followed by removal of the abasic site by apurinic/apyrimidinic endonuclease I (APE1), or to a lesser extent by tyrosyl DNA phosphodiesterase 1 (TDP1), the single-strand break core repair machinery follows four specific stages: Detection, end-processing, gap-filling and ligation (**Figure 1.3**) (Caldecott, 2007; 2003).

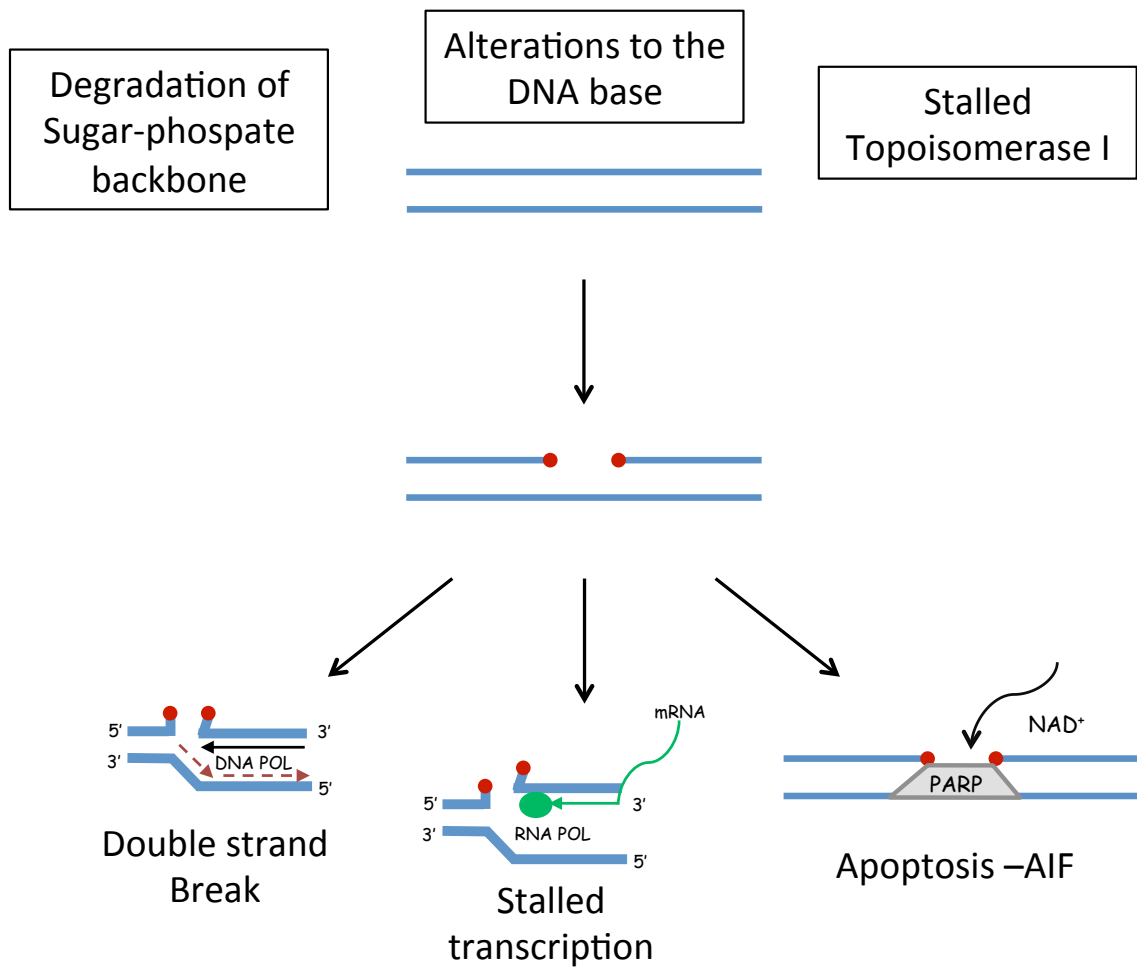


Figure 1.2 Outcome of DNA single-strand breaks if left unrepaired

DNA single-strand breaks can arise from degradation of the sugar-phosphate backbone, alterations to DNA bases as a by-product of BER and through abortive Top1-complexes. Depending on the nature of the damage various chemical modifications can arise at the 3' and 5' end of the DNA break see **Figure 1.5**. If these lesions are not repaired in a timely fashion then this can impact the cell in multiple ways. In replicating cells if the replication machinery encounters the lesion it can lead to replication fork collapse and result in the more toxic double-strand break and cell death. Collision with RNA polymerase and the transcription complex can result in stalled transcription, pre-mature termination of protein synthesis and incorporation of RNA loops. Finally, PARP, the molecular sensor of single-strand breaks, becomes activated when it encounters the break and since this requires NAD⁺ excessive activation can deplete the cell of NAD⁺ leading to apoptosis through release of the mitochondrial apoptosis inducing factor.

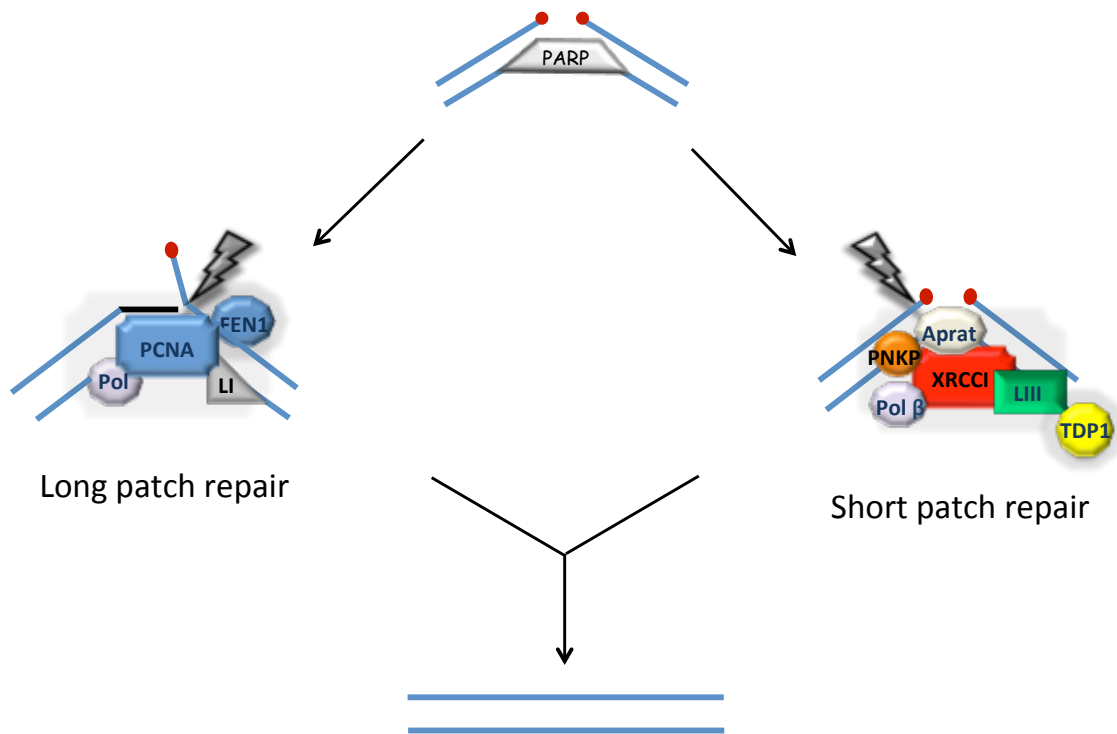


Figure 1.3 Enzymes involved in the repair of DNA single-strand breaks

DNA single-strand break repair can be split into several categories. Firstly, the single-strand break is detected by PARP, which acts as a molecular sensor. Upon interaction with the single-strand break PARP becomes activated resulting in synthesis of Poly (ADP-ribose) chains. These chains are primarily bound to PARP but are also found on other proteins, such as XRCC1. Activation of PARP allows for chromatin remodelling and sequestration of the XRCC1 chaperone protein. XRCC1, the molecular chaperone, interacts with several key proteins in the next stages of repair. End-processing, thought of as the second step in the repair pathway allows specific enzymes, depending on the type of modifications at the breaks termini, to efficiently cleave the modification and restore the required 3'-OH and 5'-P for ligation. The penultimate step is that of gap filling and this can result in the divergence of this particular pathway. If only one nucleotide is required then this divergence leads to short-patch repair where XRCC1 and Lig III α are the main players. However, if more than one nucleotide is required then the long-patch repair pathway is involved in the last step, this incorporates FEN-1 to remove the protruding flap of single-stranded DNA, PCNA and gap filling by Pol β , Pol δ or Pol ϵ before ligation. The last step in the repair pathway is the sealing of the phosphate backbone in an ATP dependent manner; a process carried out by Lig III α (short-patch repair) or Lig I (long-patch repair). Figure adapted from (Caldecott, 2007).

1.3.5.1 Detection: PARP

Initially SSBs are rapidly detected by Poly(ADP-ribose) polymerase (PARP) resulting in activation and synthesis of Poly(ADP-ribose) chains in a nicotinamide adenine dinucleotide (NAD⁺) dependent manner. These chains are primarily bound to PARP but are also found on other proteins, such as XRCC1. The major source of PAR synthesis is carried out by PARP-1, an abundant and stable component of chromatin, and to a lesser extent PARP-2 (Fisher et al., 2007). PARP activation has multiple roles including chromatin remodelling via PARylation of histones (Beneke, 2012), and generation of ATP for ligation. One main function is the sequestration of X-ray repair cross-complementing protein 1 (XRCC1) to the site of damage through binding of the central breast cancer susceptibility protein 1 domain (BRCT1) (Masson et al., 1998; Plo, 2003; Schreiber et al., 2002). The binding of PARP is a transient process and dissociation of PARP occurs through charge repulsion and catabolism of the chains by Poly (ADP-ribose) glycohydrolase (PARG) to allow access to the break for the rest of the repair proteins. XRCC1 is a 69 kDa scaffold protein, that interacts with the nicked DNA and polymerase β through its N-terminal domain (Caldecott et al., 1996), Ligase III α through a C-terminal BRCT2 domain, polynucleotide kinase phosphatase (PNKP), and several end processing factors depending on the type and location (3'- or 5'-) of the damage.

1.3.5.2 End-processing

The second step is coined “end-processing” and due to the numerous abnormal 5'- and 3'-DNA termini caused by damage (**Figure 1.4 and 1.4**) the cell has evolved several key end-processing enzymes to restore the 3'-hydroxyl and 5'-phosphate required for ligation. Several of the autosomal recessive ataxias are involved within the end processing stage of the repair, including tyrosyl DNA phosphodiesterase 1 (TDP1), aprataxin (APTX) and PNKP.

1.3.5.3 Gap filling

The third step involved in the SSB repair is gap filling which is primarily carried out by Pol β to fill in the missing nucleotide. However, instances where more than 1 nucleotide gap filling have been demonstrated. This is a sub-pathway to base excision and single-strand repair, denoted long-patch repair. Long-patch repair seems to function mainly in

S-phase and requires removal of the protruding flap by flap endonuclease 1 (FEN-1), PCNA and gap filling by Pol β , Pol δ or Pol ϵ .

1.3.5.4 Ligation

The fourth and final step requires the sealing of the DNA by DNA Ligase III α or DNA Ligase I for short patch and long patch repair respectively (Caldecott, 2007; Dianov and Parsons, 2007), although redundancy between these ligases has been demonstrated (Arakawa et al., 2012; Simsek et al., 2011).

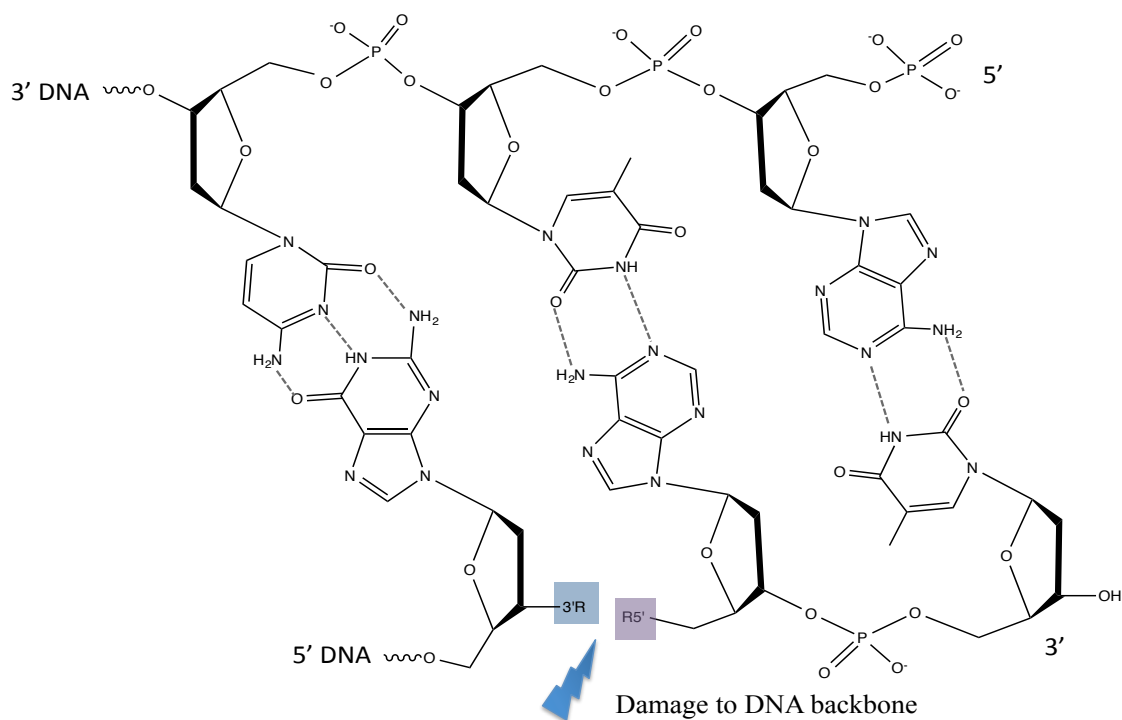


Figure 1.4 DNA structure and damage modifying 3' and 5' phosphate backbone of DNA

DNA is made up of a phosphate backbone, pentose sugar and base (guanine, cytosine, adenine and thymidine). Chemical instability leading to deamination, depurination, depyrimidination or hydrolysis of the phosphodiester backbone as well as reactive oxygen species and intermediates of DNA repair pathways can lead to single-strand breaks. These breaks can lead to different chemical modifications at the 5' (lilac) or 3' (blue) of the phosphodiester backbone.

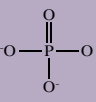
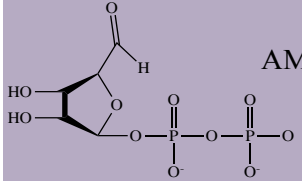
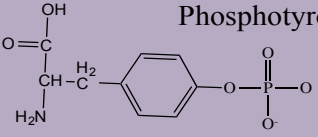
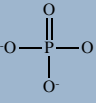
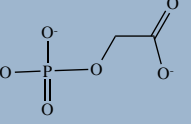
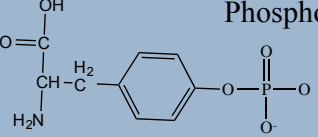
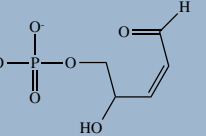
Chemical modification	Damage	Enzyme
 Phosphate	<ul style="list-style-type: none"> • Undamaged • Correct termini for ligation 	<ul style="list-style-type: none"> • Ligase – restores DNA backbone
OH Hydroxyl	<ul style="list-style-type: none"> • Abortive Top1 complexes 	<ul style="list-style-type: none"> • PNKP - kinase activity
 AMP	<ul style="list-style-type: none"> • Abortive Ligase activity 	<ul style="list-style-type: none"> • Aprataxin
 Phosphotyrosine	<ul style="list-style-type: none"> • Abortive Top2 complexes 	<ul style="list-style-type: none"> • TDP2
OH Hydroxyl	<ul style="list-style-type: none"> • Undamaged • Correct termini for ligation 	<ul style="list-style-type: none"> • Ligase – restores DNA backbone
 Phosphate	<ul style="list-style-type: none"> • Abortive Top1 • ROS • Base excision repair 	<ul style="list-style-type: none"> • PNKP phosphatase activity
 Phosphoglycolate	<ul style="list-style-type: none"> • Sugar disintegration (ROS) 	<ul style="list-style-type: none"> • APE1, TDP1
 Phosphotyrosine	<ul style="list-style-type: none"> • Abortive Top1 complexes 	<ul style="list-style-type: none"> • TDP1
 α,β unsaturated aldehyde	<ul style="list-style-type: none"> • Base excision repair (ROS, Alkylators) 	<ul style="list-style-type: none"> • APE1

Figure 1.5 Typical modifications at DNA breaks, cause and enzymes involved in repair

Highlighted in lilac or blue are common chemical modifications occurring at the 5' or 3' end of the phosphate backbone of the damaged DNA, respectively. The required chemical ends needed to seal the backbone and restore the genetic integrity of the DNA are the 5' phosphate and 3' hydroxyl. Figure adapted from (Caldecott, 2008).

1.3.6 Ligase family

The ligase family are involved in the last step of sealing the DNA phosphate backbone in an ATP dependent manner. Whether required for single-strand break repair (Lig III α , Lig I), mitochondrial repair (mito-Lig III α), sealing of okazaki fragments (Lig I) or double-strand breaks (Lig IV). Lig III α has already been shown to interact with TDP1 (Chiang et al., 2010; Das et al., 2009; El-Khamisy et al., 2005) and has since been thought of as the main candidate for the final step in TDP1-mediated repair. There are 3 Ligases found; Lig I, Lig IV and the evolutionary newer Lig III, which has two splice variants that are involved in nuclear or mitochondrial repair (**Figure 1.6**). They all contain a DNA-binding domain (DBD), adenylation domain (AdD) and an oligonucleotide-binding domain (OBD) that encircles the break in a ring-shaped architecture (Pascal et al., 2004).

The DBD binds to the minor groove of DNA interacting with the phosphodiester backbone on both sides of the DNA break. This alters the curvature of the DNA in an underwound conformation and positions the DNA, exposing the nicked ends to residues that are important for the catalytic activity of the enzyme. The DBD interacts with the catalytic core of the ligase, consisting of the AdD and OBD, which encircles the DNA and enhances the stimulation of the end-joining activity of the Ligase. The catalytic core forms a covalent interaction with ATP through a lysine residue, forming a Lys-AMP and releasing an inorganic pyrophosphate. The DNA 5' phosphate then attracts the Lys-AMP forming a DNA-AMP intermediate. Finally, the orientation of the ligatable ends allows the 3'-OH to displace the AMP intermediate, resulting in covalent joining of the nicked end and restoration of the DNA (Ellenberger and Tomkinson, 2008).

Homologs of DNA Lig I and IV are found in all eukaryotes, whereas Lig III is only found in vertebrates and is thought to be an evolutionarily newer ligase.

1.3.6.1 Ligase I

As mentioned previously Ligase I is found in all eukaryotes. Whilst a mitochondrial variant is found in yeast, only a nuclear enzyme is present in vertebrates. Human Lig I contains a replication factory targeting sequence in which it interacts with PCNA

through a PCNA interacting protein motif (PIP motif) (Levin et al., 1997; W. Song et al., 2009). Since the PCNA sliding clamp is involved in replication it makes Lig I the most suitable ligase to join okazaki fragments. It is therefore not surprising that ligase I deficiency results in an increase in okazaki fragments, although does not seem to affect cell proliferation rates (Arakawa et al., 2012; Henderson et al., 1985). The N-terminal domain also contains several phosphorylation sites which are substrates for cyclin-dependent kinase. Phosphorylation of residues increases through the cell cycle reaching a peak as a hyper-phosphorylated ligase at the G2/M. Interestingly, Lig I is dephosphorylated in response to DNA damage and replication and only casein kinase II (CK2) phosphorylation at serine 66 is required for activation of the ligase activity, due to its importance in the ligase-AMP formation (Ferrari, 2003; Prigent et al., 1992).

1.3.6.2 Ligase III

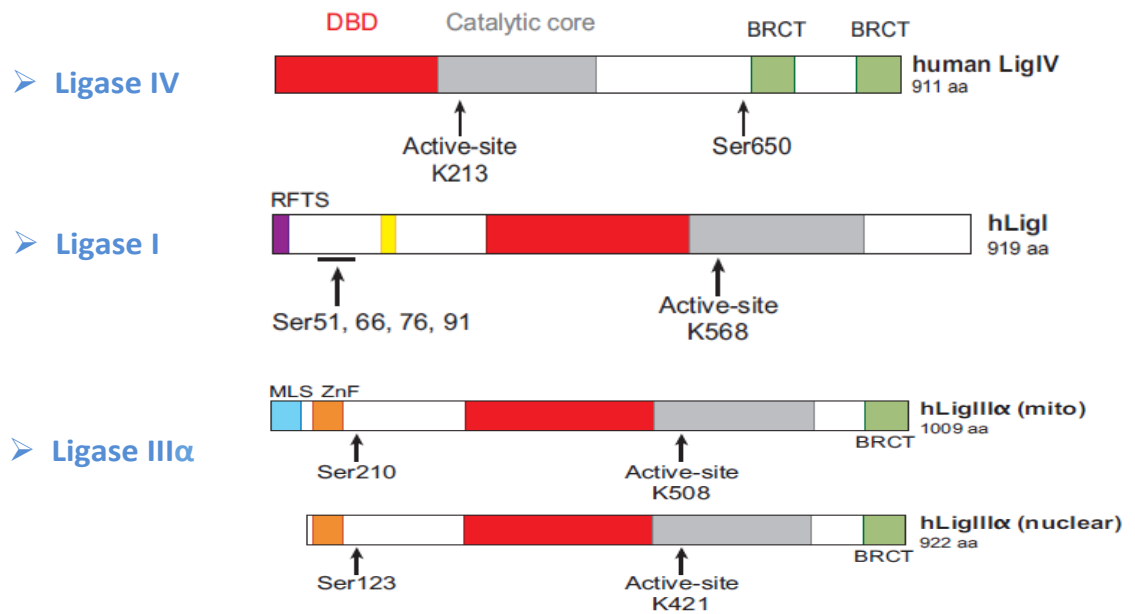
Ligase III is only found in vertebrates and has two forms, α and β , generated by alternative splicing. Lig III β is germ-cell specific and does not contain the C-terminal BRCT domain found in Lig III α . In addition, both of these splice variants also contain mitochondrial and nuclear versions; these are synthesised by alternative translation initiation mechanisms (Lakshmipathy and Campbell, 1999).

The BRCT domain, known for its role in protein - protein interactions, found in Lig III α is required for the interaction with XRCC1 in the nuclear version. This forms a nuclear Lig III α -XRCC1 complex and is required to stabilise nuclear Lig III α . This is supported by the fact that Chinese hamster ovary cell lines EM9 and EMC11 (XRCC1 deficient cell lines) have reduced levels of DNA ligase III α (Caldecott et al., 1994; 1995; Nash et al., 1997).

One other unique feature of Lig III is the zinc finger forming part of the DNA binding domain. This zinc finger is proposed to sense nicks and gaps in the DNA backbone and has been described in a jack-knife model. Lig III α is important for short-patch repair and has been implicated in alternative non-homologous end-joining. Whilst the nuclear version is not critical for cell survival, since Lig III is the only known ligase involved in mitochondrial repair, loss of the mitochondrial Lig III α results in cell lethality (Gao et al., 2011; Lakshmipathy and Campbell, 2001; Simsek et al., 2011).

1.3.6.3 Ligase IV

Unlike Lig I and III, Lig IV has only been identified to play a role in one repair pathway; the repair of double-strand breaks by non-homologous end-joining and in certain cells (B cells) V(D)J recombination. This is owing to the fact Lig IV has a propensity to ligate ends that are non-complementary. Similar to Lig III α Lig IV relies on an interaction with XRCC4 for stability. Inactivation of the mouse Ligase IV gene results in embryonic lethality and inactivation in human cell lines results in hypersensitivity to IR, due to defects in NHEJ (Barnes et al., 1998). Since homologous recombination cannot take place in post-mitotic cells it is not surprising that the developing Lig IV null embryos appear to die because of massive apoptosis in the central nervous system (Frank et al., 1998; 2000). Since Lig IV leads to embryonic lethality only hypomorphic mutations have been identified in patients, leading to microcephaly and immunodeficiency (O'Driscoll et al., 2001).



RFTS – Replication focus targeting sequence domain

MLS – Mitochondrial leader sequence

ZnF – Zinc finger domain

BRCT – BRCA1 C terminus domain

Figure 1.6 Ligase Family

The ligase family all contain a DNA binding domain and catalytic core. For stability Lig IV and IIIα (nuclear) require interactions with molecular chaperone complexes XRCC4 and XRCC1, respectively and these interactions occur through their BRCT domain. Lig I contains a replication factory targeting sequence (RFTS) that is required for its interaction with PCNA. Unique to Lig IIIα is the zinc finger domain (ZnF) used to sense gaps and nick in the DNA backbone. As opposed to non-vertebrates Lig IIIα and not Lig I has variant containing a mitochondrial leader sequence (MLS) that allows it to be targeted to the mitochondria and is essential for cell survival. Figure adapted from (Ellenberger and Tomkinson, 2008).

1.3.7 Protein-linked DNA single-strand breaks

DNA Topoisomerases are a class of enzymes that release helical tension during transcription and replication, by creating nicks within the phosphate backbone on one or both strands of the DNA (Bermejo et al., 2007; Champoux, 2001; Deweese et al., 2008). DNA topoisomerase I (Top1) has also been shown to remove RNA bases from the DNA containing strands preventing formation of R loops, which comprise of RNA/DNA hybrids, that can lead to deleterious effects if unresolved (Tuduri et al., 2009). The release of positive and negative supercoiling by the Top family is a transient process with the equilibrium strongly favouring the sealing of the DNA backbone. This is to help limit the time in which the integrity of the DNA may be compromised. However, the topoisomerase DNA complex can become trapped forming a stalled Top-DNA adduct, otherwise known as an abortive Top complex. Abortive activity of Top1, due to collision with replication/transcription machinery, or as a result of DNA damage within close proximity to the enzyme can result in clastogenic and/or lethal DNA damage. Top1 becomes trapped through a covalently linked phosphodiester bond with the 3'-terminus of DNA. Nucleolytic cleavage of the DNA by nucleases such as the MRN complex, ARTEMIS and CtIP, amongst others, have been shown to remove the stalled Top1 adducts, but this action has only been shown to be in effect on Top1 linked double-strand breaks (Connelly and Leach, 2004; Eid et al., 2010; Nakamura et al., 2010; Pommier et al., 2006; Shrivastav et al., 2008). In contrast, TDP1, a member of the phospholipase D superfamily has been shown to hydrolyse the abortive Top1 complex after degradation of the topoisomerase by the ubiquitin proteasome.

1.4 Tyrosyl DNA phosphodiesterase 1

TDP1 was first reported in yeast and the human homologue was later reported (El-Khamisy et al., 2005; S. W. Yang et al., 1996). A defect in human TDP1 was shown to cause accumulation of Top1-SSBs (El-Khamisy et al., 2005). TDP1 was later shown to be localised not just to the nucleus but also to the mitochondria (Das et al., 2010; Fam et al., 2013; Huang et al., 2013) and protects cells from a broad spectrum of DNA damage (Alagoz et al., 2013; El-Khamisy et al., 2005; 2007; Interthal et al., 2005a; Murai et al.,

2012). TDP1's primary function is to resolve abortive Top1 activity by hydrolysing the phosphodiester bond between the tyrosine residue of Top1 and the 3'-phosphate group (Davies et al., 2002a; 2002b; Murai et al., 2012). Together with TDP2, which hydrolyses the corresponding activity at the 5'-terminus (Cortes Ledesma et al., 2009; Zeng et al., 2011), are currently the only known human enzymes that display this activity in cells.

In neurons, reactive oxygen species (ROS), collision with RNA polymerases and discontinuities through an inherent instability within the DNA have been shown to cause the formation of these stalled Top1 adducts. With the high oxidative load in the brain and high rates of transcription it is therefore logical that these trapped protein-DNA links arise endogenously within cells. If left unrepaired these adducts have been shown to inhibit transcription by blocking the progression of RNA polymerase, and can lead to over activation of PARP depleting cellular levels of NAD⁺ and adenosine triphosphate (ATP), both of which could lead to neuronal cell death (**Figure 1.7**). Once the topoisomerase has been trapped on the DNA due to collision of the replication machinery or caused by damage within close proximity, the first step for removal is its degradation by the proteasome. Once degraded a small peptide fragment remains covalently attached to the 3'-phosphate. SUMOylated TDP1 is then recruited to the site of damage to hydrolyse the 3'-phosphotyrosine in a two-step reaction (Raymond et al., 2004); leaving a 3'-phosphate and 5'-hydroxyl (J. J. R. Hudson et al., 2012; Plo, 2003). Because these termini are not compatible with ligation the phosphatase and kinase activity of PNKP is required to form the 3'-hydroxyl and 5'-phosphate needed for ligation. The importance of TDP1 in neurons is typified by a hypomorphic mutation in the protein leading to the autosomal recessive ataxia SCAN1.

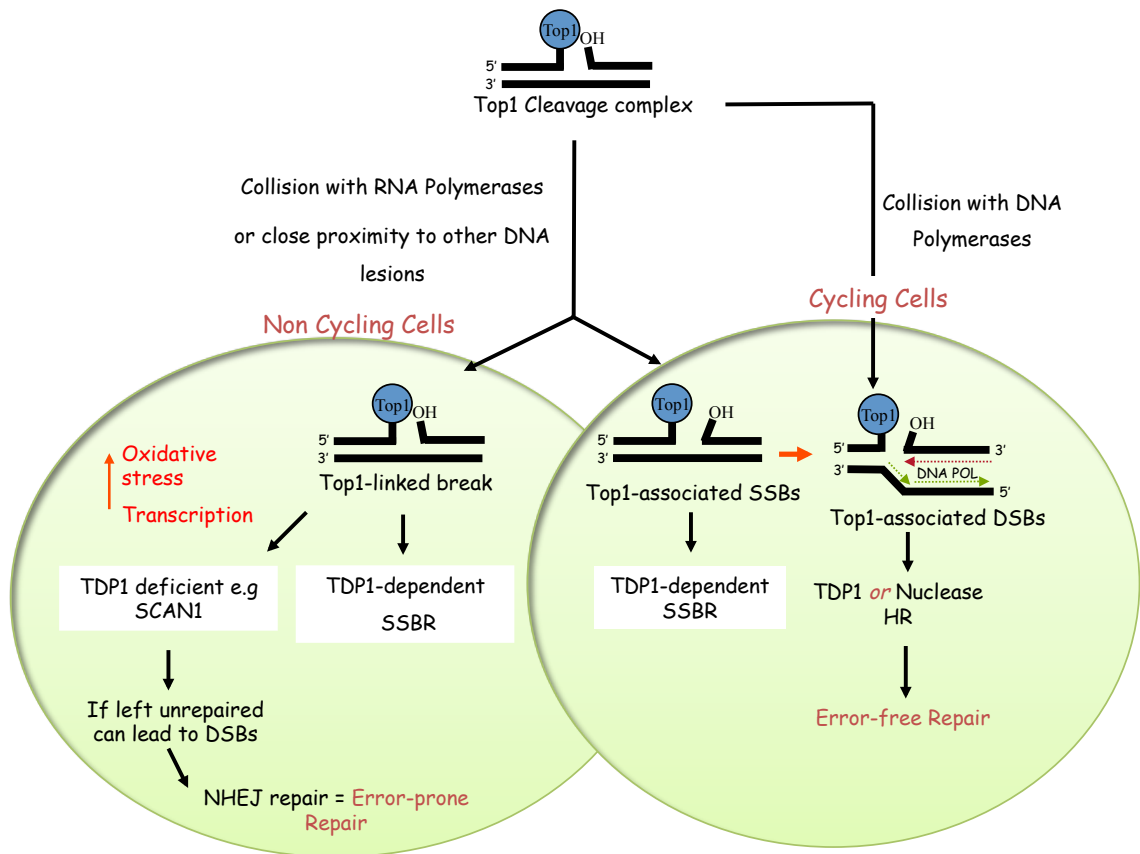


Figure 1.7 Role of TDP1 in cycling and non-cycling cells

Topoisomerase 1 is required to relieve topological strain during transcription and replication. Collision with RNA polymerases or lesions in close proximity to the Top1-DNA complex can lead to an abortive Top1-DNA complex in which the Top1 protein can become stuck to the 3' of the DNA. Normally, if this situation arises Top1 becomes ubiquitinated and degraded by the 26S proteasome, leaving a 3'phosphotyrosine and possible short peptide fragment. TDP1, a phosphodiesterase enzyme, can remove the tyrosine and peptide fragment from the 3'phosphate in a two-step reaction. In addition to this in cycling cells nucleases have also been found to remove this lesion. This then allows PNKP and single-strand break repair machinery to repair the damaged DNA and seal the phosphodiester backbone. If TDP1 is not present or its catalytic activity has been compromised due to a missense mutation, as found in SCAN1 patients, this 3'phosphotyrosine can lead to double-strand breaks, or depletion of NAD⁺ through over-activation of PARP and ultimately cell death. Post-mitotic cells also do not possess the error-free repair found in cycling cells through homologous recombination. Taken together with the high transcriptional activity and oxidative load within the brain the hypomorphic mutation of TDP1 leads to cerebellar degeneration as seen in SCAN-1 but does not seem essential in cycling cells. Figure adapted from (El-Khamisy and Caldecott, 2007).

1.5 Diseases associated with defects in single-strand break repair

Due to the high oxidative load, levels of transcription, and long lived nature of post-mitotic cells in the brain the integrity of DNA is under constant threat. We possess a wide array of end-processing enzymes involved in the removal of lesions of termini of DNA single-strand breaks in order to efficiently restore the required 3' hydroxyl and 5' phosphate needed for successful ligation. It is therefore not surprising that mutations in several of the end-processing enzymes has been linked to neurological disorders, particularly autosomal recessive ataxias. In order to understand why these defects particularly effect the brain we must first understand the neural development of the brain, the requirement for long lived cells and the DNA repair processes needed for cell maintenance.

1.5.1 Neural development

Development of the neural circuit through induction, proliferation, migration and differentiation from progenitor stems cells during embryonic development requires faithful transmission of the genetic code, which is achieved under high replication and transcription stress (Bluml et al., 2012; Caviness et al., 1995; Sidman and Rakic, 1973). In addition, the maintenance of genetic integrity in terminally differentiated post-mitotic neural tissue is constantly challenged by the high level of transcription and metabolic activity, coupled with its limited regenerative capacity (McKinnon, 2009). It is therefore not surprising that defects in the DNA repair have been implicated in a myriad of neurological disorders (Jeppesen et al., 2011; Katyal and McKinnon, 2008; McKinnon, 2009; Yüce and West, 2013). This chapter will review a subset of these disorders termed autosomal recessive ataxias. Examples of this class of disorders covered in this chapter include: ataxia oculomotor apraxia 1 (AOA1), ataxia oculomotor apraxia 2 (AOA2), spinocerebellar ataxia with axonal neuropathy 1 (SCAN1), X-linked mental retardation (XLMR) and microcephaly, infantile-onset seizures, and developmental delay (MCSZ). The neurological component is the most deleterious effect in this particular class of disorders and the above examples show no signs of

immunodeficiency or cancer predisposition. Understanding the biochemistry of how defects in particular proteins impact neuronal integrity and survival will provide us with a solid foundation for design of future therapies, for which suitable animal models will be vital.

Development of the neural circuit occurs through rapid proliferation of neural stem and progenitor cells within the ventricular and sub-ventricular zones (Sidman and Rakic, 1973). These rapidly dividing cells then undergo rounds of differentiation and migrate to specific locations where they mature to form specialized neurons and glia. Inability to maintain genomic integrity during this process will often lead to microcephaly, neurodegeneration or brain tumors (Cox et al., 2006; Jeppesen et al., 2011; Rass et al., 2007; Sailer and Houlden, 2012). During neurogenesis cells containing DNA lesions, single or double-strand breaks are readily replaced, and are therefore thought to favour apoptosis. However, by early childhood the majority of terminally differentiated cells have formed and are not as easily replaced. Within this context, the post-mitotic cells have to function throughout life under high oxidative loads making the DNA repair system extremely important (Fishel et al., 2007; Y. Lee et al., 2009; Y. Lee and McKinnon, 2007). Damage to the DNA can lead to progressive loss of neuronal structure and function resulting in a decline of neurons through apoptosis, thus leading to neurodegeneration, or accumulation of aberrant proteins that have been linked to degenerative diseases such as Alzheimer's and Parkinson's (Kulkarni and Wilson, 2008; McKinnon, 2009). Autosomal recessive ataxias are a class of inherited ataxias in which defects in DNA repair within end-processing enzymes are prominent (**Figure 1.8**).

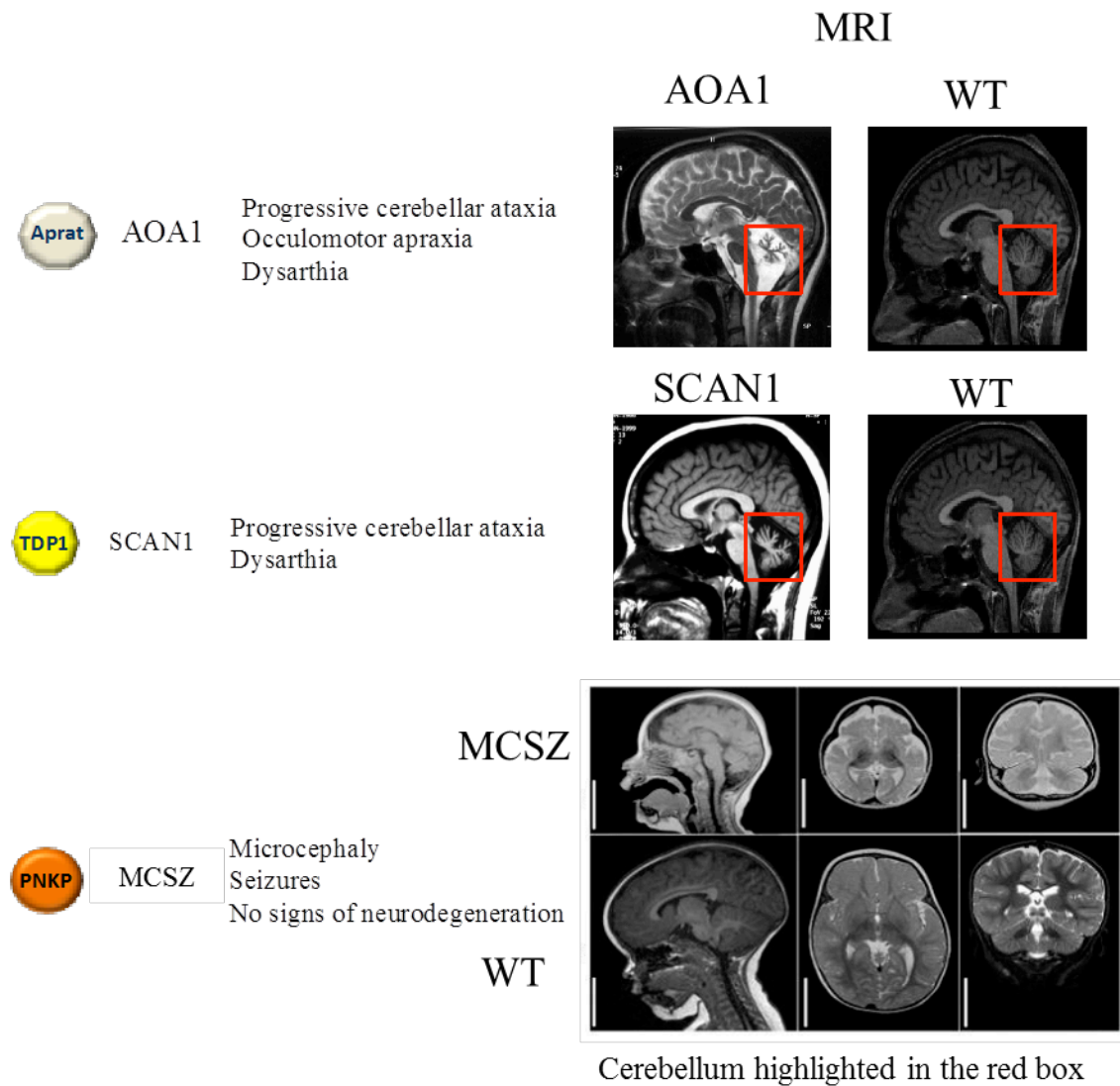


Figure 1.8 Defects in end-processing enzymes and links to disease

The plethora of chemical modifications found at the termini of DNA breaks requires a vast array of end-processing enzymes to efficiently remove the inappropriate modification and allow the restoration of the 3' hydroxyl and 5' phosphate required for ligation. Defects in these enzymes have been linked to several diseases, such as AOA1, SCAN1 and MCSZ. Whilst clinical features of patients can overlap primarily through neuronal degeneration (highlighted in the degeneration of the cerebellum in AOA1 and SCAN1 patients compared to healthy adults (WT)) there are features that do not overlap. AOA1 patients exhibit oculomotor apraxia not seen in SCAN1 patients. A larger contrast is seen when mutations in PNKP, an enzyme involved in TDP1-mediated repair, leads to microcephaly and seizures, but shows no signs of neurodegeneration. MRI images were taken from (Le Ber et al., 2003; J. Shen et al., 2010; Takashima et al., 2002).

1.5.2 Autosomal recessive ataxias

Ataxia-telangiectasia (AT) is the archetypal example and like most other recessive ataxias, AT results in severe degeneration of the cerebellum (Lakin et al., 1996; Y. Lee et al., 2001), a pathological trait it shares with ataxia oculomotor apraxia 1 (AOA1), ataxia oculomotor apraxia 2 (AOA2) and spinocerebellar ataxia with axonal neuropathy (SCAN1). In contrast to these other recessive ataxias AT patients also show a predisposition to cancer (Boultonwood, 2001; Geoffroy-Perez et al., 2001).

Dr Fogel has clinically classified autosomal recessive ataxias into three distinct groups: Friedreich ataxia-like, Friedreich ataxia-like with cerebellar atrophy and early onset ataxia with cerebellar atrophy (Fogel and Perlman, 2007). Why this phenotype is shared amongst these neurological disorders is still under investigation but a unique characteristic of the cerebellum is the extended postnatal development (Y. Lee et al., 2009).

The cerebellum is predominantly composed of three neuronal populations: granule cells, Purkinje cells and interneurons that make up three functional lobes. (1) The anterior lobe (paleocerebellum, or spinocerebellum) receives proprioceptive input from the spinal cord and controls muscles that regulate posture. (2) The posterior lobe (neocerebellum) is involved in the inhibition of involuntary movement by inhibitory neurotransmitters, such as GABA from the Purkinje cells. This plays an important role in fine motor coordination. The last lobe (3) known as the flocculonodular lobe, consists of the floccle and is involved in the maintenance of equilibrium, controlling balance and spatial orientation (Apps and Garwicz, 2005; Hibi and Shimizu, 2012).

The hallmarks of these congenital ataxias are early onset cerebellar ataxia and this progressive degeneration prevents the cerebellum from integrating the sensory and visual inputs. Unsurprisingly, patients show loss of proprioception, vibration sense, balance and gait. Some manifestations also include hypotonia (decreased muscle tone) and dysarthria (problems with speech) (Espinós-Armero et al., 2005; Fogel and Perlman, 2007). Whilst exogenous sources of damage have yet to be implicated in these

diseases there are many endogenous sources that can compromise the integrity of the genome during neural development and within the terminally differentiated cells.

1.5.3 Spinocerebellar ataxia with axonal neuropathy 1 (SCAN1)

SCAN1 is a rare form of ataxia diagnosed in a Saudi Arabian family caused by a missense mutation in the TDP1 gene (c.1478A>G) (Hirano et al., 2007; Takashima et al., 2002). This results in a histidine to an arginine substitution at position 493 (H493R) disrupting the active site and reducing the enzymatic activity by up to 25 fold. TDP1 H493R can also become trapped to the 3'-phosphatate of DNA after releasing the Top1 complex and this is proposed to be more toxic than an inactive variant (Hirano et al., 2007). The neuropathology of SCAN1 is progressive cerebellar atrophy, pes cavus, peripheral neuropathy and muscular atrophy. The clinical symptoms develop at a mean age of 15 years resulting in ataxic gait, seizures and senescence of pain, touch, and vibration senses within the extremities. Patients tend to be wheel chair bound by early adulthood. At a cellular level, patients exhibit mild hypercholesterolemia and hypoalbuminemia (Takashima et al., 2002).

To test the hypothesis that neuronal degeneration was a result of the TDP1 defects, several labs created TDP1 knockout mice. Much like cells from SCAN1 patients, Tdp1^{-/-} mice cells show a hypersensitivity to camptothecin. The neuropathology of Tdp1^{-/-} mice showed very mild age-dependent cerebellar degeneration, however, it did not fully recapitulate that of the human disorder (Katyal et al., 2007). These observations could have several possible explanations. Mice may have a redundant pathway to deal with stalled Top1 complexes not found in humans. The oxidative load in mouse neuronal cells could be considerably less than that of humans; therefore, with the average life span of mice being 2 years the accumulation of SSBs could be below the threshold to see significant cerebellar atrophy. A third possibility is that the H493R mutation is actually more toxic to the cell. Although this also seems unlikely since the Tdp1 H493R knock-in mice exhibit no or very mild phenotype (Boerkoel, personal communication with El-Khamisy). It is probable that a different explanation is more likely to explain the specific vulnerability of the human cerebellum and that degeneration is likely caused by defects in nuclear and mitochondrial repair (Fam et al., 2012).

1.5.4 Microcephaly, infantile-onset seizures, and developmental delay (MCSZ)

As previously discussed polynucleotide kinase phosphatase is a bi-functional enzyme involved in the phosphorylation of 3'-termini of DNA and de-phosphorylation of the 5'-phosphate; both intermediates in the repair of abortive Top1 complexes by TDP1. 3'-phosphates are also the most common form of DNA damage created by ROS. The mode of action allows the correct termini for ligation in the last step of the repair pathway. Mutations in PNKP have recently been established in several patients from Arabic and European origin. This recessive disorder features microcephaly, infantile on-set seizures and developmental delay (J. Shen et al., 2010). MCSZ patients do not feature an increased incidence to cancer or immunodeficiency.

At a cellular level all four mutations found in the PNKP gene resulted in lower cellular protein levels. Whilst two of the mutations also ablate cellular PNKP DNA 5'-kinase activity, reduction of 3'-phosphatase activity has yet to be demonstrated (Reynolds et al., 2012). As this enzyme is featured in the repair of Top1 linked breaks it is unsurprising that patient cell lines and A549 cells depleted for PNKP show hypersensitivity to camptothecin. Resistance of these cells to hydrogen peroxide damage is significantly lower than that of normal cells. Whilst there are some similarities between the neuropathology of MCSZ and SCAN1 patients (ultimate loss of neuronal cells) there is a striking contrast in neuronal phenotype. MCSZ patients exhibit microcephaly due to proliferation defects or cell loss during neurogenesis (J. Shen et al., 2010). PNKP must therefore have a more pivotal role during development and its interaction with XRCC1 and XRCC4 point to multiple roles in single and double-strand break repair. Interestingly defects in ligase IV, a known protein interaction partner to XRCC4, also results in widespread apoptosis during neurogenesis, resulting in microcephaly (Barnes et al., 1998; O'Driscoll et al., 2001). Whether these features would recapitulate in animal models is yet to be tested.

1.5.5 X-linked mental retardation (XLMR)

Cullins are a family of E3 ubiquitin ligases, containing a RING finger domain. These E3 enzymes are part of the hierarchical ubiquitin pathway and play role in the control of a variety of processes (H. C. Liu et al., 2012; Sarikas et al., 2011). Of interest, the cullin4 family (CUL4) consists of two paralogs, CUL4A and CUL4B that share a high sequence homology and can interact with the same protein substrates, indicating a redundancy. Interestingly, only mutations in the CUL4B protein have been linked to XLMR despite the ability to target the same substrates (Kerzendorfer et al., 2010). Topoisomerase 1 is one of the many substrates of CUL4B and this demonstrates a link between TDP1 and CUL4B for the repair of Top1 linked DNA breaks. As mentioned earlier, before removal of the 3'-phosphotyrosine by TDP1, the Top1 is degraded via the ubiquitin proteasome pathway. CUL4B patient cells and A549 cells with CUL4B knockdown show sensitivity to the Top1 poison camptothecin and increased formation of Top1 cleavage complexes (Kerzendorfer et al., 2010). Both sets of evidence pointed to CUL4B acting upstream of TDP1 to target Top1 for degradation. Although both TDP1 and CUL4B deficiency result in neural dysfunction, the cause of neural cell loss is likely to be fundamentally different. XLMR patients with CUL4B mutations are associated with mental and growth retardation, relative microcephaly and motor neuron impairment. There are no signs of cerebellar atrophy, as seen in SCAN1, but development of gait ataxia has been observed (Isidor et al., 2010; Zou et al., 2007). The substrate overlap between the two CUL4 paralogs could account for lack of cerebellar atrophy seen in XLMR patients. It may well be that CUL4A can target Top1 to the proteasome, but at a much slower rate.

Mutations within CUL4A have not been linked to XLMR and in support of this Cul4a deficient mice showed no phenotypic attributes similar to that of patients with XLMR. The Cul4a^{-/-} male mice were however sterile, and they also exhibited a stronger DNA damage response (L. Liu et al., 2009). Initially, early depletion of Cul4b mice caused embryonic lethality and therefore comparisons between XLMR patients and mouse models could not be observed. Recent work by two groups using a Cre/lox strategy enabled the creation of the Cul4b knockout mouse. Cul4b has been shown to be

important during mouse embryogenesis and although Cul4b mutant mice did not show any difference in the phenotypic attributes compared to the wild type mice, they displayed some XLMR-like behavioural phenotypes, linked to hippocampus related spatial learning, memory deficiencies, as well as increased epileptic susceptibility (Chen et al., 2012; L. Liu et al., 2012).

1.5.6 Ataxia oculomotor apraxia 1 (AOA1)

Ataxia oculomotor apraxia 1 is an autosomal recessive disorder that is caused by another mutation within an end-processing enzyme. AOA1 accounts for 10% of all recessive ataxias and has been found in patients from Japan, Europe and Africa (Moreira et al., 2001a). Onset of the disease is variable and appears anywhere between 1 and 16 (mean of 5) years of age. AOA1 is clinically represented by early onset gait ataxia, dysarthria, oculomotor apraxia, cerebellar atrophy and late axonal peripheral motor neuropathy. Similar to SCAN1 the clinical biochemistry shows hypoalbuminaemia and hypercholesterolaemia (Date et al., 2001; Moreira et al., 2001b).

Genetically, patients have shown mutations in the aprataxin gene that seem to be concentrated within the histidine triad domain. Functionally the full-length aprataxin consists of 3 domains; a forkhead associated domain, which shares homology with PNKP and can interact with XRCC4 and XRCC1, a histidine triad (HIT) domain and a DNA binding zinc finger domain. Genotyping has shown that the mutations leading to AOA1 appears to be concentrated to the HIT domain. Whilst aprataxin has been shown to cleave several DNA lesions *in vitro*, the HIT domain has a high activity for 5'-adenosine monophosphate (5'-AMP) termini, making it highly likely that this is its physiological substrate (Ahel et al., 2006; Rass et al., 2007). The last step in strand break repair is the sealing of DNA nicks, a process carried out by Ligase 1, 3 or 4 depending on the repair pathway. The sealing of the breaks requires energy and is dependent on ATP. A 5'-AMP intermediate occurs during this biochemical process and a prediction is that an incomplete ligation cycle, or premature adenylation by the ligase would yield a 5'-AMP. Premature ligation has been supported by co-deletion of aprataxin with a 3'- processing enzyme that results in a synergistic defect in the cells'

ability to repair oxidative damage (El-Khamisy et al., 2009). This modification would require removal before sealing of the DNA breaks can occur.

There is conflicting evidence about how the defects result in neurodegeneration. Post mortem brain sections and fibroblasts from patients exhibited higher levels of oxidative damage, but neither lymphoblastoid cells from patients or mouse APTX^{-/-} astrocytes showed defects in SSB repair (El-Khamisy et al., 2009; Reynolds et al., 2009). This is indicative of the low levels at which these lesions are formed endogenously. Neuronal cell death is probably due to the build-up of these lesions over time. This progressive degeneration could also account for the lack of phenotype seen in APTX^{-/-} mice. More recently, work from the Wilson lab has demonstrated that an isoform of aprataxin with an N-terminal extension localizes to the mitochondria and that these alternate transcripts were found at higher levels within the cerebellum (Sykora et al., 2011). Depletion of aprataxin in human SH-SY5Y neuroblastoma cells and primary skeletal muscle myoblasts resulted in mitochondrial dysfunction and reduced copy number. Furthermore, AOA1 patients shared striking similarities with patients linked to disorders of the mitochondria (Fogel and Perlman, 2007; Stemmler et al., 2010). Both patient types exhibit peripheral neuropathy and progressive ataxia, and at a biochemical level reported deficient coenzyme Q10 (Espinós-Armero et al., 2005), a known mitochondrial antioxidant and key component of the electron transport chain.

1.5.7 Ataxia oculomotor apraxia 2 (AOA2)

Although similar to AOA1, ataxia with oculomotor apraxia type 2 presents with a later onset, between 11 and 22 years of age. Phenotypically AOA2 shares many characteristics with AOA1, including oculomotor apraxia, sensory neuropathy, pes cavus, muscular weakness and cerebellar atrophy (in some patients). Due to the homogeneity of the disorder, apart from genotyping, differential diagnosis can be seen through deep tendon reflexes, which is actually increased compared to the decrease seen in AOA1 patients (Fogel and Perlman, 2007). In contrast to AOA1 laboratory analysis also shows normal albumin levels but high serum levels of the alpha-fetoprotein. The gene of interest has been mapped to c.9q34 and encodes for the protein senataxin. Unlike other recessive disorders, mutations within this protein have also been linked to

the autosomal dominant disorder juvenile amyotrophic lateral sclerosis (ALS4), due to a gain of function (Moreira et al., 2004).

Functionally, senataxin is a putative RNA/DNA helicase that has been shown to play a major role in the removal of R-loops (formation of RNA-DNA hybrids) and replication fork stabilization (Alzu et al., 2012; Suraweera et al., 2009). This is further supported by its direct interaction with RNA polymerase II, and that addition of RNase H1 (an enzyme that resolves DNA/RNA hybrids) or a transcription inhibitor (alpha amanitin) leads to loss of senataxin foci (Yüce and West, 2013). So how does this lead to degeneration of neurons?

Lesions through ROS, or protein-linked DNA breaks (PDBs) would lead to stalled transcription and a requirement for senataxin to resolve the DNA/RNA hybrid. Loss or reduced function of senataxin results in an increase in persistence of R-loops. This in turn can lead to DSBs and blockage of transcriptional progression (Alzu et al., 2012; Yüce and West, 2013). Whilst replicating cells can repair the double-strand breaks through HR, an error free process, neuronal cells (due to the lack of replication) would require NHEJ, increasing the chance for instability. As previously been discussed both of these consequences can result in cell death and evidence of this has previously been reported as senataxin *-/-* cells are sensitive to DNA damaging agents, particularly agents that introduce oxidative stress (Lavin et al., 2008).

1.6 Neurodegeneration: a consequence of repair defect in the nucleus, mitochondria or both?

Whilst defects in the repair of the above proteins can result in risk to the integrity of nuclear DNA, many of these enzymes have been also shown to localise and protect mitochondrial DNA (mtDNA) from damage. Mitochondria are small organelles within the cell that contain their own genetic material. They act as energy producing factories via the production of ATP through the electron transport chain (ETC) and these tiny organelles can produce approximately 65 kg of ATP everyday (Lane, 2006).

Whilst this production of energy is required for survival, a by-product of this process is an abundance of reactive oxygen species. Electrons are leaked through the electron transport machinery leading to production of intermediates that can pose a threat to DNA. Whilst the mitochondria possess a host of antioxidants and enzymes, such as superoxide dismutase (SOD), the location of the mtDNA within the inner mitochondrial membrane and their lack of protection through protective proteins such as histones make them especially susceptible to damage (Gough and Cotter, 2011). In fact it has been shown that there is up to 3-fold more oxidative damage to mtDNA and that once damaged these lesions persist for an extended period of time (E. K. Hudson et al., 1998; Yakes and Van Houten, 1997). Unlike nuclear DNA where approximately 97% of the DNA is non-coding, the majority, 93%, of mtDNA is transcribed and required for producing the components of the ETC (Taylor and Turnbull, 2005).

Due to the size and nature of the mitochondria the repair proteins are actually transcribed in the nucleus and imported to the mitochondria through leader sequences. It is important to note that there is a large portion of mitochondria within the cerebellum and that damage above a certain threshold may lead to induction of cell death through a negative feedback to nuclear DNA damage and apoptosis. If we consider high ROS compromising mtDNA through oxidative damage then the components of the ETC transcribed from mtDNA may lead to improper function. This could lead to an increase in ROS, further damaging mitochondrial and nuclear DNA. It is therefore highly conceivable that mitochondrial damage could contribute to the pathology seen in these disorders. It would be interesting to create nuclear or mitochondrial knock-in proteins of an animal model that shows degeneration of the cerebellum with that particular knockout. Would either of these knock-in models alleviate, or completely remove the pathology seen in these knockouts?

1.7 Aims and objectives

A defect in the catalytic activity of TDP1 leads to accumulation of DNA single-strand breaks in cells (El-Khamisy et al., 2005; Katyal and McKinnon, 2007) and *in vitro* a comparatively high concentration of TDP1 is required to achieve removal of the synthetic lesion (El-Khamisy and Caldecott, 2007; Raymond et al., 2004). So how does TDP1 overcome these issues *in vivo*?

Also in higher eukaryotes TDP1 contains an N-terminal domain of ~ 150 amino acids that has not been crystallised with the catalytic C-terminal domain. Why is this N-terminal domain required in higher eukaryotes and how does this affect the overall structure of TDP1?

The main aim of the thesis is to characterize and identify novel mechanisms for the repair of DNA topoisomerase-mediated and oxidative stress-induced DNA breaks. Using *in vitro* biochemical analysis combined with *in vivo* studies I will investigate how TDP1 is recruited and retained at the sites of damage, through post-translational modifications and protein – protein interactions. I will also study the role for the N-terminal evolutionarily domain of TDP1 and if it is required for optimal protection against genotoxic agents, and what function it may have in the cellular environment.

Chapter II

Materials and Methods

2.1 Materials

Solutions were made up with distilled water unless otherwise stated. Materials and solutions were autoclaved at 125 °C for 15 minutes for sterilisation where possible. Filter sterilisation was carried out through a 0.2 µm filter (Nalgene). Storage was at room temperature unless stated.

LB Broth
(Autoclaved)

LB Agar
(Autoclaved)

TAE (50x)
2 M Tris
50 mM EDTA
5.71 % (1M) Glacial acetic acid

TBE (10x)
0.9 M Tris
0.9 M Boric acid
20 mM EDTA (pH 8.0)

Towbin (10x)
0.25 M Tris
1.92 M Glycine

Protein loading buffer (5x)
250 mM Tris-Hcl (pH 6.8)
500 mM DTT
10 % w/v SDS
0.5 % Bromophenol blue

50 % v/v Glycerol

stored at -20 °C

Tris-glycine Electrophoresis buffer (running buffer)

25 mM Tris

250 mM Glycine

1 % SDS

Western transfer buffer

Towbin (1x)

20 % methanol

DNA loading dye

0.25 % bromophenol blue

40 % sucrose

stored at 4 °C

Phosphate-Buffered Saline (PBS)

(Autoclaved)

SOB medium

0.5 % (w/v) yeast extract

2 % (w/v) tryptone

10 mM NaCl

2.5 mM KCl

20 mM MgSO₄

Transformation buffer

250 mM KCl

15 mM CaCl₂·H₂O

10 mM PIPES

(Sterile filtered) stored at 4 °C

2.2 Cloning and Molecular methods

2.2.1 PCR, restriction digests and ligations

PCR was carried out using FastStart Taq DNA Polymerase (Roche) according to the manufacturers guidelines. Amplified products were purified using the Qiagen PCR purification kit and protocol. Restriction endonuclease digests were performed with enzymes from NEB or Roche, and digests were set up according to the conditions recommended. Typically digests were left for 2.5 hours at 37 °C and Alkaline phosphatase (Roche) was added to the sample containing the vector after 2 hours. This was to prevent re-ligation of the plasmid without the DNA insert. Digests were run on a 1 % agarose gel (1 % w/v in 1X TBE) and the appropriate bands were excised. DNA was then purified using the Qiagen gel extraction kit. Ligations were carried out using T4 DNA ligase (NEB) and samples were left at 16 °C for 3 hours. The ligated products were then transformed into *E.coli* DH5 α cells (section 2.2.3) before being sent for sequencing (GATC).

2.2.2 Site-directed mutagenesis

Site-directed mutagenesis (SDM) reactions were carried out using the QuikChange method (Stratagene) on a plasmid DNA template. Primers homologous to the gene of interest that contained the desired mutation were used to amplify the whole gene containing plasmid by PCR. The reactions, total volume (25 μ l), were then digested with 0.5 μ l of Dpn1 (NEB) and left overnight at 37 °C. This was to ensure the methylated DNA template was fully digested and only the DNA containing the required mutation remained. The plasmid containing the mutation was then transformed into *E.coli* DH5 α cells and sent for sequencing (GATC). Primers used for PCR and SDM's are listed in table 2.2.1.

Table 2.2.1 DNA plasmids created and oligonucleotide sequences

Plasmid	Insert	Species	Cloning sites	Resistance	Oligonucleotides*
pET16b-10HT	TDP1 ν 2	BL21 (DE3), DH5 α	XhoI/BamHI	Amp	
pET16b-10HT	TDP1 S81E	BL21 (DE3), DH5 α	XhoI/BamHI	Amp	
pET16b-10HT	TDP1 Nter (1-150)	BL21 (DE3), DH5 α	BamHI/EcoRI	Amp	N-termTDP1 FWD +REV2
pET16b-10HT	TDP1 Cter (151-608)	BL21 (DE3), DH5 α	BamHI/EcoRI	Amp	C-termTDP1 FWD2 +REV
pET16b-10HT	TDP1 H493N	BL21 (DE3), DH5 α	XhoI/BamHI	Amp	
pET16b-10HT	TDP1 H493R	BL21 (DE3), DH5 α	XhoI/BamHI	Amp	
pET16b-10HT	TDP1 (1-150) K111R	BL21 (DE3), DH5 α	SDM	Amp	N-termK111R FWD +REV
pET16b-10HT	TDP1 (1-150) K139R	BL21 (DE3), DH5 α	SDM	Amp	N-termK139R FWD +REV
pET16b-10HT	TDP1 (1-150) K111R, K139R	BL21 (DE3), DH5 α	SDM	Amp	N-termK111R/139R FWD +REV
pET16b-10HT	TDP1 SIM2	BL21 (DE3), DH5 α	SDM	Amp	SIM2FWD +REV
pET16b-10HT	TDP1 SIM3	BL21 (DE3), DH5 α	SDM	Amp	SIM3FWD +REV
pET16b-10HT	TDP1 SIM4	BL21 (DE3), DH5 α	SDM	Amp	SIM4FWD +REV
pET16b-10HT	TDP1 SIM5	BL21 (DE3), DH5 α	SDM	Amp	SIM5FWD +REV
pCI-myc-puro	TDP1 H493R	DH5 α	SDM	Amp/puro	PCI493R FWD +REV
pCI-myc-puro	TDP1 (281-608)	DH5 α	XhoI/NotI	Amp/puro	281 FWD + PCI CTDP1 REV
pCI-myc-puro	TDP1 (368-608)	DH5 α	XhoI/NotI	Amp/puro	368 FWD + PCI CTDP1 REV
pCI-myc-puro	TDP1 (150-535)	DH5 α	XhoI/NotI	Amp/puro	PCI NTDP1 FWD + 535 REV
pCI-myc-puro	TDP1 (150-416)	DH5 α	XhoI/NotI	Amp/puro	PCI NTDP1 FWD + 416 REV
pFlag-CMV	Lig1	DH5 α	NotI/EcoRV	Amp/neo	
pCI-myc-puro	TDP1 K111R	DH5 α	SDM	Amp/puro	N-termK111R FWD +REV

SDM = Site-directed mutagenesis

Oligonucleotides*	Sequence
N-termTDP1 FWD	CGAGGATCCGATGTCTCAGGAAG
N-termTDP1 REV2	GCCGAATTCTCACTCCCCTGATGTCTCATACTC
C-termTDP1 FWD2	GCCGGATCCGGGCCAGGACATTTGG
C-termTDP1 REV	CAAGAATTCTCAGGAGGGCAC
N-termK111R FWD	GTGGTGATCAGAAAGGA
N-termK111R REV	CTTTCTCCTTTCTGATCA
N-termK139R FWD	CACAGGCTCAGAGAGGA
N-termK139R REV	CTTCCTCCTCTCTGAGCC
SIM2 FWD	CTAGAAGAAGGCGCCGCGTTGTCATACAC
SIM2 REV	GTGTATGACAACCGCGGCGCCTTCTTCATAG
SIM3 FWD	CTCAAGGAGTGGGCGGCTGTTCATTACACAAG
SIM3 REV	CTTGTGAATGACAGCGGCCCACTCCTTGAG
SIM4 FWD	CTCTGAAACAAATGCTGCTCTTATTGGTTC
SIM4 REV	GAACCAATAAGAGCAGCATTTGTTTCAGAG
SIM5 FWD	GGAAAAAGCTCTGTTTCCTGCTGCCTTGATCTATCC
SIM5 REV	GGATAGATCAAGGCAGCAGGAACAGAGCTTTTTC
PCI493R FWD	GCAATGCCATGCCAAGGATTAAGACATATA
PCI493R REV	TATATGTCTTAATCCTTGGCATGGCATTGC
281 FWD	GCCCTCGAGACCTCCAACCTCATCCATGC
368 FWD	GCCCTCGAGGGAAGTCAAAAAGATAAATTGG
PCI CTDP1 REV	GGCGCGCGCTCAGGAGGGCACCCAC
PCI NTDP1 FWD	GCCCTCGAGGGCCAGGACATTTGGGAC
535 REV	GGGCGGCCGCTCAGCGGATCATCAGCTGGGTGC
416 REV	GGGCGGCCGCTCACAGCATGCTCTCTTTAAACTC

2.2.3 Competent cells and transformations

2.2.3.1 *E.coli* DH5 α cells

500 ml of SOB media (section 2.1) was inoculated with a streak of DH5 α cells in a 2 L flask. This was left to grow at 18 °C and 225 rpm to reach an OD₆₀₀ of 0.6 (approx. 35-40 hours). The cells were incubated on ice for 10 minutes then spun in a pre-chilled centrifuge (Sorvall legend RT centrifuge) at 3220 g for 10 minutes at 4 °C. The supernatant was discarded and cells were re-suspended in 160 mls of ice-cold transformation buffer (section 2.1) and placed back on ice for a further 10 minutes. Cells were centrifuged as above and supernatant discarded. Cells were re-suspended in 20 mls of transformation buffer containing 7 % DMSO, left on ice for 10 minutes and then dispensed into pre-chilled Eppendorf tubes. Cells were then frozen in liquid nitrogen and stored at – 80 °C.

2.2.3.2 Transformations

DNA was transformed into DH5 α cells for cloning whilst BL21 (DE3) cells were used for protein expression. Both followed the same transformation protocol described.

50 μ l of cells were transformed with either 1 μ l of miniprep plasmid DNA or 5 μ l of ligation product. Once the DNA had been added to the cells the sample was left on ice for 10 minutes, followed by a heat shock treatment at 42 °C for 45 seconds, then placed back on ice for a further 10 minutes. 100 μ l of LB broth was then added to the cells and left to incubate at 37 °C and 225 rpm for 1 hour. 50 μ l (from miniprep) or 150 μ l (from ligation) was then spread onto LB agar (LB solidified with 1.5 % agar) plates, containing the appropriate antibiotic, and left overnight at 37 °C. Ampicillin or kanamycin was added to the LB agar plates at a final concentration of 50 μ g/ml.

2.3 Electrophoresis and western blot analysis

2.3.1 Electrophoresis of DNA

DNA was typically resolved on a 1 % agarose gel (1 % agarose w/v dissolved in TBE), stained with ethidium bromide (~0.3 µg/ml). The gel was run at 100 V for 40-60 minutes (depending on the resolution needed) in 1X TBE (section 2.1). Samples were loaded in 1X DNA loading buffer (section 2.1) and run alongside a 1 kb or 100 bp ladder. DNA was visualized by UV illumination using the Syngene InGenius Bioimaging system. For gel extraction gels were placed over a UV box and bands excised with a scalpel (section 2.2.1).

2.3.2 Electrophoresis of proteins

Proteins were resolved via Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE gels), alongside precision plus dual colour protein standards (Biorad), using the XCell surelock gel system (Invitrogen). Samples were denatured in a final concentration of 1X protein loading buffer (section 2.1) and incubated at 99 °C for 5 minutes before loading. Gels were run at 190V through a 5 % stacking gel and resolving gel. The stacking gel consisted of 5 % acrylamide from a 30 % acrylamide stock (National Diagnostics), 0.125 M Tris pH 6.8, 0.1 % (w/v) SDS, 0.1 % (w/v) ammonium persulphate and 0.1 % (v/v) TEMED). The resolving gels were made up with 8, 10, 12 or 14 % acrylamide from a 30 % acrylamide stock (National Diagnostics). The resolving gel also consisted of 0.375 M Tris pH 8.8, 0.1 % (w/v) SDS, 0.1 % (w/v) ammonium persulphate and 0.04 % (v/v) TEMED (N,N,N',N'-Tetramethylethylenediamine). The ammonium persulphate and TEMED were added to mixtures last, to allow polymerisation of the gel (Sambrook and Russell 2001)

Gels were run in 1X running buffer (Materials, section 2.1) and then either transferred for western blot analysis, or stained.

2.3.2.1 Coomassie blue staining

Gels were stained with Coomassie blue solution (50 % methanol, 10 % acetic acid, 0.05 % (w/v) Coomassie blue)

2.3.2.2 Silver staining

Silver staining was carried out using the Biorad Silver Stain Plus kit according to manufacturer's instructions.

2.4 Western blots

Protein samples were resolved on an appropriately sized gel. The gel was then washed with distilled water and transferred to a Hybond-C extra membrane (Amersham biosciences). The membrane was then transferred at 30V for 70 minutes using the XCell II Blot module kit (Invitrogen) in 1X transfer buffer (Materials, section 2.1). Membranes were then blocked for 1 hour in 5 % (w/v) non-fat dried milk (Marvel), dissolved in PBS and 0.1 % (v/v) Tween (PBST). This was followed by incubation with the primary antibody, made up in blocking buffer, for 1 hour at room temperature or 4 °C overnight. Membranes were then washed with PBST for 5 minutes, and this step was repeated 3 times. The appropriate Horseradish peroxidase (HRP)-conjugated secondary antibody (diluted in blocking buffer) was then added to the membranes and left for 1 hour at room temperature. A summary of the antibodies used can be found in table 2.2.2. The membranes were then washed 3 times for 5 minute periods in PBST before detection with addition of ECL chemiluminescent reagents (GE Healthcare) as per manufacturers guidelines. Emission was captured with the autoradiograph film using the Xograph compact 4 automatic X-ray film processor. Quantification of blots was carried out by Gene Tools (SynGene), or ImageQuant software (GE Healthcare).

Table 2.2.2 Antibodies

Name	Source Species	WB (dilution)	MW (kDa)	Source
Actin (AC-40)	Mouse	1 in 2000	44	Sigma (A4700)
APE1 (NB100-101)	Rabbit	1 in 1000	37	Novus Biologicals (NB100-101)
PCNA (PC10)	Mouse	1 in 1000	29	Alan Lehman
Flag	Mouse	1 in 3000	N/A	Sigma (F3165)
His	Mouse	1 in 1000	N/A	Sigma (H1029-0.2mL)
pLigase 1	Mouse	1 in 1000	125	Neomarkers
Ligase 3 (TL25)	Rabbit	1 in 2000	103	Keith Caldecott
Myc (9B11)	Mouse	1 in 2000	N/A	Cell Signaling (2276)
Sumo1 (D-11)	H, M, R	1 in 500 can detect in lysate	11-20	Santa Cruz (sc-5308)
Sumo1	Rabbit	1 in 4000 not for detection in lysate	16-18	Active Motif (40120)
Sumo2/3	Rabbit	1 in 4000 not for detection in lysate	15, 18	Active Motif (40220)
Sumo2/3	Rabbit	1 in 3000	15, 18	Abcam (ab80986)
TDP1	Rabbit	1 in 2000	68	Abcam (ab4166)
TDP1	Rabbit	1 in 500 (detects C-terminus)	68	Eurogentec

2.5 Protein expression and purification

2.5.1 Protein expression

20 mls of LB broth in a 250 ml flask (containing 50 µg/ml of the appropriate antibiotic) was inoculated with a streak of colonies from BL21 (DE3) cells containing the plasmid DNA of interest. The flask was then left overnight at 37 °C and 225 rpm for approximately 18 hours until a 10 fold dilution of the culture (diluted in LB broth) gave an OD₆₀₀ between 0.3 and 0.5. An inoculum volume of 20 mls was then added to 1 L of terrific broth (Melford), containing 50 µg/ml of the appropriate antibiotic and left at 37 °C and 225 rpm until an OD₆₀₀ between 0.9 and 1.0 was reached. The culture was then placed at 20 °C and 225 rpm for 1 hour and IPTG was then added to make a 1 mM final concentration. The culture was placed back in the 20 °C incubator and left overnight. A final OD reading was taken in the morning and cells were spun down at 5,000 rpm for 15 minutes at 4 °C in a pre-chilled centrifuge. The supernatant was then discarded and the cell pellet was re-suspended in 20 mls of buffer (25 mM HEPES and 0.5 M NaCl pH 8.0) and stored at -20 °C

2.5.2 Protein purification

2.5.2.1 Cell lysis

Cell pastes were taken from the -20 °C to defrost. 20 mls of buffer A, (25 mM HEPES, 500 mM NaCl, 10 mM imidazole and 10 % Glycerol at pH 8.0), and 240 µl of 100 mM PMSF (Phenylmethanesulfonyl fluoride) were added to re-suspend the cell pellet. The re-suspended cells were then placed on ice and sonicated at 40 % amplitude for 30 seconds and then left to cool for 30 seconds. This process was repeated a further 6 times. The sample was then spun at 10,000 rpm for 15 minutes @ 4 °C in a Beckman Coulter Avanti J-20XP centrifuge. A 400 µl aliquot was taken from the soluble fraction and labeled the load.

2.5.2.2 Immobilised metal affinity chromatography (IMAC) purification

The supernatant was loaded onto a Ni⁺ charged 1 ml HiTrap column (GE healthcare) attached to an FPLC in the cold cabinet. Before loading the column was washed with distilled water and equilibrated with 5 column volumes of buffer A. The protein sample was loaded onto the column at 1 ml/min and unbound proteins were removed with 10 column volumes of buffer A. The column was then washed with 10 column volumes of 80 mM imidazole (dissolved in buffer A) and 8 column volumes of 250 mM imidazole (dissolved in buffer A). 200 µl of each wash, termed flow through, 80 mM imidazole wash and 250 mM imidazole wash were reserved and a protein gel was run to identify which fractions contained the protein of interest. 25 µl or 75 µl of fractions, depending on UV trace, were added to 20 µl of 5X protein loading buffer and made up to a final volume of 100 µl with column buffer. Samples heated at 80 °C for 5 minutes before 3 – 20 µl, depending on fraction, were loaded onto the gel. Fractions containing the fusion protein were dialysed against 0.5 L of pre-chilled 25 mM HEPES, 130 mM NaCl, 10 % Glycerol, 1 mM DTT pH 8.0 for 1.5 hours at 4 °C. This step was then repeated for a further 1.5 hours in fresh dialysis buffer. The sample was then snap-frozen in liquid nitrogen and stored at -80 °C. The Protein concentration was determined by Nanodrop (Thermo Scientific).

Table 2.2.3 Proteins expressed and purified

Batch	Purification Column	Concentration (mg/ml)	Protein	Molar Concentration (μ M)	Comments
OW110308a	1ml HiTrap HP	0.51	TDP1 S81A	7.15	
OW110308b	1ml HiTrap HP	1.01	TDP1 S81E	14.17	
OW110308C	1ml HiTrap HP	0.78	TDP1 H493N	10.94	
OW111201B	1ml HiTrap HP	1.20	TDP1 v2	16.83	
OW110225	1ml HiTrap HP	0.70	K111R	9.78	
OW110127A	1ml HiTrap HP	0.38	TDP1(1-150) K111R	19.2	N-terminal truncations
OW110127B	1ml HiTrap HP	0.55	TDP1(1-150) K139R	27.79	N-terminal truncations
OW110127C	1ml HiTrap HP	0.80	TDP1(1-150) K111/139R	40.37	N-terminal truncations
OW101201	1ml HiTrap HP	0.53	TDP1(1-150)	26.82	N-terminal truncations
NA101201	1ml HiTrap HP	0.81	TDP1(151-608)	14.76	
OW111021A	1ml HiTrap HP	n/a	TDP1 SIM1 ILLV - AALV	n/a	No expression of protein
OW110811A	1ml HiTrap HP	0.26	TDP1 SIM2 LRVV - AAVV	3.65	Potential contaminant at 65 kDa
OW110811B	1ml HiTrap HP	0.31	TDP1 SIM3 IDVI - AAVI	4.35	Potential contaminant at 65 kDa
OW110815A	1ml HiTrap HP	0.32	TDP1 SIM4 VYLI - AALI	4.49	Potential contaminant at 65 kDa
OW110815B	1ml HiTrap HP	0.28	TDP1 SIM5 LYLI - AALI	3.93	Potential contaminant at 65 kDa
OW111201C	1ml HiTrap HP	2.02	K111R	28.23	

2.6 Biochemical assays

2.6.1 SUMOylation assay

The SUMOylation assay used to identify the covalent interaction of TDP1 and its variants with SUMO 1, 2 and 3 was carried out following the SUMOlink SUMO-1 kit (Activemotif) manual. Purified His-TDP1 or His-TDP1 mutants (430 nM) were incubated with purified SAE1/SAE2 (50 nM), UBE2i (500 nM), and SUMO 1, 2 or 3 (30 μ M) in 50 mM Tris-HCl, pH 8.0; 50 mM NaCl; 5 mM MgCl₂; 10% glycerol; and 0.5 mM DTT, and 5 mM ATP (Activemotif). Reactions were incubated at 30 °C for 3 hours unless otherwise indicated, followed by fractionation on SDS-PAGE and analysis by western blotting using 1:4000 dilution of anti-SUMO1 (Activemotif), or a 1:200 dilution of anti-TDP1 polyclonal antibodies (ab4166; abcam). Quantification of SUMOylated TDP1 was conducted using ECL and ImageQuant LAS4000 software, and normalised to total TDP1 immunoreactive signal.

2.6.2 SUMO and TDP1 non-covalent binding studies

2.6.2.1 Recombinant TDP1 binding to SUMO immobilised resin

5 μ g SUMO 1, 2 or 3 immobilised to beads (Enzo life sciences) were equilibrated by washing 3 times in 200 μ l binding buffer (1 % (v/v) Triton, 300 mM NaCl, 25 mM HEPES and 10 % (v/v) Glycerol pH 8.0). 2.5 μ g of recombinant His-TDP1 and TDP1

truncations were then added to the beads in a 500 µl volume (made up in binding buffer) and left to bind end-over-end for 3 hours at 4 °C. The samples were then spun in a pre-chilled centrifuge at 1,500 rpm, 4 °C for 2 minutes and supernatant was discarded. The beads were washed in 200 µl of binding buffer, re-spun, and supernatant removed. This process was repeated for 4 washes followed by the addition of 20 µl of 2X protein loading buffer. Sample boiled for 15-20 minutes, resolved on an SDS-PAGE gel and transferred for western analysis (see 2.4).

2.6.3 Casein Kinase 2 phosphorylation

3 µg of recombinant His-TDP1 was added to 2 µl of CK2 (NEB) in 10X CK2 reaction buffer (NEB), 400 µM ATP and samples made up to 10 µl in 25 mM HEPES, 130 mM NaCl at pH 8.0. Reactions were left for 2 hours at 30 °C before addition of protein loading buffer to stop the reaction. Varying concentrations of TBB (SIGMA) were pre-incubated with CK2 for 1 hr at 30 °C before addition of recombinant TDP1. Samples were analysed through immunoblots or SDS-PAGE gel.

For incubation with whole cell extracts 3 µg of recombinant His-TDP1 was added to 5 µl of cell extract (HEK 293 cells) with the addition of ATP and magnesium chloride to final concentrations of 5 mM and 2 mM respectively. Samples left to incubate for 30 mins at 30 °C. Samples were fractionated and analysed by immunoblotting (see 2.4).

2.7 Protein structure and folding

2.7.1 Circular dichroism

Recombinant proteins were dialysed in 30 mM sodium phosphate buffer, pH 7.4, and filtered immediately before the spectrum was obtained through 0.2-micron filters (Millipore). The concentration of recombinant proteins was accurately determined immediately before, during and after the CD scans, to ensure accurate secondary structure estimations. Samples were normalised to the lowest concentration (0.6 mg/ml), placed in a 0.2-mm quartz cuvette (Starna, Essex, UK) and measurements taken using a JASCO J-715 spectropolarimeter (JASCO). The CD spectra of the buffer alone was

subtracted from that of the sample, and time constant was set to 4 seconds with a scan rate at 50 nm/min. The bandwidth was 1 nm and the sensitivity set to standard. Scans were performed from 260 to 190 nm with a 0.1 nm data pitch and continuous scan mode. A Peltier device was used to maintain a temperature at 10 °C. The buffer baselines were subtracted and data represent the average of four independent experiments \pm standard error of the mean (s.e.m.)

2.7.1.1 Quantitative analysis of circular dichroism by Dichroweb

Secondary structure analysis was conducted using three different algorithms; the variable selection method (CDSSTR) with reference data set 4, K2D and CONTIN from the Dichroweb server, which gave the best goodness-of-fits.

2.7.2 Protein denaturation curves

45 μ l of 1 μ M protein solution (25 mM HEPES, 130 mM NaCl, 10% Glycerol and 1 mM DTT at pH 8.0) was added to 15 μ l of 20 x SYPRO-Orange; resulting in a final protein concentration of 0.75 μ M. A control of 45 μ l of buffer (25 mM HEPES, 130 mM NaCl, 10% Glycerol and 1 mM DTT at pH 8.0) with 15 μ l of 20 x SYPRO-Orange was also carried out. 20 μ l of each sample was aliquoted in triplicate into a 96 well plate. Protein melting experiments were carried out using the LightCycler 480 System II (Roche). The instrument was set up with a detection format of 465 nm as the excitation wavelength and 580 nm as the emission wavelength to detect SYPRO-Orange-specific signal. Melting curve fluorescent signal was acquired between 20 °C and 70 °C using a ramping rate of 0.03 °C s⁻¹, and an acquisition of 20 data points per degree Celsius.

Analysis

Melting temperatures (T_m) were determined in Prism by using a variation of the Boltzmann model from (Ericsson, Hallberg et al. 2006).

$$Y = (a_n X + b_n) + \frac{(a_d X + b_d) - (a_n X + b_n)}{1 + e^{\frac{T_m - X}{m}}}$$

Where a_n and a_d are the slopes of the native and denatured baselines and b_n and b_d are the y-intercepts of the native and denatured baselines. T_m is the melting temperature; m describes the slope of transition. Data were normalised to a fraction of protein in its denatured state and presented as normalised relative fluorescence \pm standard deviation.

2.8 Mammalian cell culture

2.8.1 Maintenance of cell lines

Human HEK293T cells, were cultured at 37 °C and 5 % CO₂ in Dulbecco's essential medium (Gibco, Invitrogen) supplemented with 10 % foetal calf serum (FCS), 2 mM l-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. Chicken DT40 cells, were cultured at 39 °C and 5 % CO₂ with RPMI 1640 supplemented with 10 % FCS, 1 % chicken serum, 2 mM l-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin with 10⁻⁵ β-Me.

2.8.2 Transfection

2.8.2.1 Calcium phosphate transfection

HEK293T cells from a fully confluent T75 flask were split and diluted 1 in 10 in fresh media (total 10 mls) per 10 cm plate. Plates were placed in the incubator (37 °C and 5 % CO₂) overnight. 2X HBS (section 2.1) was thawed at room temperature and 500 µl aliquots were made per transfection. Into a separate Eppendorf a maximum of 10 µg DNA (irrespective of number of plasmids used per transfection), 61 µl of CaCl₂ (2 M stock) was added and made up to 500 µl with distilled water. Whilst bubbling the 2X HBS the DNA mixture was added drop-wise. The final mixture was then added drop-

wise over the 10 cm plate and the cells were placed back in the incubator and left for 24 hours (precipitate in the plate will be visible through the microscope. The media was then removed, washed with PBS and replaced with fresh media. Cells were then left in the incubator overnight ready for harvest the following day.

2.8.2.2 Transfection by electroporation

Stable cell lines. DT40 Tdp1^{-/-} cells were transfected following method 1: Stable non-targeted transfection of DT40 by Julian E. Sale (Reviews and Protocols in DT40 Research, Subcellular Biochemistry Volume 40; Jean-Marie Buerstedde and Shunichi Takeda). 5 ng of linearized pcD2E vector containing G418 resistance and 15 ng of linearized Myc – tagged hTDP1, c-TDP1 or empty vector were mixed with 5.12×10^6 cells in 0.5 mls ice cold PBS. Samples were then transferred to a Gene Pulser cuvette (Bio-Rad) with a 0.4 cm electrode gap and left for 10 mins at 4 °C. Using the Bio-Rad Gene Pulser samples were electroporated at 550 V and 25 uFD before being chilled on ice for 10 mins. Cells were then transferred to a 10 cm dish containing 20 mls of pre-warmed media and left overnight at 39 °C. Cells were diluted in 80 mls of media supplemented with 2mg/ml G418 for selection and aliquoted into 200 µl samples in 4 x 96 well plates. Cells were left for 7 days until colonies could be seen through the bottom of the wells. Once at approx. 2 mm diameter cells were transferred to 24 well plates with 1.5mls of pre-warmed media and left until full confluent (3 days). Cells were then harvested, samples fractionated on an SDS-PAGE gel and stable clones were detected through immunoblots (see 2.4).

2.9 Cell harvest and lysis

Media was removed and cells washed in ice cold PBS. Lysates were prepared from cells by re-suspension in 200 µl lysis buffer (20 mM HEPES pH 7.4, 0.5% NP40, 40 mM NaCl, 2 mM MgCl₂, 1X protease inhibitor cocktail [Roche], 1X protease inhibitor cocktail [Roche] 20 mM NEM, 25 U/ml benzonase [Merck]) per 10 cm plate. Lysates were incubated for 30 minutes on ice, cleared by centrifugation at 13000 rpm at 4 °C for 10 minutes. Supernatant was collected and the protein concentration determined by Bradford assay. Protein loading buffer was then added to lysate for western analysis, or

the NaCl concentration adjusted to 300 mM for co-immunoprecipitation, to examine protein-protein interaction.

2.10 Co-immunoprecipitation

Anti-Myc antibody (9E11; cell signalling), or Anti-Flag antibody (F3165; Sigma) was added to the lysate (section 2.9) and the mixture was rotated at 4 °C for 1 hour. Following centrifugation at 1500 rpm for 3 minutes, the supernatant was then added to protein G sepharose beads (Sigma), and allowed to bind end-over-end at 4 °C for 2 hours. Beads were then washed 3 times with wash buffer (20 mM HEPES pH 7.4, 300 mM NaCl) and re-suspended in protein loading buffer. Samples were boiled for 15-20 minutes, resolved on an SDS-PAGE gel and transferred for western analysis.

2.11 Creating stable cell lines in DT40 Tdp1^{-/-} cells

DT40 Tdp1^{-/-} cells were complemented with Myc-TDP1, Myc-TDP1¹⁵¹⁻⁶⁰⁸ or Myc vector. 20 µg of protein encoded plasmid linearized along with 20 µg of pcD2E (containing G418 resistance). Samples were purified using ethanol precipitation and the concentration was measured by NanoDrop. 5.12×10^6 cells/transfection were re-suspended in final volume 500 µl of ice cold PBS. 15 µg of DNA containing either the Myc-TDP1, TDP1¹⁵¹⁻⁶⁰⁸ or Myc vector and 5 µg of pcD2E was added to cells and transferred to a Gene Pulser cuvette (biorad) with a 0.4 cm electrode gap. Samples were left on ice for 10 mins before electroporation at 550 V and 25 uFD. Cells were placed on ice for a further 10 mins before addition of 20 mls of pre-warmed culture media to a 10 cm plate. Plates incubated overnight at 39 °C and 5 % CO₂. Samples were made up to 80 mls (total volume) with culture media containing 2 mg/ml G418 and aliquoted in 200 µl fractions to a 96 well plate and left for 7 days until colonies were visible through the bottom of the plate. Wells containing 1 colony were then made up to 3 mls of media with 2 mg/ml G418 and aliquoted into 12 well plates for incubation at 39 °C. When fully confluent (3 days) the samples were processed and immunoblots analysed for clones expressing the protein of interest.

2.12 Viability assays

Measurement of cellular sensitivity to DNA damaging agents. To measure the sensitivity of cells to CPT and MMS in the presence or absence of 0.5 μ M PARP inhibitor (Olaparib) 0.5 ml (1×10^5 cells in 10 ml) cells were mixed with 0.5 ml of media containing various concentrations of the drugs. Five hundred cells/well were seeded in triplicate into 96 well plates with 100 μ l of medium/well. Plates were incubated at 39 °C for 72 hrs and cell viability determined using the CellTiter-Blue kit. 20 μ l of CellTiter-Blue solution added to each well, fluorescence (560(20)Ex/590(10)Em) measured using the GloMax®-Multi Detection System (Promega) and data analysed using SigmaPlot. Survival of untreated cells set to 100 %. Error bars represent standard error from 3 independent repeats.

2.13 FACS analysis

1×10^6 cells for each cell line were collected and pellets obtained by centrifugation at 250 g in a benchtop centrifuge for 5 mins. The supernatant was removed and cells were washed twice with cold PBS. The cells were then re-suspended in 50 μ l PBS and 1 ml of, ice cold, 70 % ethanol was added whilst vortexing. Aliquots were stored at –20 °C overnight. Cells were spun down, washed twice with PBS, and stained with 500 μ l PBS containing 20 μ g/ml propidium iodide and 50 μ g/ml of RNase A. Re-suspended cells were analysed for DNA content on a FACS Canto flow cytometer and data processed with FACS Diva software (Becton Dickinson)

2.14 DNA single-strand break assays

2.14.1 Alkaline comet assay

The alkaline comet assay was used to analyse induction of DNA strand breaks after treatment with various chaotropic agents. The assay conducted in alkaline conditions measures SSBs, alkali labile sites and DSBs, but because of their frequency the assay

primarily measures SSBs. Samples were processed on ice and in the dark to prevent induction of unwanted DNA damage, or repair.

2.14.1.1 Preparation

Frosted slides (Fisher) were coated with 0.6 % (w/v) electrophoresis grade agarose (Invitrogen) dissolved in 1X PBS. Cells were suspended in pre-chilled PBS and mixed with equal volume (150 μ l) of 1.2 % low-gelling-temperature agarose (Sigma, type VII) maintained at 42 °C. Cell suspension was immediately layered onto pre-chilled frosted glass slides (Fisher) pre-coated with 0.6 % agarose and maintained in the dark at 4 °C until set, and for all subsequent steps. Slides were immersed in pre-chilled lysis buffer (2.5 M NaCl, 10 mM Tris-HCl, 100 mM EDTA pH 8.0, 1 % Triton X-100, 1 % DMSO; pH 10) for 1 hr, washed with pre-chilled distilled water (2 x 10 min), and placed for 45 min in pre-chilled alkaline electrophoresis buffer (50 mM NaOH, 1 mM EDTA, 1 % DMSO). Electrophoresis was then conducted at 1 V/cm for 25 min, followed by neutralization in 400 mM Tris-HCl pH 7.0 for 1 hr. Finally, DNA was stained with Sybr Green I (1:10,000 in PBS) for 30 min. Average tail moments from 50 cells/sample were quantified from blindly coded slides using Comet Assay IV software (Perceptive Instruments, UK).

2.14.2 Gyrosyl assay

2.14.2.1 Preparation of oligonucleotide substrates:

Oligonucleotides of 13mer with a 3'phosphotyrosyl bond conjugated with an FITC molecule were purchased from Midland Certified Reagent (Midland, TX, USA).

2.14.2.2 TDP1 fluorescence assay:

The assay was developed in 384 well black plates in a final reaction volume of 15 μ l. Assay buffer was composed of 50 mM tris 8.0, 5 mM MgCl₂, 80 mM KCl, 0.05% Tween 20 and 1mM DTT. 10 μ l of TDP1 was dispensed into each well. Reactions were then initiated by addition of 5 μ l substrate and plates were incubated at 25°C. Final concentrations were 6.25pM of TDP1 and 10nM of substrate. Addition of quench reagent (2 μ l Sensor + 30 μ l Enhancer buffer) (Gyrasol Technologies, Kansas, USA) stopped the reaction and fluorescence was subsequently monitored with excitation and

emission wavelengths of λ_{ex} 490nm and λ_{em} 520nm using a BMG Labtech Pherastar plate reader. *In vitro* 3'-tyrosyl-DNA phosphodiesterase 1 activity of whole cell extracts (WCE) was determined using the TDP1 fluorescence assay format in a 15 μ l reaction volume containing 8 μ l assay buffer, 5 μ L of 30 nM oligonucleotide substrate and 2 μ l of the indicated amounts (μ g) of WCE diluted in lysis buffer.

2.15 Mass spectroscopy

Bands of interest were cut from the gel, using a clean scalpel that was rinsed with methanol and transferred to separate Eppendorfs. Once excised the instant blue stain was removed using a de-staining protocol:

Samples were washed in 300 μ l 50% MeCN at room temperature for 15 mins in a shaking incubator. Supernatant was then removed followed by the addition of 300 μ l 25 mM NH_4HCO_3 , samples were again left to incubate at room temperature for 15 mins. The supernatant was removed and samples were washed with 300 μ l 50% MeCN in 25 mM NH_4HCO_3 and left shaking at room temperature for 15 mins. Samples were dehydrated using a speedvac for 5 mins with no heat. Once dehydrated the samples were rehydrated in 10 mM DTT solution (~50 μ l or enough to cover the gel pieces) and left to incubate for 45 min at 50°C before undergoing alkylation.

Alkylation:

Samples were allowed to return to room temperature before the supernatant was removed and replaced with iodoacetamide solution. Samples were then left to incubate for 45 min in the dark, at room temperature. The supernatant was then removed before samples were again subject to dehydration, i.e. three changes of 50% MeCN, 25 mM NH_4HCO_3 , with 5 min shaking between each change followed by 5 mins in speedvac with no heat.

Samples were rehydrated in trypsin solution adding just enough to fully rehydrate so gel pieces grew back to their original size. Gel pieces were then covered with 25 mM NH_4HCO_3 and left overnight at 37°C.

Formic acid was added to each sample to a final concentration of ~ 5%. Samples were

then vortexed, spun down and the supernatant (peptides) transferred to a clean tube. 1X volume of MeCN, (equal volume to volume of gel pieces) was added to each sample and left to shake for 5 min. Supernatant was then added to the peptides and samples containing the peptide eluate was reduced to a final volume of $\sim 8 \mu\text{l}$, using a speedvac, with no heat. The peptide eluates were analysed using LTQ-OrbitrapXL following standard operating procedure.

Chapter III

Results 1: The N-terminal evolutionarily driven domain of TDP1 is required for optimal cellular protection against genotoxic stress

3.1 Introduction

3.1.1 TDP1; discovery, substrate and mechanism of action

TDP1 was first discovered by Nash *et al* (S. W. Yang *et al.*, 1996) as they demonstrated a partially purified protein, extracted from *Saccharomyces cerevisiae*, that was capable of removing an abortive topoisomerase 1–DNA complex. This enzyme was able to hydrolyse the tyrosine-DNA phosphodiester linkage between DNA and Top1, leaving a 3' phosphate and 5' hydroxyl group (Davies *et al.*, 2002b; 2002a; Debéthune *et al.*, 2002; Interthal *et al.*, 2001) (**Figure 3.1.1**).

Unlike many other “end-processing enzymes” the product resulting from hydrolysis by TDP1 does not leave the appropriate 5' hydroxyl and 3' phosphate required for ligation of the DNA backbone. Therefore PNKP, chaperoned by XRCC1, is also required: a protein that contains kinase activity, to phosphorylate the 5' end of DNA and phosphatase activity to remove the 3' phosphate and replace it with a hydroxyl group (Jilani *et al.*, 1999; Plo, 2003; Weinfeld *et al.*, 2011). This allows the corresponding ATP dependent ligase to re-ligate the DNA, restoring its integrity.

TDP1's primary substrate is that of the stalled Top1 complex (El-Khamisy *et al.*, 2005; Katyal *et al.*, 2007; S. W. Yang *et al.*, 1996). The Top1 protein is an enzyme required to relieve supercoils within the DNA as a result of replication and transcription. The enzyme accomplishes this by creating a single-strand break within one strand of the DNA, forming a DNA-protein linked complex, and allows controlled rotation of the other strand to alleviate the supercoils (Champoux, 2001; Deweese *et al.*, 2008; Leppard and Champoux, 2005). Top1 has also been shown to be a player in the prevention of R-loops: RNA strands that get entwined with single-stranded DNA (as a result of transcription) by removing RNA bases from the DNA containing strands. This can lead to deleterious effects if untreated (Tuduri, Crabbe *et al.* 2009).

The breaking of DNA by Top1 is a transient one and is required for the removal of helical torsion. Once the stress has been alleviated Top1 uses the OH group at the 3'-break of the DNA for re-ligation, without the requirement for a DNA ligase, and thus

recycling the Top1 protein. Unfortunately, if lesions are in close proximity to the Top1-DNA adduct (Pommier, 2003; Pourquier and Pommier, 2001; Sordet et al., 2004), or if synthetic compounds such as camptothecin (CPT) (Hsiang et al., 1985) are used, Top1 becomes unable to remove itself from the DNA and forms a Top1-linked DNA single-strand break. It is also important to note that crystal structures of the Top1-DNA bound complex shows the 3' phosphotyrosine residue buried within the DNA-Top1 complex, making its accessibility for hydrolysis difficult (Redinbo et al., 1998). Thus, it is predicted that once the abortive Top1-complex forms it is ubiquitinated and then degraded by the 26S proteasome (Deb  thune et al., 2002; Desai et al., 1997; 2003; Interthal et al., 2005a). This mechanism leaves a tyrosine residue, and possibly small peptide sequence, covalently attached to a 3' phosphate at the nick of the DNA. It is this substrate that can then be hydrolysed by TDP1.

The abortive Top1 complex is not what causes toxicity to the cell; if this abortive complex remains, collision with the replication machinery, or transcriptional polymerases can result in double-strand breaks, DNA single-strand breaks and can impede the production of cellular proteins (El-Khamisy, 2006). Either of these outcomes can lead to cell death and defects in the repair pathway; stalled Top1 complexes have been implicated in several diseases (El-Khamisy et al., 2005; Kerzendorfer et al., 2010; Takashima et al., 2002).

The phosphodiesterase domain of hTDP1 was crystallised by Davies *et al* in 2002 (Davies et al., 2002b) and shows a tertiary structure composed of two α - β - α domains that contain the catalytic pocket with which the enzyme hydrolyses the 3' tyrosine phosphate (**Figure 3.1.1**). Analysis of the protein sequence shows this enzyme belongs to the phospholipase D (PLD) superfamily (Interthal et al., 2001), all of which catalyze a phosphoryl transfer reaction. These families contain conserved HKD motifs that are responsible for the similar reaction chemistry and follow the sequence HxK(x)₄D. Unlike other members of the PLD family TDP1 does not contain that aspartate residue but does have two motifs with conserved histidines and lysines at His263, Lys265, His493, and Lys495 (Interthal et al., 2001). Other conserved amino acid groups between TDP1 orthologs are Asn283, Gln294, Asn516, and Glu538. Due to the pseudo-twofold axis of symmetry residues Asn283 and Gln294 work in conjunction with His493 and Lys495, whilst Asn516, and Glu538 co-operate with His263 and Lys265 (Davies et al.,

2002b; 2002a; Raymond et al., 2004). These residues help balance out the proton transfer and allow H493 to activate a water molecule by stabilizing the reaction intermediates.

The mechanism of TDP1 has been elucidated and shown to be a two-step process: binding of the TDP1 substrate positions the phosphate moiety in the center of the active site where the histidine 263 can initiate the 1st nucleophilic attack, cleaving the phosphotyrosine and forming a phospho-TDP1 intermediate. The second step of the reaction involves H493, this histidine residue acts as a base that can activate a water molecule to hydrolyse the 3'phospho-His263 linkage. This restores the catalytic site of TDP1 and results in the release of the DNA containing a 3' phosphate (**Figure 3.1.2**).

It is not surprising then, that if the H263 is mutated to an alanine the TDP1 enzyme becomes catalytically dead, and no intermediate is formed. Evidence has been documented supporting the fact that the H493 is involved in the second step of the reaction, as mutation to an asparagine or arginine results in a significant reduction in catalytic activity, but the TDP1-DNA intermediate still occurs (Raymond et al., 2004). The latter histidine to arginine mutation at position 493 has been linked to spinocerebellar ataxia with axonal neuropathy 1 (SCAN1) (El-Khamisy et al., 2005; Hirano et al., 2007; Takashima et al., 2002).

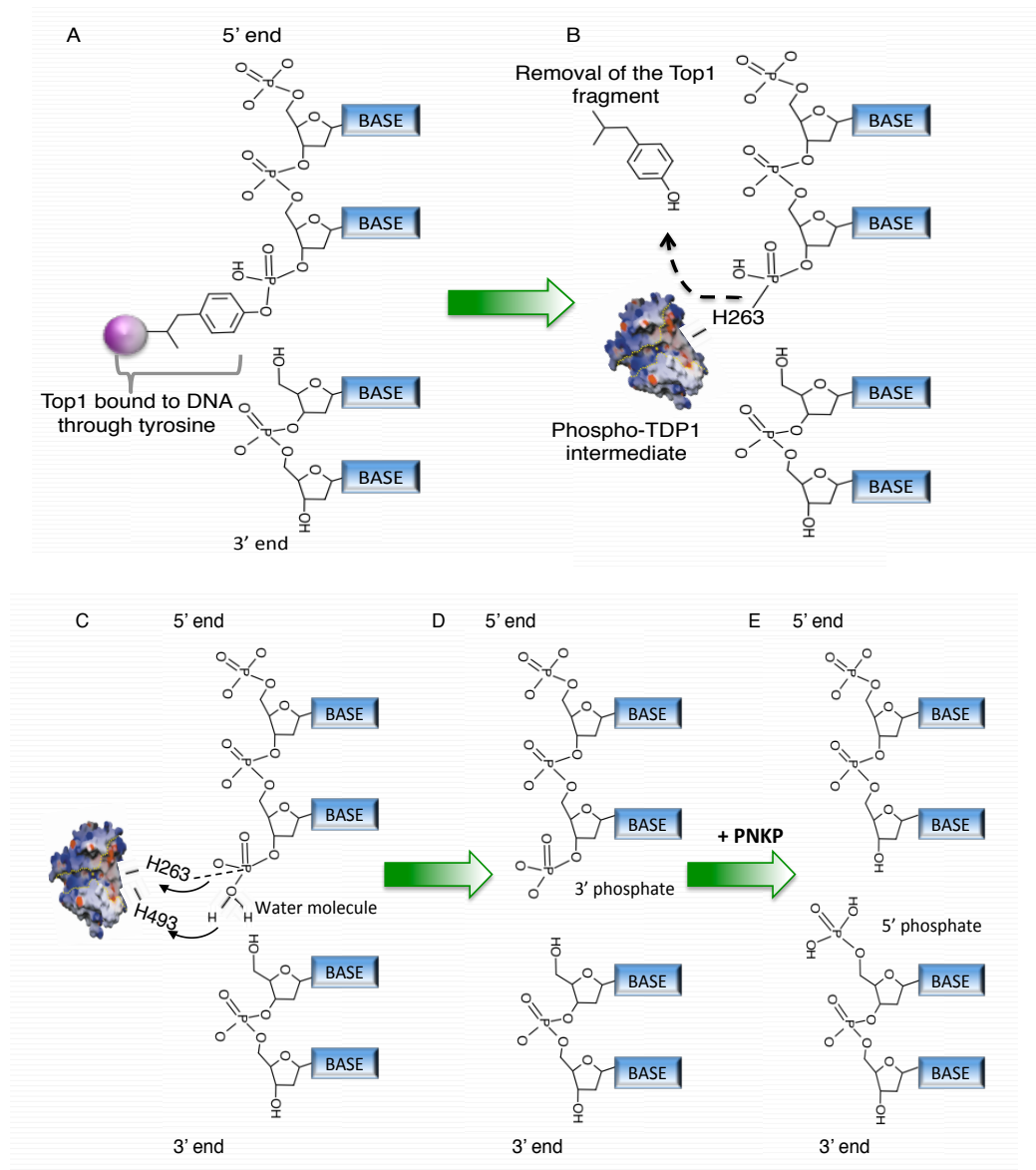


Figure 3.1.1 Cleavage of abortive Top1-DNA complexes by TDP1

Topoisomerase 1 becomes trapped, due to lesions near the Top1-DNA adduct or chemicals that prevent the enzyme from recycling. This causes a prolonged covalent bond between the tyrosine of Top1 and phosphate on the 3- end of the DNA break (A). Top1 is then degraded to a small peptide fragment of unknown length with the 3'phosphotyrosine residue remaining. To repair the stalled moiety, the H263 of TDP1, through nucleophilic substitution, removes the Top1 fragment by donating a proton, hydrolysing the tyrosine residue and forming a 3-phosphohistidine bond (B). The H493 of TDP1 within its negative environment would then act as a base, abstracting a proton from a nearby water molecule (C). This 'activated' water molecule would then act as the second nucleophile, resulting in the release and restoration of the H263 of TDP1, thus removing the protein from the DNA and leaving a 3'phosphate (D). These DNA ends at either side of the DNA nick (3'phosphate and 5-hydroxyl) are not suitable for ligation. Consequently PNKP, an enzyme with phosphatase and kinase capabilities, is required in the last step to leave a 3'OH and 5'P for ligation (E).

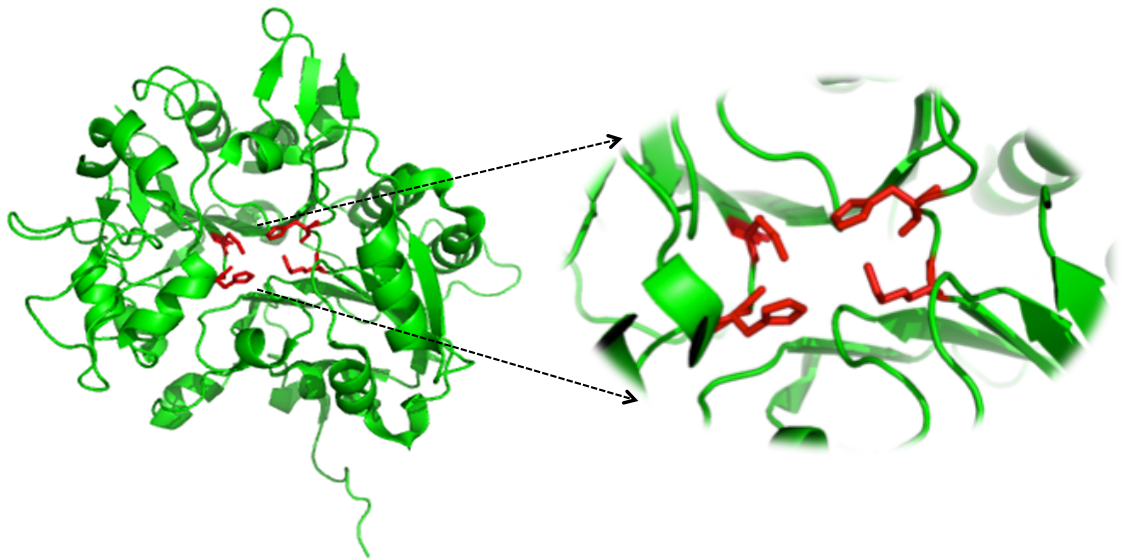


Figure 3.1.2 Crystal structure of TDP1

The crystal structure of TDP1 depicts the catalytic domain (left) and a magnification (right) of the catalytic pocket containing the histidines at positions 263 and 493, lysines at 265 and 495 and aspartic acid used in the two step nucleophilic attack to remove abortive Top1 complexes and other damaged DNA at the 3' end (Davies et al., 2002b).

3.1.2 TDP1's role in SCAN1

The importance of TDP1 was highlighted when a mutation at position 493 which changed a histidine to arginine was found and linked to the autosomal recessive disease SCAN1. At the molecular level this mutation was shown to have a 25-fold decrease in catalytic activity compared to the wild type protein (El-Khamisy et al., 2005; Interthal et al., 2005a). Furthermore, this mutation has an effect on the second nucleophilic substitution involved in the enzymatic action of TDP1. This second step is required for the recycling of TDP1 by the removal of the TDP1 protein from the bound DNA. As a result, this particular mutation can also cause TDP1 to become trapped on DNA, a process thought to be more toxic than the initial stalled topoisomerase complex (Hawkins et al., 2009; Interthal et al., 2005b).

It was thought that the combination of the reduced ability for TDP1 to remove abortive Top1 complexes and the stalled TDP1-linked DNA led to degeneration of the cerebellum and the pathologies seen in SCAN1. This disease was initially thought to be

a result of defects in TDP1 mediated nuclear repair, but further investigations by our lab (unpublished) and others (Das et al., 2010) show a role for TDP1 in the mitochondria, leading to speculation that either, or both could be contributory to the pathology seen (Fam et al., 2013; 2012; Weissman et al., 2007).

Patients with SCAN1 show progressive cerebellar atrophy, pes cavus, peripheral neuropathy and muscular atrophy. Clinically symptoms develop at a mean age of 15 years and result in ataxic gait, seizures and senescence of pain, touch, and vibration senses within the extremities. Patients tend to be wheel chair bound by early adulthood. At a cellular level, patients exhibit mild hypercholesterolemia and hypoalbuminemia (Takashima et al., 2002)

3.1.3 Chapter objectives

TDP1 in yeast and higher eukaryotes have been shown to hydrolyse stalled Top1-DNA complexes, they also possess a highly conserved catalytic domain that contains the activity of the enzyme. However, in higher eukaryotes TDP1 has evolved a domain at the N-terminus, which possesses a much lower conservation between species, and in human's spans 150 amino acids (25 kDa) in length (**Figure 3.1.3**). So if TDP1 in yeast can adequately remove Top1 from DNA why has the N-terminal in humans evolved?

Since TDP1 has been shown to have a role within the mitochondria, has the domain evolved for this sole purpose? It is a possibility since the N-terminus has been shown to interact with Ligase III (Chiang et al., 2010; Das et al., 2009), the only ligase that can access the mitochondria. There are however arguments that the N-terminus has evolved for more than just a mitochondrial role, as ligase III can also interact with XRCC1 for nuclear repair (Plo, 2003). This TDP1-Ligase III-XRCC1 mediated repair has been a common and accepted pathway of repair for Top1 damage within the nucleus (Caldecott, 2007).

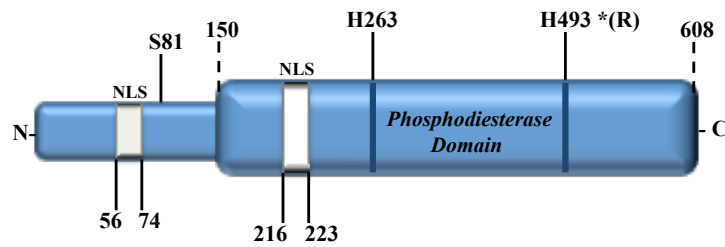


Figure 3.1.3 Schematic of TDP1

TDP1 consists of two domains, a catalytic domain denoted as the phosphodiesterase domain, which spans from amino acid 151 to 608 and is highly conserved throughout species. It contains a nuclear localization sequence and has histidine residues at positions 263 and 493 that are crucial for its catalytic activity. The H493 can be mutated to an arginine and has been found in patients with SCAN-1. The N-terminal domain (amino acids 1-150) is only found in higher eukaryotes and has a lower sequence homology between species. The N-terminal domain contains a nuclear localization sequence and in humans is phosphorylated at serine 81 by ATM and DNA-PK.

This first chapter looks into what possible benefit and role the N-terminal domain of TDP1 may possess. To better understand the role of the N-terminal domain of TDP1 in vertebrate cells. I engineered DT40 Tdp1 $-/-$ (a kind gift from Yves Pommier) by complementing them with Myc-tagged, human, full-length TDP1 (hTDP1) or the human catalytic domain of TDP1 (TDP1¹⁵¹⁻⁶⁰⁸). Because chicken DT40 cells have well characterised repair pathways (Yamazoe et al., 2004) we could study the effect the loss of the N-terminal domain had on the protection against various DNA damaging agents. There is also a high degree of amino acid sequence identity between human and chicken Tdp1 (69 %) with all catalytic residues conserved (**Figure 3.1.4**); therefore, hTDP1 is predicted to act in a similar fashion to that of the endogenous chicken Tdp1.

Human	1	MSQEGDYGRWTISSSDESEEEKPKPKDPSTSSLLCARQGAANEPRYTCSAQKAAHKRKI	60
Chicken	1	M QEG +GRWT+SSS++S EE +KPSTSSLL A + A+ P+Y CSEA+K AHKRK	60
		MLQEGAHGRWTVSSSEDSTEENSDESEKPKSTSSLLSAPRSEASGPQYPCSEARKVAHKRKA	60
	61	SPVKFSNTD--SVLPKRQKSGSQEDLGWCLSSSDDELQPEMPQKQAEKVVIKKEKDISA	118
		SP++FS+T + PP ++ QE LGWCLSSSD+E PE +K K +K+EK	
	61	SPLRFSDTHLPAQTTPAEKQRPQGEGLGWCLSSSDEE--PEDREKHTRKETLKEEK----	114
	119	PNDGTAQRTENHGAPACHRLKEEEDYEYETSGEGQDIWMDLKGNNPFQFYLTRVSGVKPKY	178
		D ++ ++H + ++ ED E GE QD WD+L GNP F+LT+V G++ Y	
	115	-CDAPREQPQSHCKDEHSEKNEKAEDYNEVLGEPQDTWDLSSGGNPFQFFLTQVKGIEQSY	173
	179	NSGALHIKDILSPLFGTLVSSAQFNYCFDWDLVKQYPPEFRKKPILLVHGDKREKAHL	238
		NSGALHIKDILSPLFGTLVSSAQFNYC DV WLW+QYP E+RKKP+L+VHG+KRE+KA L	
	174	NSGALHIKDILSPLFGTLVSSAQFNYCIDVAWLVRQYPQEYRKKPLLIVHGKRESKAEL	233
	239	HAQAKPYENISLCQAKLDIAFGTHHTKMMLLLYEEGLRVVIHTSNLIHADWHQKTQGIWL	298
		AQA+P+ENIS CQAKLDIAFGTHHTKMMLLLYEEGLRVVIHTSNLI DHWQKTQGIWL	
	234	LAQARPFENISFCQAKLDIAFGTHHTKMMLLLYEEGLRVVIHTSNLIAEDWHQKTQGIWL	293
	299	SPLYPRIADGTHKS-GESPTHFKADLISYLMAYNAPSLKEWIDVIHKKDLSETNVYLIGS	357
		SPLYPR+ G+ S GES T+FK+DLISYLMAY++P LKEWID+I +HDLSET VYL+GS	
	294	SPLYPRLPQGSSDSAGESETNFKSDLISYLMAYSSPVLKEWIDLIREHDLSETRVYLLGS	353
	358	TPGRFQGSQKDNWGHFRLKLLKDHASSMPNAESWPVVGQFSSVGSGLGADESKWLCSEFK	417
		TPGR+QG K+ WGH +L+KLLKDHASS+P ESWPVVGQFSS+GSLGAD SKWLCSEF+	
	354	TPGRYQGIDKEKWGHLKLRKLLKDHASSIPAQESWPVVGQFSSIGSLGADGSKWLCSEFQ	413
	418	ESMLTLGKESKTPGKSSVPLYLIYPSVENVRTSLEGYPAGGSLPYSIQTAEQKNWLHSHYF	477
		ES++ G K VP++L+YP+V NVR SLEGYPAGGSLPYSIQTA+KQ WLHSHYF	
	414	ESLVAAGSGVAALLKCDVPIHLVYPTVSNVRQSLEGYPAGGSLPYSIQTAQKQLWLHSHYF	473
	478	HKWSAETSGRSNAMPHIKTYMRPSPDFSKIAWFLVTSANLSKAAWGALEKNGTQLMIRSY	537
		HKWSAE SGRS+AMPHIKTYMRPS DF KIAWFLVTSANLSKAAWGALEKNGTQLMIRSY	
	474	HKWSAEVSGRSHAMPHIKTYMRPSHDFQKIAWFLVTSANLSKAAWGALEKNGTQLMIRSY	533
	538	ELGVFLFLPSAFGLDS--FKVKQKFFAGSQEPMATFPVPYDLPPELYGSKDRPWIWNIPYV	595
		ELGVFLFLPSAFGLD F VK + ++ +FPVP+DLPPE YGSKD+PWIWNIPY	
	534	ELGVFLFLPSAFGLDKGYFHVKGNNMLSEGKDSATSFVPVFDLPPELYGSKDQPWNIWNIPT	593
	596	KAPDTHGNMWVPS 608	
		APDTHGNMWVPS	
	594	SAPDTHGNMWVPS 606	

Figure 3.1.4 Alignment of Human and Chicken Tdp1 amino acid sequence

The amino acid sequences for human TDP1 (accession NP_060789) and chicken Tdp1 (accession XP_421313) share 69 % identity (letters) and 11 % similarity (+) giving a total of 80 % positive hits according to the NCBI BLASTP sequence alignment program.

3.2 Results

3.2.1 Full-length TDP1 and the catalytic domain show no difference in activity, structure or stability

To assess the impact of the N-terminal domain, full-length recombinant hexa-His tagged TDP1 and TDP1¹⁵¹⁻⁶⁰⁸ were expressed in *E.coli* and purified using a nickel charged column (**Figure 3.2.1**). These recombinant proteins were then used to determine if the N-terminal domain had an impact on the activity, structure or stability *in vitro*.

3.2.1.1 The C-terminal domain does not demonstrate a reduction in enzymatic activity *in vitro*

Whilst the catalytic sites of TDP1 reside in the C-terminal catalytic domain the N-terminal domain could have an effect on the overall structure of the catalytic pocket by opening up the cleft and making it more accessible to the target substrate. Alternatively it could be important in the binding of the DNA phosphate backbone, thereupon orientating the catalytic sites into a better position to remove the 3' lesion from DNA. To investigate the role the evolutionarily driven N-terminal domain may have on TDP1's activity, recombinant TDP1 or TDP1¹⁵¹⁻⁶⁰⁸ were added to a fluorescence based activity assay.

This assay was developed in-house and contains an oligo with a 3' phospho-tyrosine (**Figure 3.2.2**), which mimics a stalled Top1 substrate. The tyrosine residue is covalently bound to a fluorescent molecule. This assay uses a trivalent metal ion sensor that is bound to the DNA phosphate backbone so it can quench the fluorescence of any fluorophore within close proximity by absorbing the electron transfer. Hence removal of the fluorophore containing tyrosine from the DNA results in fluorescence that can be recorded.

6.25 pM of either full-length TDP1, TDP1¹⁵¹⁻⁶⁰⁸ or assay buffer was mixed with 10 nM of the oligo substrate. The reaction was left for 10 minutes at room temperature before fluorescence was recorded. Incubation of His-TDP1 or His-TDP1¹⁵¹⁻⁶⁰⁸ with the Top1-

synthetic substrate resulted in an increase in fluorescence and therefore enzymatic cleavage of the 3'-phosphotyrosine. Results indicate that TDP1 and TDP1¹⁵¹⁻⁶⁰⁸ show no significant difference in enzymatic activity with respective average fluorescence units of 751 to 813 (30 mins), 1626 to 1426 (60 mins) or 2036 to 1917 (90 mins) (**Figure 3.2.2**). This indicates that the N-terminus has no impact upon activity of TDP1 *in vitro*. The H493N mutant gave readings of 183, 195 and 324 at 30, 60 and 90 minute time points. This shows that mutation of H493 to an asparagine disrupts the catalytic activity of the enzyme, dramatically lowering the enzymes efficiency to hydrolyse the 3'phosphotyrosine in the synthetic substrate. This decrease in activity supports previous data (El-Khamisy et al., 2005; Lebedeva et al., 2012; Raymond et al., 2004).

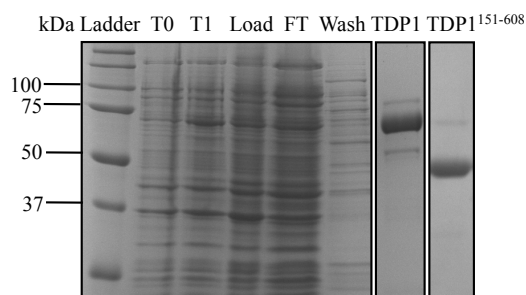


Figure 3.2.1 Purification of recombinant full-length human TDP1 and TDP1¹⁵¹⁻⁶⁰⁸

N-terminal His-tagged recombinant hTDP1 and the truncated catalytic domain were expressed in *E.coli* under a T7 promoter. Bacterial cells were cultured (T0) before induction of expression by IPTG (T1). The lysed cells were then loaded onto a pre-packed nickel charged HiTrap column (Load). Unbound protein was then washed off in buffer (FT) and weakly bound proteins were eluted in buffer containing 100 mM imidazole (Wash). The protein of interest was eluted in 25 mM HEPES, 250 mM imidazole, 500 mM NaCl at pH 7.2 (lane TDP1 and TDP1¹⁵¹⁻⁶⁰⁸).

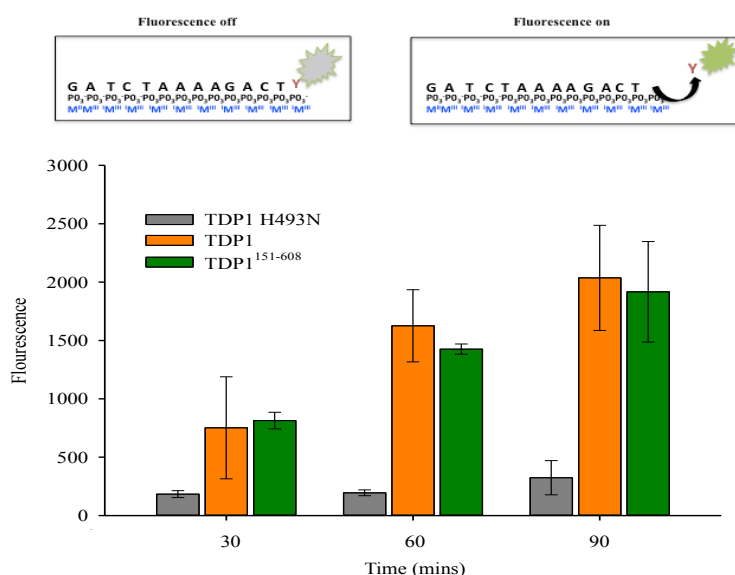


Figure 3.2.2 Removal of the N-terminal domain does not impact upon activity in *vitro*

A 13-mer oligo with a 3'-phosphotyrosine conjugated to a FITC molecule was synthesised. This also had a small molecular, non-fluorescent trivalent metal ion sensor (Mn^{III}) (Gyrasol technologies, USA) that binds the phosphate backbone. This trivalent metal would quench the fluorescence of any fluor by electron transfer. Therefore removal of the fluorophore containing Tyrosine from the oligo, by an active TDP1 molecule, would allow fluorescence from the molecule to be recorded. 6.25 pM recombinant TDP1, TDP1 H493N or TDP1¹⁵¹⁻⁶⁰⁸ was mixed with 10 nM of the substrate. Reactions were then quenched at several time points and fluorescence was recorded using PHERAstar software on a BMG Labtech Pehrastar Plate reader. Analyses of data show no difference in activity between the full-length and catalytic domain of TDP1, whilst the H493N mutation severely inhibited activity. Error bars represent standard deviation from 3 independent repeats.

3.2.1.2 TDP1 and TDP1¹⁵¹⁻⁶⁰⁸ have no difference in thermal stability, or structure

Next, we tested the thermal stability of the full-length TDP1 and TDP1¹⁵¹⁻⁶⁰⁸ using a dye that fluoresces when bound to hydrophobic regions of a protein. Recombinant wild type and TDP1¹⁵¹⁻⁶⁰⁸ were subject to Thermal denaturation profiling: a technique used to determine the stability and profile the unfolding of a protein (Rehman et al., 2011). Fluorescence values were obtained using the LightCycler 480 (Roche) and melting points determined using a variation of the Boltzmann model (**Figure 3.2.3 B**) (Ericsson et al., 2006). By comparing the denaturation curves proteins showed a similar unfolding pattern and resulted in T_m of 55.5 and 54.3 °C for the truncated and wild-type proteins respectively, indicating no loss in stability or difference in the unfolding pattern of the two proteins (**Figure 3.2.3 A**).

We also compared the circular dichroism absorption spectrum of TDP1 and TDP1¹⁵¹⁻⁶⁰⁸, to investigate the secondary structure, through circular dichroism spectroscopy; a well-known method for structural comparison (Lawton, 2002; Whitmore and Wallace, 2008). The two proteins resulted in two similar spectral graphs with negative bands at ~210 and 220 nm, and positive bands at ~195 nm (**Figure 3.2.4 A**). High-tension voltage for both spectra was under 600 V (**Figure 3.2.4 B**) at the lowest analysed wavelength, a value required for interpretation of suitable data (Kelly et al., 2005). Subtraction of the TDP1¹⁵¹⁻⁶⁰⁸ plot from the wild-type gave a spectra not associated with any structure, when analysed using algorithms for secondary structure (**Figure 3.2.4 C**).

To obtain a quantitative read out analysis of secondary structure using DichroWeb by the variable selection algorithm (CDSSTR), which provides superior fits for globular proteins (Lobley et al., 2002; Sreerama and Woody, 2000; Wallace, 2003; Whitmore and Wallace, 2008), revealed no major difference in α -helical or β -sheet content, suggesting no apparent change in structure between the catalytic domain and full-length TDP1 (**Table 3.1**). From these analyses we conclude that there is no significant change in structure or unfolding of the truncated protein compared to wild-type TDP1. Interestingly, the catalytic domain had already been crystallised (Davies et al., 2002b); so a comparison of the known alpha helix, beta sheet and unstructured content could be drawn with the quantitative read-out of DichroWeb. Alpha helix, beta sheet and unstructured content between the known crystal structure and interpretation of CD

spectra were quite similar (**Table 3.2.2**). This means that secondary structure analysis of the full-length protein by DichroWeb could be considered to be reasonably accurate.

These results indicate that the majority of the N-terminal TDP1 domain is unstructured, a statement supported by *in silico* analysis of the theoretical structure of this domain when analysed by the 3D structural prediction software Phyre2 (**Figure 3.2.5**) (Kelley and Sternberg, 2009). Phyre2 software prediction of the TDP1 N-terminal amino acid sequence alone shows, mainly, an unstructured domain with some alpha helical content (**Figure 3.2.5 A**). Interestingly, when analysis of full-length TDP1 is carried out the software predicts no helical content within the N-terminal domain of TDP1 (**Figure 3.2.5 B**). Both models for prediction support the CD data and DichroWeb analysis strengthening the idea that the evolutionarily driven domain is predominantly unstructured.

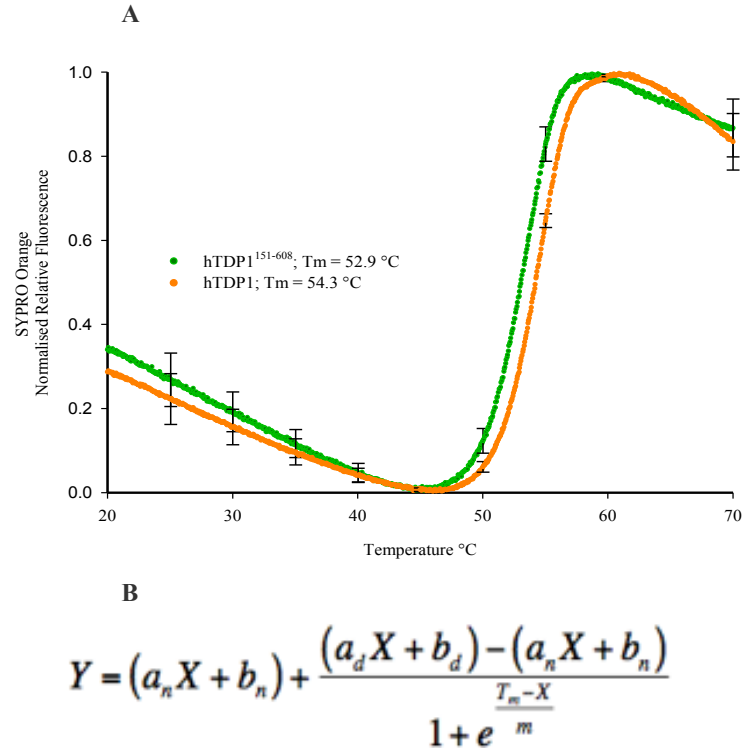


Figure 3.2.3 Loss of the N-terminal domain does not impact on stability *in vitro*

0.75 μM of recombinant hTDP1 or hTDP1¹⁵¹⁻⁶⁰⁸ was mixed with SYPRO-orange, and melting profiles were obtained between 20 and 70 °C, using a ramping rate of 0.03 °C s⁻¹. Protein denaturing experiments were carried out using the LightCycler 480 System II (Roche). Data were normalized to a fraction of protein in its denatured state and presented as normalized relative fluorescence. Melting temperatures (T_m) were determined according to the Boltzmann model (Ericsson et al., 2006) (**B**) using the Prism software. Error bars represent standard deviation from 3 independent replicates. Loss of the N-terminal domain does not diminish the stability, unfolding pattern or melting temperature of the protein.

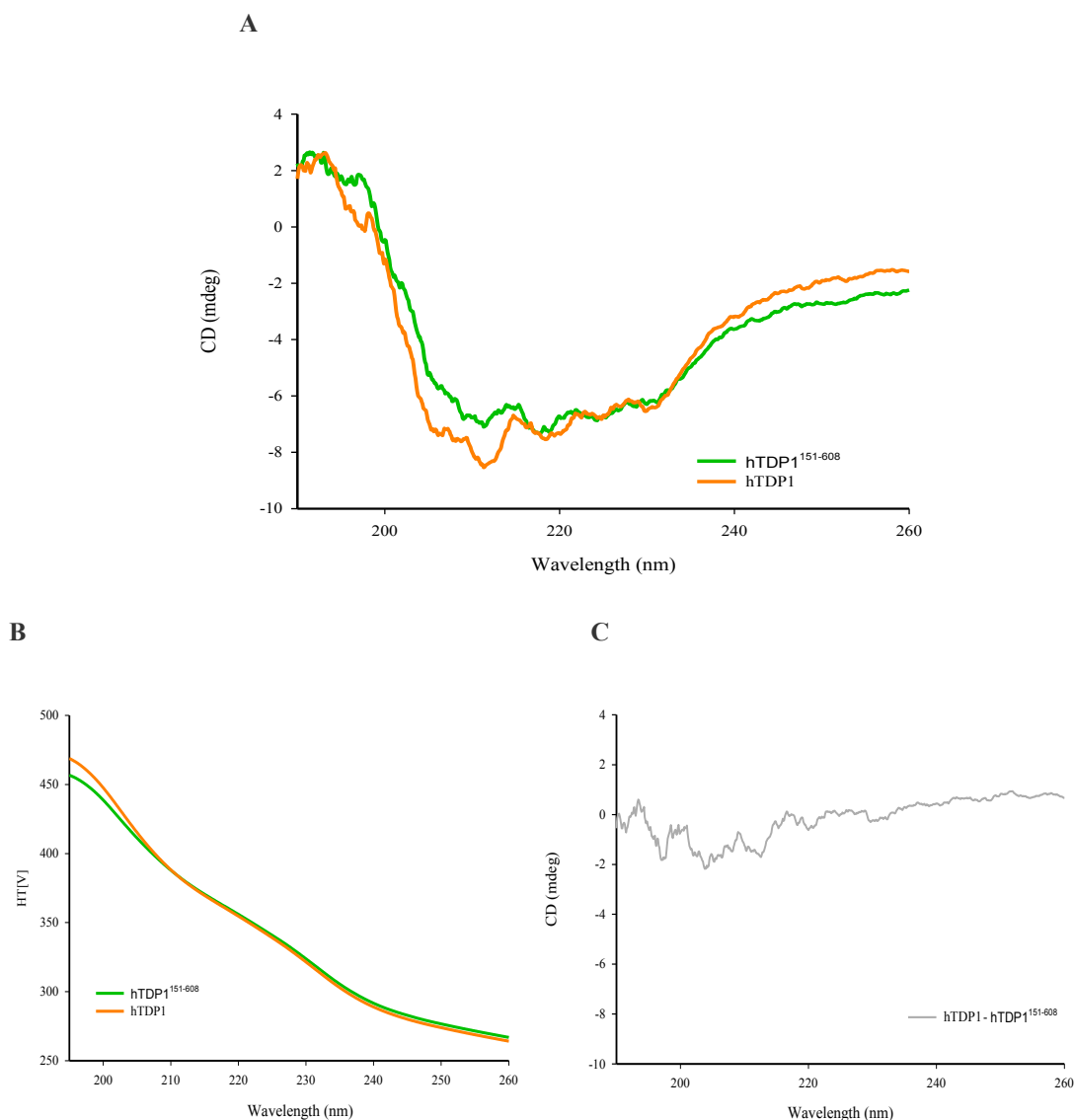


Figure 3.2.4 Comparison of hTDP1 and hTDP1¹⁵¹⁻⁶⁰⁸ using CD shows no gross structural difference Recombinant TDP1 or hTDP1¹⁵¹⁻⁶⁰⁸ in 30 mM sodium phosphate buffer (pH 7.4) was placed in a 0.2-mm quartz cuvette and the circular dichroism spectrum determined using a JASCO J-715 spectropolarimeter. Scans were performed from 260 to 195 nm and buffer baselines were subtracted (**A**). Data only recorded if the high-tension voltage (HT [V]) < 550 V (**B**). Subtraction of the hTDP1¹⁵¹⁻⁶⁰⁸ spectra from that of hTDP1 (**C**) results in a plot that analysis with secondary structure predicts no gross structure present.

	Helix		Sheet		Others*		NRMSD# (<0.25)	
CDSSTR Reference set 4	TDP1	TDP1 ¹⁵¹⁻⁶⁰⁸	TDP1	TDP1 ¹⁵¹⁻⁶⁰⁸	TDP1	TDP1 ¹⁵¹⁻⁶⁰⁸	TDP1	TDP1 ¹⁵¹⁻⁶⁰⁸
Secondary Structure (%)	19.2	19.4	28.4	28.8	52.4	51.8	0.22	0.24

* Loops, turns, and unstructured

Normalised root mean square deviation

Table 3.2.1 Bioinformatics analyses of core secondary structure using CDSSTR algorithm

Secondary structure analysis was conducted using the variable selection method (CDSSTR) with reference data set 4 from the Dichroweb server, which gave the best goodness-of-fits. NRMSD values were 0.22 and 0.24 for TDP1 and hTDP1¹⁵¹⁻⁶⁰⁸ respectively.

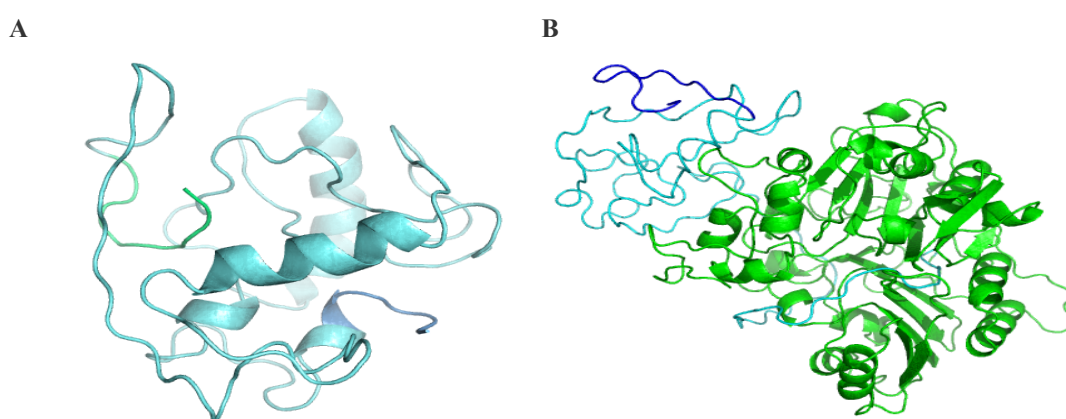


Figure 3.2.5 Phyre2: N-terminus of TDP1 is mainly unstructured

Amino acid sequences for TDP1¹⁻¹⁵⁰ (A) and TDP1 (B) were analysed by the protein folding prediction site Phyre2. The hypothetical model shows the N-terminal protein of TDP1 to be largely unstructured, with some small alpha helices (A). Whilst in the full-length model the N-terminus, highlighted in cyan (residues 1-10 highlighted in purple), is completely unstructured (B). This indicates, as hypothesised, that the N-terminal evolutionarily driven domain is predominantly unstructured in terms of secondary structure.

	Helix		Sheet		Others*	
	TDP1 ¹⁵¹⁻⁶⁰⁸ Crystal structure	TDP1 ¹⁵¹⁻⁶⁰⁸ Circular dichroism	TDP1 ¹⁵¹⁻⁶⁰⁸ Crystal structure	TDP1 ¹⁵¹⁻⁶⁰⁸ Circular dichroism	TDP1 ¹⁵¹⁻⁶⁰⁸ Crystal structure	TDP1 ¹⁵¹⁻⁶⁰⁸ Circular dichroism
Secondary Structure (%)	24	19.4	22	28.8	54	51.8

* Loops, turns, and unstructured

Normalised root mean square deviation

Table 3.2.2 Comparison between the crystal structure and predicted structure from CD

Comparison of the secondary structure between the crystal structure (UniProtKB: Q9NUW8) and the predicted structure from CD data using CONTIN show similar levels of secondary structure.

3.2.2 Full-length hTDP1 is required for optimal protection against DNA damaging agents in DT40 cells, whilst hTDP1¹⁵¹⁻⁶⁰⁸ gives partial protection

Since the catalytic domain and full-length TDP1 show no differences in activity, stability, or structure *in vitro*, it would be interesting to investigate what cellular function this domain may have. To study the *in vivo* role of the N-terminal domain of TDP1, stable cell lines expressing Myc-tagged full-length human TDP1 and Myc-TDP1¹⁵¹⁻⁶⁰⁸ (catalytic domain) were created. These cell lines were made in DT40 cells for three reasons. Firstly, we already had Tdp1 -/- cells available in the lab; secondly, DT40 cells have a high ratio of targeted to random integration of transfected gene constructs (Buerstedde and Takeda, 1991) and finally they have well characterised repair pathways (Yamazoe et al., 2004).

Once the cells had been transfected with linearized Myc-hTDP1, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ or Myc vector, single colonies were screened by immunoblotting, using anti-Myc (9B11; Cell signalling), to detect the protein of interest (**Figure 3.2.6**). Positive clones were further analysed for cells expressing similar levels of Myc-hTDP1 and Myc-hTDP1¹⁵¹⁻⁶⁰⁸, compared to the internal PCNA control, detected by anti-PCNA (PC10; Alan Lehman). Clone 13 (hTDP1) and clone 1 (hTDP1¹⁵¹⁻⁶⁰⁸) were selected due to similar levels of protein expression (**Figure 3.2.6**).

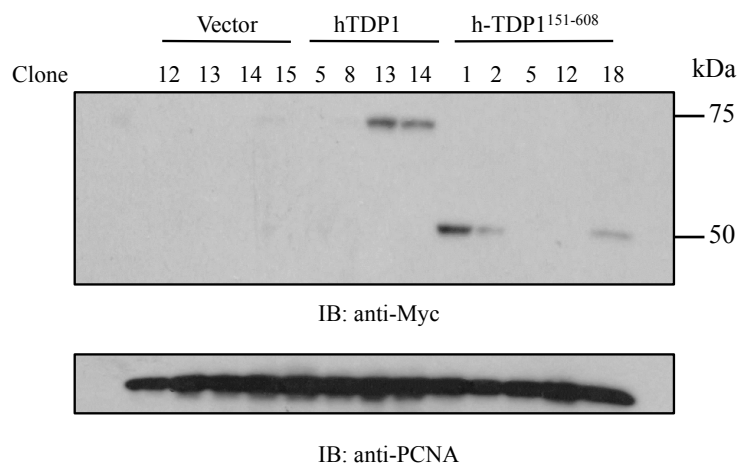


Figure 3.2.6 Complementation of DT40 Tdp1^{-/-} with Myc-hTDP1, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ or Vector
 DT40 TDP1^{-/-} cells were complemented with Myc-TDP1, Myc-TDP1¹⁵¹⁻⁶⁰⁸ or Myc vector. Immunoblots analysed with anti-Myc antibody (9B11; cell signalling) show expression of full length Myc-hTDP1 in clones 13 and 14 whilst expression of Myc-hTDP1¹⁵¹⁻⁶⁰⁸ could be detected in clones 1, 2 and 18. Clone 13 for Myc-hTDP1 and clone 1 for Myc-TDP1¹⁵¹⁻⁶⁰⁸ were selected due to similar expression levels when compared to the PCNA loading control.

3.2.2.1 Tdp1^{-/-} DT40 cells, complemented with Myc, Myc-hTDP1 or Myc-hTDP1¹⁵¹⁻⁶⁰⁸ maintain similar proliferation rates and cell cycle progression

To rule out that any effects seen between the cell lines in the presence of DNA damaging agents would be due to defects in the repair to damage and not due to the detrimental effects caused by knocking out TDP1 in the chicken cells, proliferative rates of these cell lines were measured by counting cells over a period of days using a haemocytometer. **Figure 3.2.7 A** shows similar proliferation rates between all cell lines used in this chapter, indicating that removal of Tdp1 and complementation with Myc, Myc-hTDP1 or Myc-hTDP1¹⁵¹⁻⁶⁰⁸ had no detrimental effect on replication.

To show that cell lines also had a similar fraction of their population in a particular phase of the cell cycle, cells were stained with 20 µg/ml propidium iodide and analysed using the FACS Canto flow cytometer (**Figure 3.2.7 B**). This was repeated on 3 independent occasions. The graphs show that there was no significant difference in the stage of cell cycle between cell lines. The probability of any difference between stages of the cell cycle between the generated cell lines when compared to the WT cells was > 0.44 (Students' t-test). Taken together, these results indicate that all cell lines are replicating at a similar rate and that there is no significant difference in fractions of cells in different phases of the cell cycle. Consequently, differences between cell lines in the presence of DNA damaging agents would not be a result of differing replication rates or cell cycle populations.

3.2.2.2 TDP1¹⁵¹⁻⁶⁰⁸ does not exhibit reduced activity in cell extracts

Whilst recombinant full-length TDP1 showed no difference when compared to the catalytic domain *in vitro*, protein interactions with TDP1 or post-translational modifications of TDP1 within the cell could impact upon the activity of the enzyme. To investigate this possibility 1.2 µg of whole cell extracts expressing similar levels of Myc-hTDP1, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ or Myc vector were then subjected to the fluorescence based activity assay previously described. Results indicate no significant difference in cleavage of the 3' phosphotyrosine between the full-length and catalytic domain (**Figure 3.2.8**) with arbitrary fluorescent units of $3087 \pm \text{s.d } 112.4$ and $3259.67 \pm \text{s.d } 403.8$ for the full-length and C-terminal domain respectively. This was not due to differences in expression levels of TDP1 as immunoblots show similar levels of expression (**Figure 3.2.8**). As expected, whole cell extract expressing Myc alone

showed little activity above basal amounts, with a reading of 281.33. Whilst nucleolytic cleavage of DNA could remove the 3'phosphotyrosine conjugated to a fluorophore, a result supported by previous work that demonstrates a capacity for other enzymes to carry out this role (Connelly and Leach, 2004; Eid et al., 2010; Nakamura et al., 2010) this has only ever been demonstrated to work on double-stranded DNA breaks (Shrivastav et al., 2008). We can conclude that the Myc vector and other proteins within the cell have no or little effect on the hydrolysis of the synthetic substrate. This data suggests that there is no difference in catalytic activity in cells extracts on a Top1-mimic synthetic substrate.

Since the activity of the full-length TDP1 and catalytic domain was similar, the next step was to see if they could provide equal protection against various DNA damaging agents. To determine this, cell viability assays were set up using a cell titre blue read-out (Promega). Viable cells can metabolise resazurin to resorufin, which causes a colour change from blue to pink (**Figure 3.2.9**) this colour change emits a fluorescence at 590 nm that can be measured using the **GloMax®**-Multi Detection System (Promega). The assay set up and concentrations of drugs were based on a previously published paper by the Pommier lab (Murai et al., 2012). Drug treatments were chronic and cells were left for 72 hours at 39 °C and 5 % CO₂ before addition of cell titre blue. Once added, viability of cells with or without the drug were measured after a further 3 hours at 39 °C and 5 % CO₂.

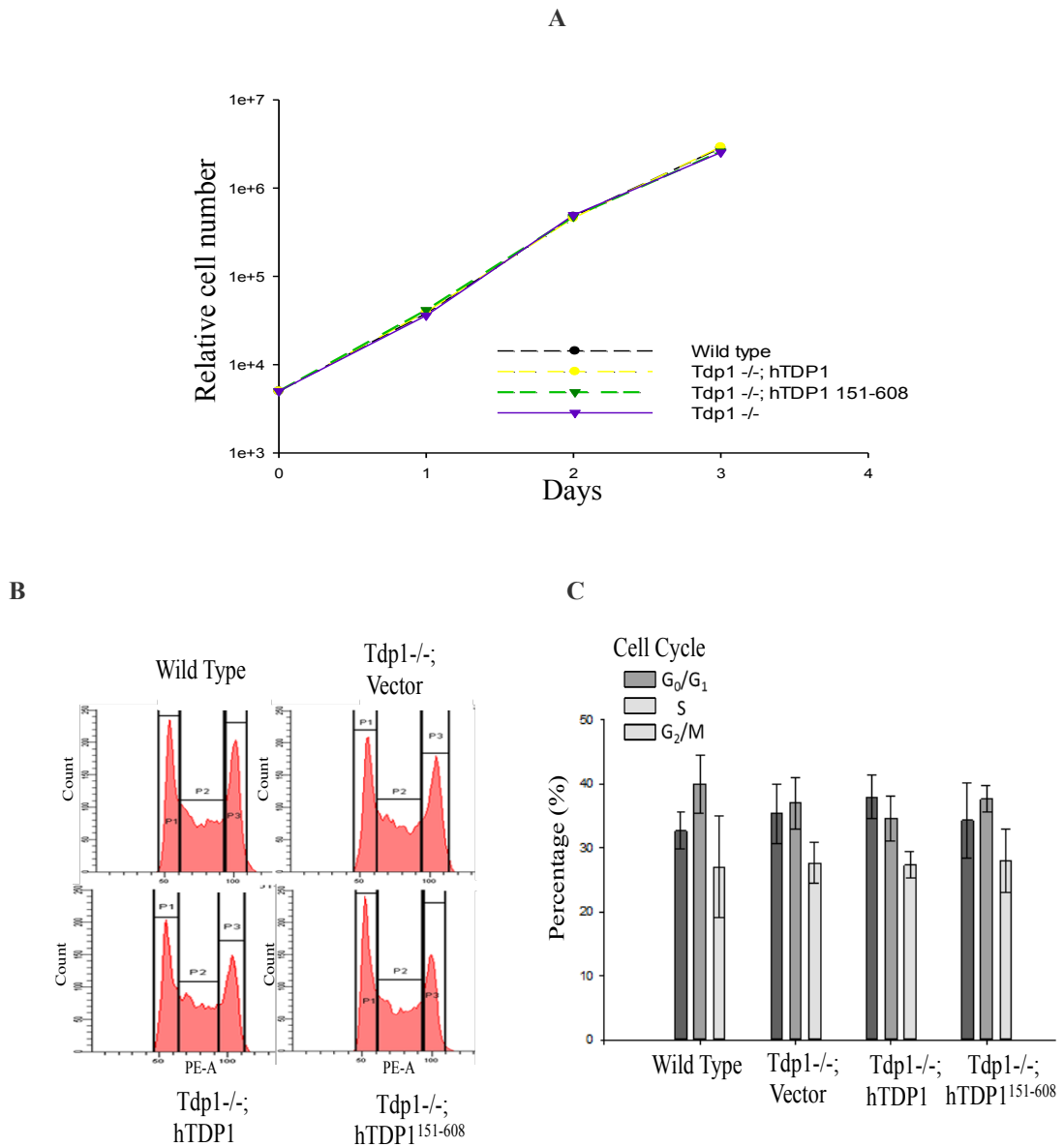


Figure 3.2.7 Characterisation of DT40 cells for cell cycle and proliferation

(A) Cell proliferation rates were studied for 72 hrs for each cell line. Results indicate that TDP1 is not required for maintenance of DT40 growth as no difference in doubling times can be seen. (B) FACS analysis: 1×10^6 cells for each cell line collected and stained with propidium iodide. Cells were analysed for DNA content on a FACS Canto flow cytometer and data show that knocking out Tdp1 or expression of hTDP1 does not change the cell cycle profile in DT40 cells. (C) Bar charts for average proportion of cells in each phase of the cycle from 3 independent FACS analyses and error bars are represented by standard deviation (B).

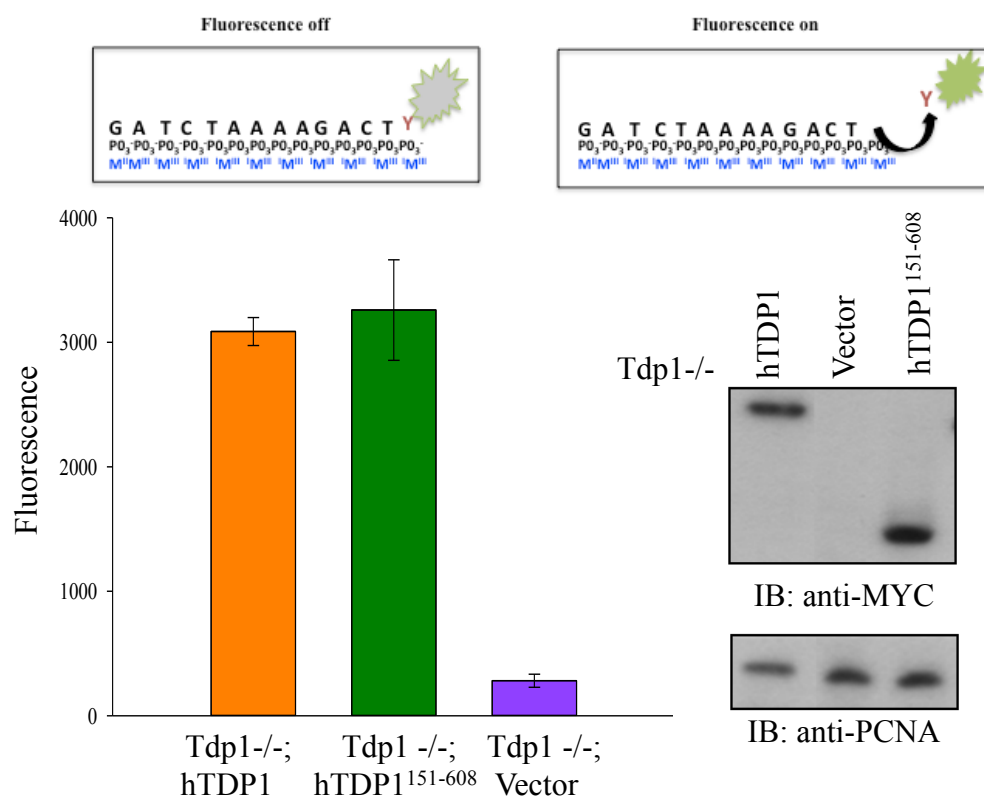


Figure 3.2.8 Removal of the N-terminal domain does not impact upon activity *in vivo*

A 13-mer oligo with a 3'-phosphotyrosine conjugated to a FITC molecule was made. This also had a small molecular, non-fluorescent trivalent metal ion sensor (Mn^{III}) (Gyrasol technologies, USA) that binds the phosphate backbone. 1.2 µg of whole cell extracts expressing similar levels of Myc-hTDP1, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ or Myc (indicated by immunoblots) were mixed with 10 nM substrate. Analyses of data show no difference in activity between the full-length and truncated catalytic domain of TDP1, whilst cells not expressing Tdp1 had little to no effect on the activity.

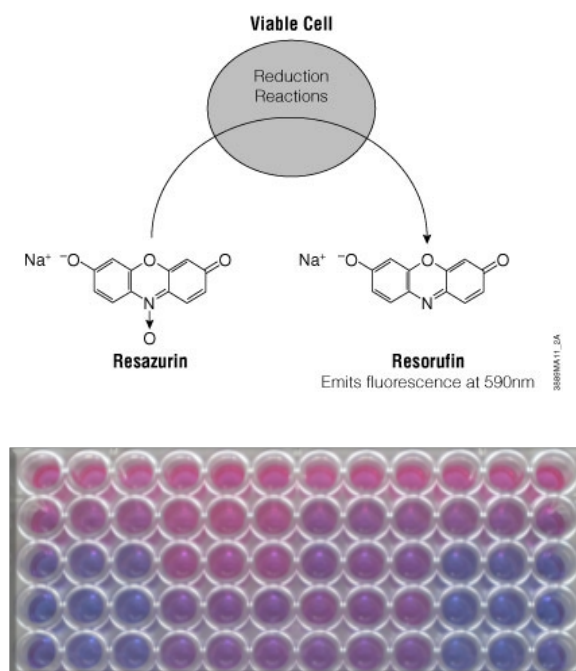


Figure 3.2.9 Cell viability assay: healthy cells metabolise Resazurin to Resorufin

Cell titre blue was used to determine viable cells. The healthy cells are able to metabolise Resazurin to Resorufin, the latter emits fluorescence at 590 nm which can be detected using the **GloMax®**-Multi Detection System (Promega). This can also be seen through a colour change from blue to pink.

3.2.2.3 Protection of cell lines treated with different genotoxic agents shows that the TDP1¹⁵¹⁻⁶⁰⁸ gives partial resistance when compared to the full-length protein

The first drug treatment was camptothecin (CPT): this is a plant alkaloid whose derivatives (topotecan, irinotecan and belotecan) are clinically approved anticancer agents for a broad spectrum of tumours (Hsiang et al., 1985; Liao, 2006; van Waardenburg, 2004). The molecular target is Top1 and the drug acts by blocking the re-ligation step of the DNA backbone, trapping Top1 to the DNA resulting in an abortive Top1-DNA complex.

Figure 3.2.10 A shows that in the presence of increasing concentrations of CPT wild-type DT40 cells show resistance at the low nano-molar range. Even at 10 nM CPT 58 % of wild-type cells were still viable. In stark contrast, DT40 Tdp1 ^{-/-} cells showed a marked increase in sensitivity to the drug with as little as 30 % of cells still viable at the lowest dosage (2.5 nM CPT) and 3 % of cells viable in the presence of 10 nM CPT. DT40 Tdp1 ^{-/-} cells complemented with Myc-hTDP1 gave a resistance to the drug to similar levels of wild-type, with 47 % of cells still viable at 10 nM treatment. This shows that the high sequence similarity between chicken and human TDP1 allows the human homologue to protect cells against camptothecin and therefore Top1 induced breaks to a similar level in the DT40 cell line. Strikingly, cells complemented with Myc-hTDP1¹⁵¹⁻⁶⁰⁸, with no difference in catalytic activity (**Figure 3.2.8**) or protein expression (**Figure 3.2.10 B**) showed an intermediate resistance between Tdp1 knockout cells and the hTDP1-complemented cells, with viability down to 6 % when treated with 10 nM CPT. A significant increase in sensitivity to CPT could be seen in Myc-hTDP1¹⁵¹⁻⁶⁰⁸ compared to the hTDP1 complemented cells.

I next examined the effect of methyl methanesulphonate (MMS): a strong electrophile that attacks the nucleophilic centers in DNA, resulting in DNA base methylation (Kedar et al., 2012; Pegg, 1984). This can cause nuclear and mitochondrial damage to DNA. Methylation predominantly occurs on N7-deoxyguanosine and N3-deoxyadenosine, it can also methylate other oxygen and nitrogen atoms in DNA bases, although to a lesser extent. In the context of TDP1 mediated repair methylated bases can trap Top1 causing abortive Top1-complexes (Alagoz et al., 2013), demonstrating a requirement for TDP1 to remove these trapped complexes. In addition to this, TDP1 has also been shown to contain a weak lyase activity that could be involved in the removal of AP sites and 3'-

dRP's that can form as a result of base excision repair (Lebedeva et al., 2012; Nilsen et al., 2012). This would mean that TDP1 could protect the cell in a Top1 independent manner and will be discussed in a later chapter.

In the presence of the DNA damaging agent MMS the DT40 Tdp1^{-/-} cells showed hypersensitivity compared to that of wild-type cells, whilst once again ectopic expression of the hTDP1 showed resistance to the drug to similar levels of the wild-type (**Figure 3.2.10 C**). Cell viability for wild-type cells and hTDP1 were 79 to 79 %, 65 to 58 %, 38 to 30 % and 17 to 13 % respectively, for increasing amounts of MMS. Interestingly, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ showed an intermediate level of protection with cell viabilities of 70 %, 37 %, 13 % and 6 %. Since TDP1 can remove stalled Top1 and has also been shown, *in vitro* and recently *in vivo*, to have a weak lyase activity (Lebedeva et al., 2012; Nilsen et al., 2012) it cannot be distinguished if this intermediate level of repair by the catalytic domain is due to defect in just Top1 associated repair or if the repair through the lyase activity is also somewhat compromised. To test this, knock down of Top1 in both full-length and the truncated cell line would need to be carried out, followed by treatment of these cells with MMS.

Next I examined the effects of ionising radiation (IR), which results in a multitude of damage: abasic sites, single-strand breaks terminated by 3'-phosphoglycolate esters as well as cytotoxic double-strand breaks (Ward, 1988), and TDP1 has been shown to play a role in protecting cells against IR (El-Khamisy et al., 2007; Inamdar et al., 2002). Unlike the other two DNA damaging agents hTDP1-complemented cells did not protect DT40 cells to the same level to that of the wild-type cells (**Figure 3.2.10 D**). This was also noted in the Pommier lab (Murai et al., 2012) and shows that hTDP1 does not fully function in all the repair pathways in the DT40 environment. It should also be noted that differences in cell protection during these drug treatments between cells lines are not due to a difference in expression of protein (**Figure 3.2.10 B**).

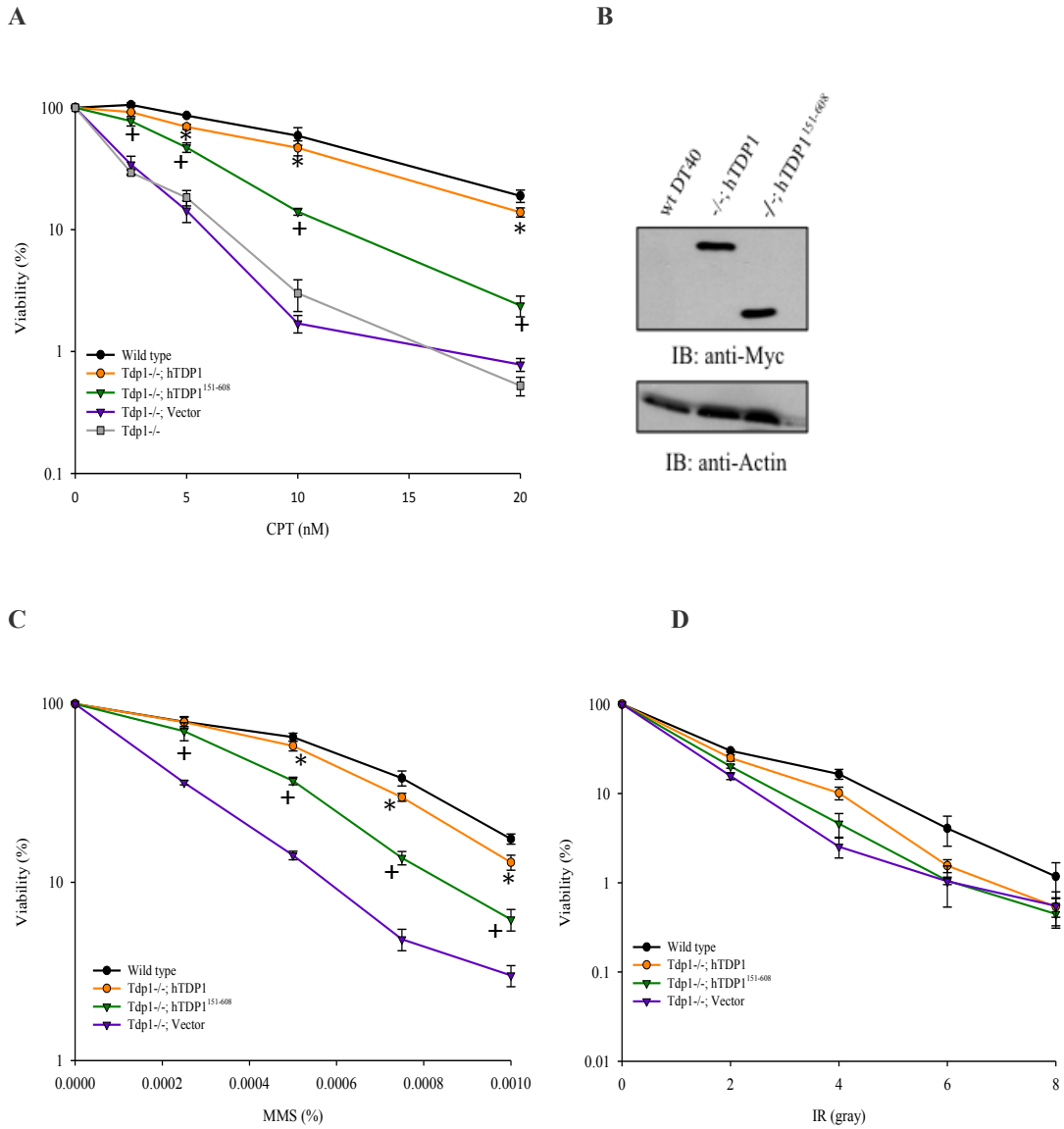


Figure 3.2.10 hTDP1 protects cells against a variety of DNA damaging agents, whilst the catalytic domain TDP1¹⁵¹⁻⁶⁰⁸ provides partial protection

Wild type DT40 cells and DT40 Tdp1^{-/-} expressing hTDP1 showed similar levels of protection against CPT (**A**) and MMS (**C**), but not IR (**D**). DT40 Tdp1^{-/-} cells were hyper sensitive to all 3 DNA damaging agents. Expression of hTDP1¹⁵¹⁻⁶⁰⁸ led to an intermediate protection against CPT and MMS and was significantly more sensitive when compared to cells complemented with hTDP1 (see asterisks). (**B**) Immunoblots of the Myc-TDP1¹⁵¹⁻⁶⁰⁸ and Myc-TDP1 stable cell lines probed with anti-Myc antibody (9B11; cell signaling) show no difference in expression levels of proteins. Error bars represent standard error of the mean of 3 independent experiments. Statistical analysis was done using the Student's t-test, '*' denote $P < 0.05$ between hTDP1 and TDP1¹⁵¹⁻⁶⁰⁸ and '+' denote $P < 0.05$ between TDP1¹⁵¹⁻⁶⁰⁸ and Tdp1^{-/-}; Myc.

3.2.3 Accumulation of DNA single-strand breaks in hTDP1¹⁵¹⁻⁶⁰⁸ compared to hTDP1, is in part transcription dependent

3.2.3.1 DT40 Tdp1^{-/-} cells, and cells complemented with hTDP1¹⁵¹⁻⁶⁰⁸ accumulate more single-strand breaks than full-length hTDP1

Since there is a difference in viability of cells between cell lines in the presence of DNA damaging agents the next step was to look at how the truncation and deletion of the TDP1 protein affects the repair of DNA at the molecular level. Alkaline Comet assays were employed to look at the response to CPT between these cell lines in terms of single-strand breaks (McArt et al., 2009). Cells expressing Myc-hTDP1, Myc-hTDP1¹⁵¹⁻⁶⁰⁸ or Myc were treated with camptothecin for 1 hour. Cells were then lysed, DNA denatured (in alkali conditions), and neutralised before being run in a comet tank at 25 V for 30 minutes. Tail moments were then counted under a microscope after the addition of SYBR green. As expected, Tdp1^{-/-} cells gave a significant increase in single-strand breaks compared to that of hTDP1-complemented DT40 cells with tail moments of 5.5 and 2.3 respectively, whilst treatment with DMSO had no effect (**Figure 3.2.11**). As in the viability assays, previously described, hTDP1¹⁵¹⁻⁶⁰⁸ demonstrated an intermediate effect, with less DNA single-stranded breaks than that of the knock out, but significantly more DNA single-stranded breaks than hTDP1 and resulted in a tail length of 3.2. The chicken genome sequence is about a third the size (1.2 billion bases) of the average mammalian genome (Homo sapiens 3 billion bases) and contains few macrochromosomes and many microchromosomes (Schmutz and Grimwood, 2004; Wallis et al., 2004). The difference between genome sizes could be responsible for the much smaller overall tail length seen in the alkaline comet assay when compared to human cell lines. Whilst TDP1¹⁵¹⁻⁶⁰⁸ only results in a small, but significant ($p < 0.005$; Student's t-test), increase in single-stranded DNA breaks the defects seen still lead to a large increase in viability.

3.2.3.2 DT40 Tdp1^{-/-} cells complemented with hTDP1¹⁵¹⁻⁶⁰⁸ accumulate more single-strand breaks than full-length hTDP1, and these appear to be transcription dependent

To try and narrow down the type of Top1-breaks, which the catalytic domain alone could not repair, the comet assay was repeated with the transcription inhibitor DRB (Yamaguchi-Iwai et al., 1999; Zandomeni et al., 1982). Interestingly, upon addition of DRB (Zeng et al., 2012), prior to treatment with CPT, the level of single-strand breaks in the hTDP1¹⁵¹⁻⁶⁰⁸ drops to that of the hTDP1, whilst the difference in CPT treated cells without DRB was still significant ($P < 0.005$, t-test) (**Figure 3.2.12**). This indicates that the sub-optimal repair by removing the N-terminal domain is due to a defect in the repair of Top1 associated transcriptional breaks.

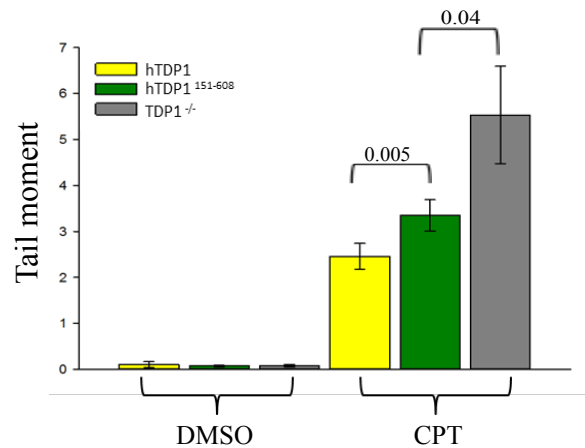


Figure 3.2.11 hTDP1¹⁵¹⁻⁶⁰⁸ accumulates a higher amount of DNA single-strand breaks than full-length TDP1

DT40 Tdp1^{-/-} cells and cells stably expressing Myc-hTDP1 or Myc-hTDP1¹⁵¹⁻⁶⁰⁸ were mock treated with DMSO or 50 μ M camptothecin ‘CPT’ for 1 hour and breaks quantified by the alkaline comet assay. The data presented is the average of $n = 3$ biological replicates \pm s.e.m. Results show that there is significant difference in the amount of breaks between ‘CPT’ treated hTDP1 and hTDP1¹⁵¹⁻⁶⁰⁸ (Student’s t-test; $P = 0.005$) and between hTDP1¹⁵¹⁻⁶⁰⁸ and Tdp1^{-/-} cells ($P = 0.04$).

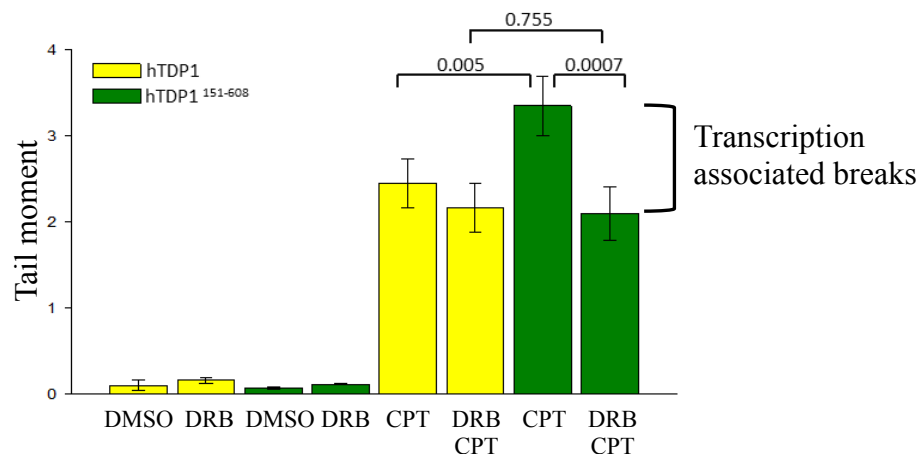


Figure 3.2.12 The increase in accumulation of DNA single-strand breaks in hTDP1¹⁵¹⁻⁶⁰⁸ compared to the full-length protein is, in part, transcription dependent

DT40 Tdp1^{-/-} cells and cells stably expressing Myc-hTDP1 or Myc-hTDP1¹⁵¹⁻⁶⁰⁸ were mock treated with DMSO or 50 μ M camptothecin ‘CPT’ for 1 hour; with or without previous incubation with 50 μ M DRB ‘DRB’ for 2 hours. Breaks quantified by the alkaline comet assay. Mean tail moments were quantified for 50 cells per sample per experiment, and the data presented is the average of $n = 3$ biological replicates \pm s.e.m. Statistics analysis was conducted using the Student’s t-test; results show that there is significant difference in the amount of breaks between ‘CPT’ treated hTDP1 and hTDP1¹⁵¹⁻⁶⁰⁸ (Student’s t-test; $P = 0.005$), and upon addition of DRB the significance of breaks between the full-length and truncated TDP1 is lost ($P = 0.755$).

3.3 Discussion

During neurogenesis the cells within the brain undergo rapid proliferation, migration and differentiation. This process accounts for the brain mass of the individual and diseases associated with this process lead to microcephaly (O'Driscoll et al., 2001; J. Shen et al., 2010). Once matured, cells become terminally differentiated and the requirement for replicative repair proteins is no longer required. These cells have to function and survive throughout the individual's life, and are not readily replaced. This then puts a high importance on the repair of transcription associated single-strand DNA breaks and protection against oxidative damage within these cells. The importance of this is typified by TDP1, as a H493R mutant is linked to SCAN-1, a neurodegenerative disease that leads to progressive degeneration of the cerebellum (El-Khamisy et al., 2005; Takashima et al., 2002). The phenotype of this disease is thought to be due to the lower activity of TDP1 H493R, thus lowering the capability to remove stalled Top1-DNA complexes in a timely manner. These un-resolved protein-linked single-stranded DNA breaks can lead to collision and stalling of transcription machinery and prevent progression of RNA polymerases to produce full-length proteins required for normal cellular activity. Over time accumulation of these breaks leads to cell death and the demise of post-mitotic tissue.

The results reported indicate a role for TDP1's N-terminal domain in the efficient and timely repair of transcriptionally associated breaks, highlighting the importance for this domain in the maintenance of post-mitotic cells. Whilst replication-associated double-strand breaks are the most cytotoxic to the cell, other enzymes, termed nucleases; such as the MRN complex, XPF/ERCC1, Mus81, CtIP and ARTEMIS (Connelly and Leach, 2004; Eid et al., 2010; Hamilton and Maizels, 2010; Nakamura et al., 2010; Hartsuiker et al, 2009) have been proposed to possess roles involved in the removal of stalled-Top1 complexes; however, these have primarily been implicated in double-strand break repair (Shrivastav et al., 2008). Therefore, the N-terminal domain of TDP1 could be of particular importance to non-dividing cells, especially neurons. This proposition is enforced by my data here showing that, removal of the N-terminal domain results in a

marked defect in the protection conferred by TDP1 in DT40 cells, following camptothecin and MMS damage.

We have shown for the first time that removal of the N-terminal domain does not impact upon activity of the enzyme *in vitro*, and that CD spectra and thermal denaturation profiling shows no difference in the unfolding pattern or core structure between the full-length TDP1 and catalytic domain. Thus, differences in ability to repair transcriptionally related Top1-linked DNA breaks, and potentially Top1 independent damage by MMS, is not due to loss of activity, change in structure or instability of TDP1¹⁵¹⁻⁶⁰⁸. These results demonstrate a requirement for the N-terminal domain of TDP1 in transcription associated DNA single-strand break repair and suggests a possible mode of action (**Figure 3.3.1**).

Since removal of the N-terminal domain does not change activity, structure or stability it is plausible that the non-optimal cellular protection against DNA damaging agents seen when the N-terminal domain of TDP1 is removed, could be due to loss of a protein to protein interaction, or post-translational modification within this domain. These in turn could allow increased localisation to sites of damage, efficiency in detecting damage, or an ability to localise other required proteins to sites of damage allowing faster repair kinetics within the cell.

One known interaction within the N-terminal domain of TDP1 is that of Lig III α but work discussed in chapter 5 and previously published work (Chiang et al., 2010; Das et al., 2009) demonstrates that loss of this interaction is not the reason for the defect in transcriptional associated TDP1 repair. Whilst working on this project an interaction of TDP1 with UBC9 (an E2 enzyme in the SUMO pathway) had been identified in our lab. I therefore tested the hypothesis that SUMO modification of TDP1 may occur at the N-terminal domain, potentially explaining some of the phenotypes presented in this chapter. This work will be the focus of my next chapter.

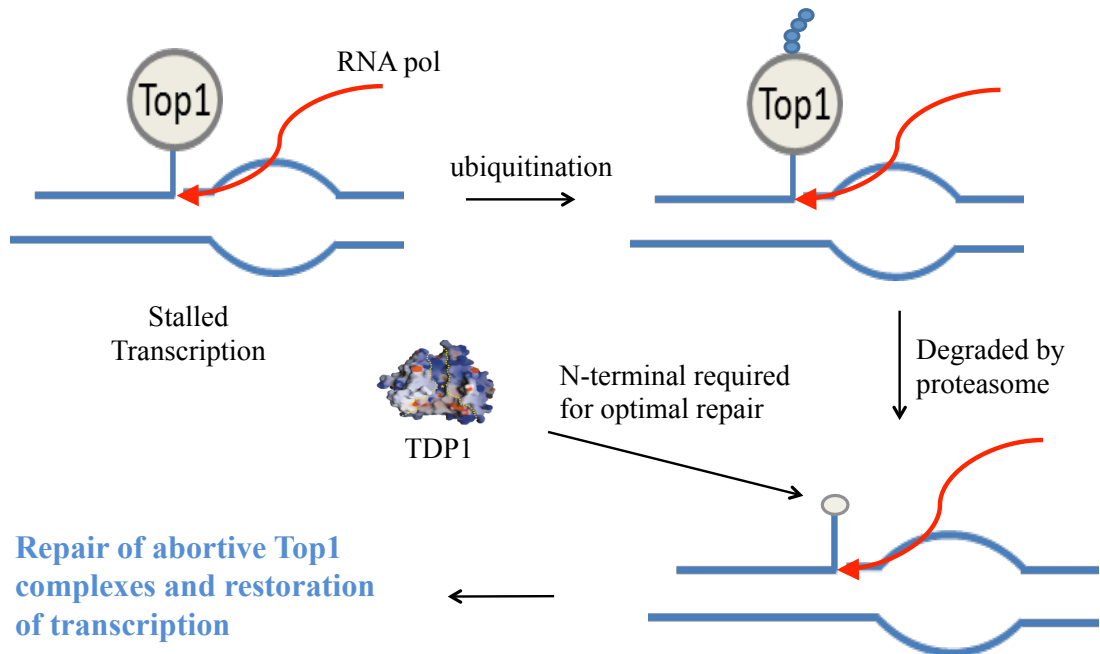


Figure 3.3.1 The N-terminal domain of TDP1 is required for optimal DNA single-strand break TDP1 mediated transcription repair

Collision of elongating RNA polymerase II with Top1 intermediates results in stalling of transcription and triggers partial degradation of Top1 to a small peptide. TDP1 can then remove the stalled Top1 peptide from DNA termini. The N-terminal domain of TDP1 allows for optimal repair and restoration of transcription. The mechanism by which the N-terminus of TDP1 aids the efficient removal of aborted Top1 complexes during transcription is not known, but is not due to an increase in catalytic activity of TDP1.

Chapter IV

Results 2: SUMOylation facilitates TDP1 driven chromosomal break repair

4.1 Introduction

4.1.2 Post-translational modification by SUMOylation

Whilst the previous chapter established a role for the N-terminal domain of TDP1 to provide optimal protection against DNA damaging agents through TDP1 mediated repair, primarily through transcriptional repair, it does not give an insight or mechanism of how this is provided.

Our lab identified an interaction between TDP1 and UBC9, the E2 ligase in the SUMO pathway. I therefore examined whether SUMOylation is a player in TDP1 mediated repair. This chapter shows that TDP1 is indeed modified by SUMO1, within the N-terminus at lysine 111 and that this repair is directed through transcriptional TDP1 mediated repair. Evidence for a non-covalent interaction between TDP1 and SUMO2/3 is also described in this chapter.

4.1.3 SUMO

Post-translational modifications, such as phosphorylation, acetylation, ubiquitylation and SUMOylation provide a dynamic platform for altering the surface area of proteins, allowing signalling and interactions with other macromolecules to occur adding a dynamic switch for regulation of proteins and other molecules within cells. These modifications are essential for cell maintenance, division and viability and allow an intricate method for control of protein function.

Small ubiquitin like modifier (SUMO) was discovered in 1996 (Boddy et al., 1996; Matunis et al., 1996; Okura et al., 1996; Z. Shen et al., 1996) and was named as ubiquitin-like protein because of the conserved ubiquitin-fold seen in the crystal structure (Bayer et al., 1998; Melchior, 2000). Although SUMO has a low sequence

homology to ubiquitin the SUMO pathway has many distinct features that are similar and can act as an antagonist to ubiquitylation (Bergink and Jentsch, 2009).

Several genes encoding SUMO proteins have been identified, only 3 subtypes are known to be involved in the DNA damage repair pathway; SUMO1, SUMO2 and SUMO3. SUMO1 shares little sequence homology to SUMO2 and 3 whilst the latter have a high similarity and can be thought of as a subclass within the SUMO family (Kamitani et al., 1998).

The interaction between proteins and members of the SUMO family has been observed via two distinct mechanisms: (1) SUMOylation, or (2) via SUMO interacting motifs (SIMS). SUMOylation allows the SUMO to covalently attach to a lysine residue within a protein through an isopeptide bond; the SIM interaction is a non-covalent one.

4.1.4 SUMOylation

Whilst the ubiquitin pathway contains a hierarchical structure: one E1 enzyme, several E2 enzymes and hundreds of E3 enzymes that confer selectivity to the target lysine of a particular protein, the SUMO pathway contains an E1 activating enzyme (AOS1/Uba2), but only one E2 conjugating enzyme (UBC9) and several E3 ligase enzymes (**Figure 4.1.1**). Unlike the ubiquitin pathway, *in vitro* at least, no E3 enzyme is needed for SUMO modification of the target protein and selectivity seems to follow a consensus sequence within the target molecule ΨKxE , where Ψ is a large hydrophobic residue and K the acceptor lysine (Bossis and Melchior, 2006). The E3 enzyme within the SUMO pathway has been proposed to increase the efficiency of binding of the E2-SUMO conjugate to the acceptor lysine (Melchior, 2000) and could possibly be attributed to SUMOylation of proteins to other lysine residues that do not conform to the above sequence, termed non-consensus sequences.

4.1.5 SUMO family

Both SUMO2 and SUMO3 proteins have similar homology and contain consensus sequences that allow SUMO2/3 poly chains to form, whilst SUMO1 only appears to

form poly-SUMO chains *in vitro* (Saitoh and Hinchey, 2000). All three SUMO proteins contain a conserved ubiquitin-fold common to ubiquitin-like proteins and the three dimensional folds of SUMO1 and ubiquitin can be superimposed (Melchior, 2000), furthermore both possess a C-terminal double glycine allowing covalent attachment to an acceptor lysine residue forming an isopeptide bond.

E1 enzyme

Although the Ubiquitin pathway contains one E1 protein the SUMO E1 enzyme consists of a heterodimer; Uba2 and AOS1, the AOS1 component of SUMO E1 is similar to that of Ubiquitin's E1 N-terminal and its' C-terminal domain has a likeness to that of Uba2, which also contains the active site cysteine (Desterro et al., 1999; Johnson et al., 1997; Okuma et al., 1999). This enzyme is known as the activation enzyme, it requires ATP to adenylate the C-terminal Glycine residue of SUMO. The formation of a thioester bond with a cysteine residue in Uba2 occurs through a transesterification reaction (Tatham, 2001). This complex then moves to the next stage of the cycle by transferring SUMO to the E2 enzyme (UBC9) (**Figure 4.1.1**).

E2 enzyme

UBC9 allows the transfer of SUMO onto itself and, *in vitro* at least, is the final step needed to SUMOylate a target protein (Okuma et al., 1999). The SUMO is transferred to the active site cysteine in UBC9 and once this has taken place UBC9 has been shown to contain a cleft that binds the consensus sequence ΨKxE in the target protein (Bernier-Villamor et al., 2002; Tatham et al., 2003). The fact that UBC9 is known to bind these sequences is the reason people believe that the E3 enzyme is not required *in vitro*. This interaction means that proteins can be screened for potential SUMOylation through the yeast two-hybrid by looking at interactions with UBC9.

E3 enzyme

E3 enzymes are a family of enzymes that are thought to be required, *in vivo*, to increase the efficiency of SUMOylation and facilitate the transfer of SUMO to consensus and non-consensus sites within the target substrate. The E1 and E2 enzymes are sufficient to cause SUMOylation *in vitro*, but E3 enzymes (PIAS family, Pc2, TOPORS and RanBP2 amongst others) have been identified and have demonstrated the capability to

enhance SUMOylation *in vitro* and *in vivo* (Johnson and A. A. Gupta, 2001; Kagey et al., 2003; Kirsh et al., 2002; Pichler et al., 2002; Weger et al., 2005).

SENPS

Sentrin specific proteases (SENPS) are the last component of the SUMO pathway and these enzymes have two roles within this enclosed cycle. Firstly, SUMO proteins are transcribed as precursor that needs to be cleaved to allow exposure of the di-glycine on the SUMO C-terminal tail. This di-glycine is required for the conjugation of SUMO to target proteins. Mutation of these glycine residues prevents covalent attachment of SUMO and are used *in vitro* as a negative control. Secondly, the SENPS are needed to remove SUMO from the target protein to restore the free pool of SUMO maintaining the equilibrium of SUMOylated and non-SUMOylated proteins. On top of this SUMOylation can also be regulated through phospho-dependent SUMO motifs (PDSM'S) (Hietakangas et al., 2006).

4.1.6 Phospho-dependent SUMO motif

Other post-translational modifications, or sequences of negative amino acids proximal to the acceptor lysine can help regulate the levels of SUMOylation and act as a molecular switch for SUMO modification to occur. These sites are termed phosphorylation-dependent SUMOylation motifs and negatively charged amino acid-dependent SUMOylation motifs (Anckar and Sistonen, 2007). The ability of other modifications to regulate SUMO modification allows an intricate pathway to allow interactions of other post-translational modifications such as ubiquitin and acetylation that may occur on the same lysine residue.

4.1.7 SUMO interacting motifs – SIMS

Proteins cannot only be modified covalently by the SUMO family but can also interact non-covalently; this occurs through SUMO interacting motifs (SIMS). This discovery meant that SUMOylation of one protein could create an interface for interaction with other proteins (Hannich et al., 2005; Minty et al., 2000). First characterised by Minty *et al* an SXS triplet was thought to be crucial for SIM interactions (Minty et al., 2000).

Further studies elucidated a core hydrophobic region consisting of the motif **V/I-X-V/I-V/I** (J. Song et al., 2004). Flanking acidic regions were also later discovered to play a role in binding (Hannich et al., 2005; Hecker et al., 2006). Even if SIM interacting proteins have been observed to interact with SUMO2 and 3 it does not necessarily mean that they will interact with SUMO1, data by Hecker *et al* suggests that this could be down to charge of flanking proteins proximal and distal to the SIM site, and lists multiple proteins with confirmed SIM sites and their recognition motif (Hecker et al., 2006).

4.1.8 Molecular consequences of SUMOylation and SUMO-SIM interactions

SUMOylation of a particular protein can have multiple outcomes that are difficult to predict (**Figure 4.1.2**). The effect can range from: changes in cellular localisation as seen in XRCC4 (Bergink and Jentsch, 2009), as an antagonist to ubiquitylation and, therefore protection against ubiquitin-mediated degradation, such is the case for I κ B α (Desterro et al., 1998), to attenuation of transcription by the recruitment of histone deacetylases (HDAC) (Shiio and Eisenman, 2003; S.-H. Yang and Sharrocks, 2004). Perhaps the most interesting outcome of SUMOylation was found for thymine DNA glycosylase (TDG). TDG is involved in the base excision repair pathway and is required to bind and excise mismatched bases. TDG binds tightly to the reaction product, the abasic site, and becomes trapped. SUMOylation causes a conformational change by interacting with an internal SIM site in TDG, thus reduces its affinity to DNA, releasing the enzyme into the nucleoplasm. SENPS then remove SUMO, enabling TDG to target new mismatched DNA (Baba et al., 2005; Hardeland et al., 2002; Smet-Nocca et al., 2011).

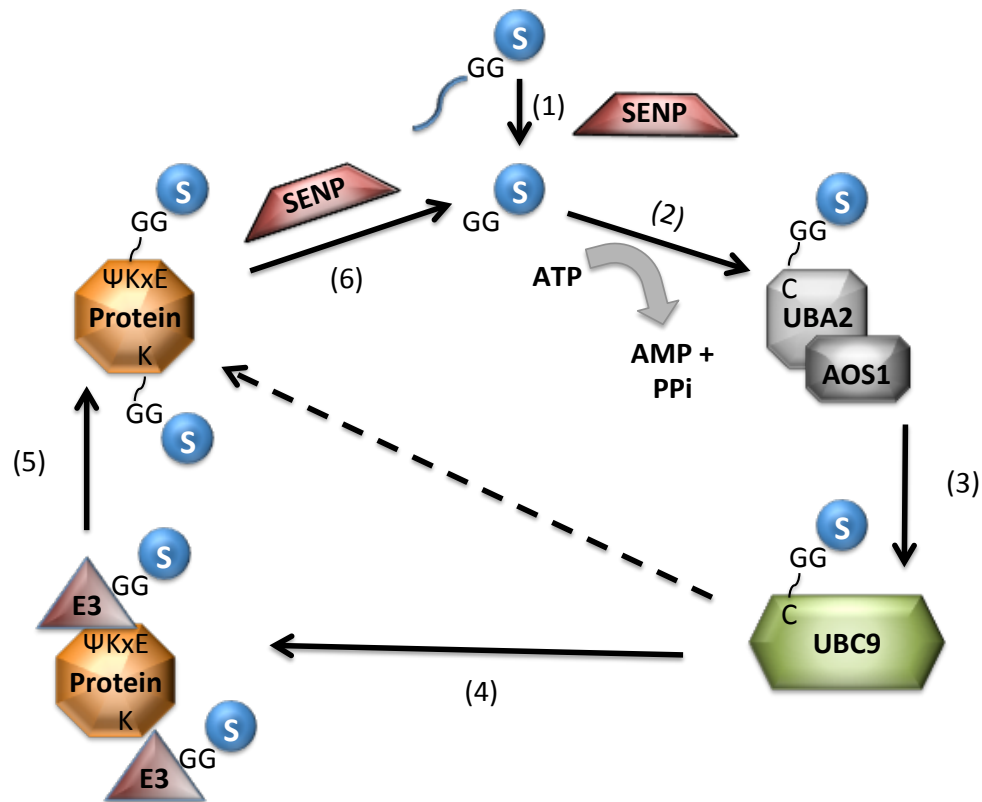


Figure 4.1.1 SUMO conjugation and de-conjugation of target substrates

(1) SUMO is translated as a precursor protein with a short C-terminal extension, which is cleaved off by SENP proteins to reveal a di-glycine C-terminal. (2) The mature SUMO is activated by the heterodimeric E1 enzyme (Uba2-AOS1). Activation is ATP dependent, and proceeds via a SUMO-adenylate intermediate, resulting in a covalent thioester linkage between SUMO and a cysteine residue in Uba2. Activated SUMO is then transferred to a cysteine in UBC9, the E2 conjugating enzyme (3). UBC9 recognises and binds the consensus ΨKxE in the target protein and is sufficient to SUMOylate the target protein *in vitro* (dashed line). (4) *In vivo*, in conjunction with one of many substrate recognising E3 SUMO ligases SUMO is transferred to a specific lysine residue within the target protein forming an isopeptide bond between the SUMO glycine tail and the acceptor lysine (5). The E3 enzyme is thought to allow for SUMOylation at non-consensus sites allowing for a wider range of targets to be SUMOylated. (6) De-conjugation of SUMO from the target substrate occurs through sentrin specific proteases (SENP) to allow SUMOylation to be a reversible and dynamic process and to maintain a free pool of SUMO proteins. Figure is adapted from (Dohmen, 2004; Geiss-Friedlander and Melchior, 2007).

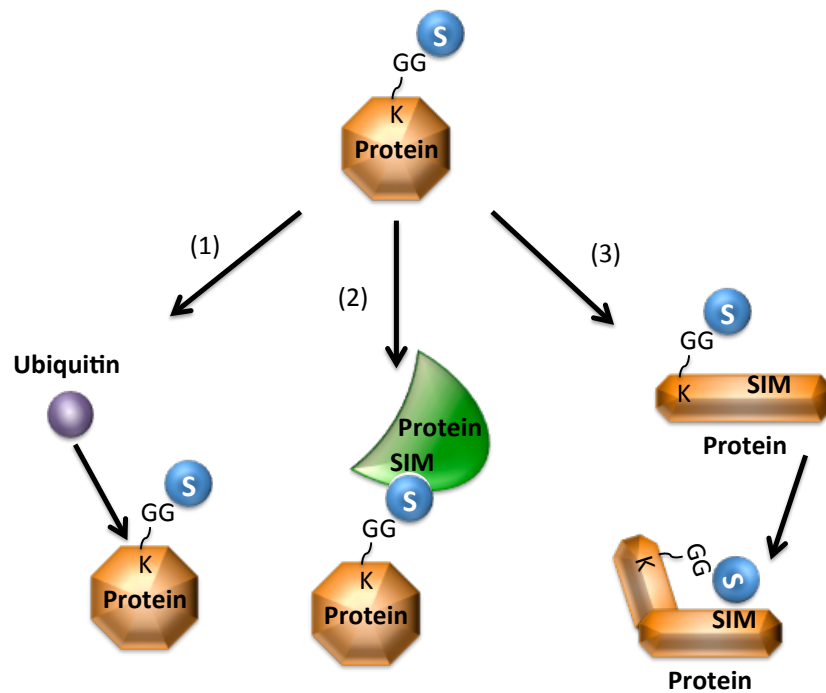


Figure 4.1.2 Molecular consequences of SUMOylation and SUMO-SIM interactions

Sumoylation can have several consequences for the modified protein. (1) SUMOylation can block interactions with target proteins and chemical modifications. One example of this is ubiquitylation that also occurs at lysine residues within a protein. Inhibition of ubiquitylation can prevent degradation of a particular protein. (2) SUMOylation can provide a binding site for other proteins via a SUMO-SIM interaction, this allows reversible and controlled interactions with other proteins. Finally, SUMOylation can result in a conformational change of the target protein, this occurs through interaction of the SUMO with an internal SIM site within modified protein (3). Adapted from (Geiss-Friedlander and Melchior, 2007)

4.1.9 Chapter objectives

This chapter investigates the mechanisms, if any, by which TDP1 is regulated covalently or non-covalently by SUMOylation. Previous work in our lab has shown that TDP1 interacts with UBC9 using the yeast-two hybrid system. Does this mean SUMOylation regulates TDP1? To investigate this we first tested if a member of the SUMO family modifies TDP1, and then studied the role this may have in cells. In order to do this we mapped the TDP1 site/s of covalent attachment by SUMO, mutated these sites to prevent SUMOylation, and investigated what role SUMOylation of TDP1 has in the cell. To demonstrate that any effect seen was due to ablation of SUMOylation it was also necessary to rule out that the mutations in TDP1 does not affect the activity, stability or folding of the protein.

I then examined if TDP1 could interact non-covalently with the SUMO family via internal SIM/s, and if so where these motifs are located in TDP1. These data showed for the first time that TDP1 is SUMOylated by SUMO1 and that this is required for localization of TDP1 to sites of DNA damage, that are in part transcription dependent.

4.2 Results

4.2.1 TDP1 is SUMOylated at lysine 111 by SUMO1 *in vitro* and *vivo*

4.2.1.1 TDP1 is SUMOylated *in vitro*

In order to determine whether TDP1 could be modified by SUMO1, I expressed and purified recombinant His-hTDP1 in bacteria. Recombinant TDP1 was then incubated with components of the SUMO pathway (**Figure 4.2.1**); SAE1/SAE2 (50 nM), UBE2I (500 nM) and either 30 μ M of SUMO1, SUMO2 SUMO3 or mutant versions, which harbours a di-glycine to di-alanine mutation to prevent covalent SUMO conjugation. Immunoblots were conducted using anti-TDP1 (Abcam) and anti-SUMO1, 2/3 antibodies (Activmotif). **Figure 4.2.2 B** shows that TDP1 antibodies cross-reacted with a 72 kDa protein (representing unmodified TDP1) and one discrete prominent band was observed (~105 kDa), consistent with SUMO modification. To consolidate this, blots were additionally probed with anti-SUMO1, and SUMO2/3 antibodies. Notably, the ~105 kDa band was strongly detected by anti-SUMO1, suggesting that TDP1 is primarily SUMOylated by SUMO1. Multiple faint bands were also detected in the anti-SUMO blots when TDP1 was incubated in the presence of SUMO1, 2 or 3. These higher molecular weight bands were not detected in the mutant SUMO controls 'MT', suggesting that these bands were dependent on covalent modification by the SUMO proteins. Since these higher molecular weight bands were observed in the anti-SUMO blots but not the anti-TDP1 blot it is probable that they are due to contaminant proteins co-purified with recombinant TDP1 that are also targets for SUMOylation.

Since p53 is known to be SUMOylated (Gostissa et al., 1999; Rodriguez et al., 1999), I employed recombinant p53 as a positive control. I observed a higher molecular weight band when p53 was incubated with SUMO1, but no band was detected when incubated with mutant SUMO1 (**Figure 4.2.2 A**). Thus, the reaction conditions enabled p53 to be covalently modified and that the di-glycine mutation was sufficient to ablate covalent attachment.

Although higher molecular weight bands are detectable when TDP1 was incubated with any of the WT SUMO proteins, quantification by ImageQuant of the slower migrating bands versus the total TDP1 protein showed that more than twice as much TDP1 was modified by SUMO1 (26 %) compared to SUMO2 (12 %) and SUMO3 (6 %) (**Figure 4.2.2 C**). Taken together, these results suggest that TDP1 is primarily SUMOylated by SUMO1 *in vitro*. It is important to note that SUMO1 has a predicted molecular weight of ~11 kDa although it actually migrates at ~17.5 kDa on SDS-PAGE (Antoine *et al* 2005); this means that if TDP1 is modified by one SUMO moiety at one lysine residue this should give a combined molecular weight of 89.5 kDa. Three possible scenarios might produce a SUMO-modified TDP1 of ~105 kDa: (i) two distinct lysine residues within TDP1 are equally modified at the same time ($17.5 + 17.5 + 72 = 107$ kDa), (ii) a di-SUMO conjugate is linked to a single lysine ($((17.5 \times 2) + 72 = 107$ kDa) or (iii) that the addition of a single SUMO1 retards the gel mobility of TDP1 in a greater than additive manner. We favour the latter explanation since mobility in SDS-PAGE for branched peptides or proteins is not linear, and largely depends on the spatial conformation adopted by the peptide.

Unlike SUMO2 and SUMO3, SUMO1 does not contain the ψ KXE sequence necessary to form SUMO polymers and, as a result, can only form a monomer *in vivo* (Tatham, 2001). However, it is worth noting that *in vitro* experiments have demonstrated that poly-SUMO1 chains can form, although this is probably an artefact due to non-physiological amounts of SUMO1 being present (Bossis and Melchior, 2006; Ulrich, 2008). If this is the case, then the higher molecular weight band in **Figure 4.2.2 B** could be attributed to aberrant di-SUMO1 chain formation. However, I consider the third possibility outlined above as the most likely explanation as I did not observe any band at ~90 kDa that would fit with the modification of TDP1 at a single lysine with a single SUMO1 migrating at anticipated molecular weights.

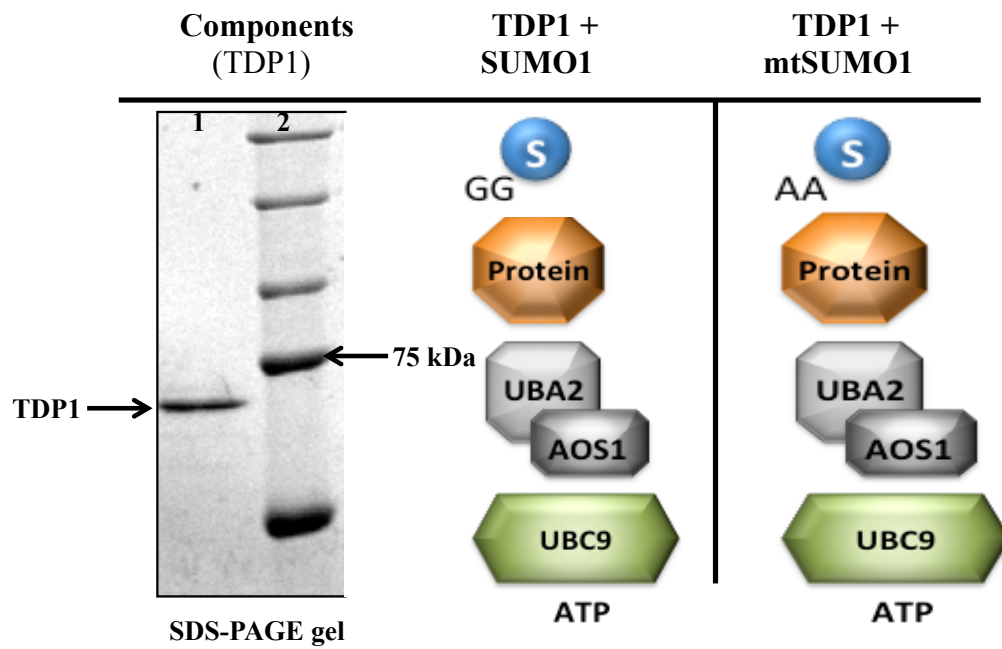


Figure 4.2.1 *In vitro* components of the SUMOylation assay

SUMOylation of protein requires: recombinant TDP1 (lane 1 shows purified TDP1 on an SDS-PAGE gel), E1 enzyme (SAE1/SAE2), E2 enzyme (UBE2I) and SUMO -1/2 or 3. The reaction is energy dependent and requires ATP. For a control a mutant SUMO containing a di-glycine to alanine mutation was used. This mutant prevents the isopeptide bond forming between the SUMOylated target lysine and glycine of SUMO.

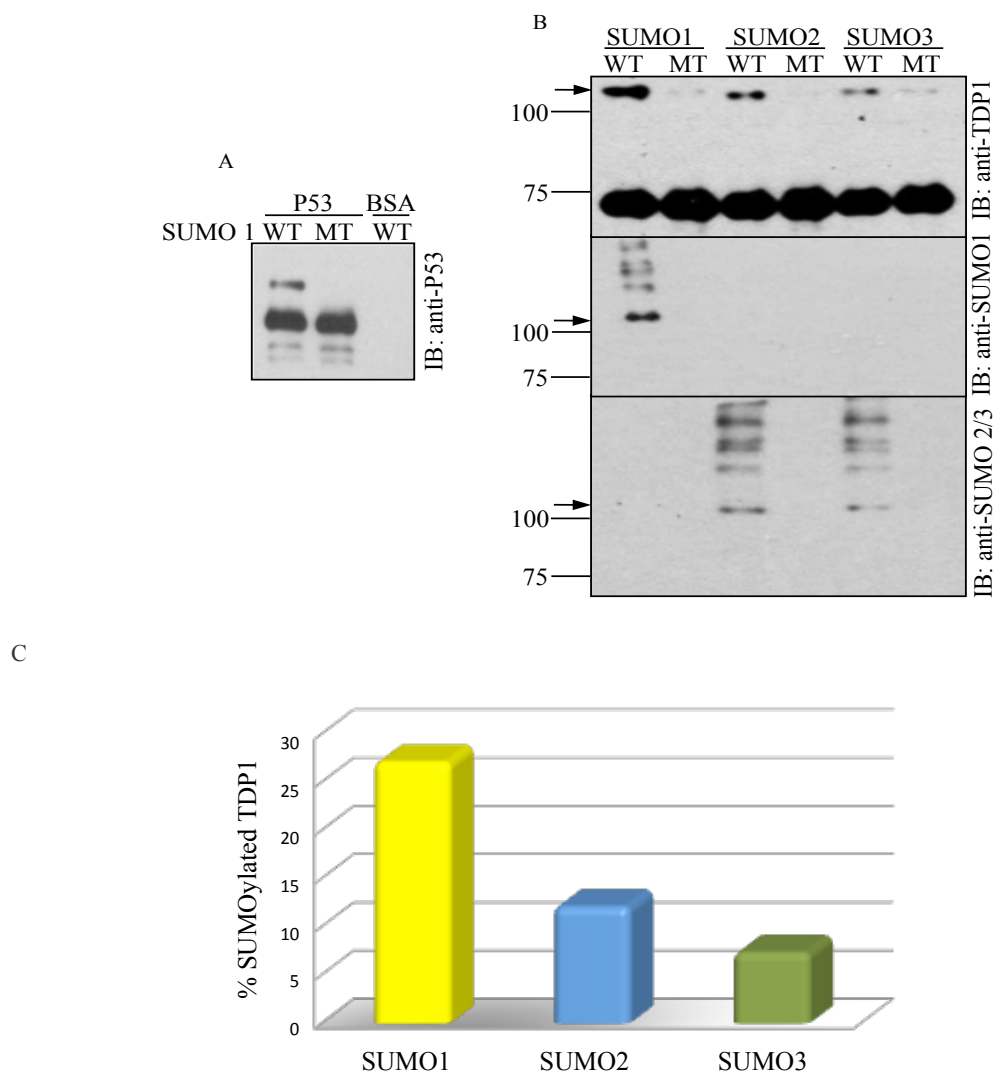


Figure 4.2.2 TDP1 is primarily SUMOylated by SUMO1 *in vitro*

(A) Incubation of P53, a known target for SUMOylation, with SUMO1 results in a slower migrating band (SUMOylated P53) that could not be detected in the mutant SUMO1 reaction 'MT'. This demonstrates that the reaction conditions used were suitable to obtain SUMOylated protein. (B) TDP1 is detected at 72 kDa with anti-TDP1. Slower migrating bands are also present when SUMO1, 2 or 3 are included in the reaction (marked by arrows, above 100 kDa). These higher molecular weight bands were not present in the mutant SUMO reactions 'MT' and immunoblots with anti-SUMO1, 2/3 confirmed the bands to be SUMOylated TDP1. (C) Quantification, using ImageQuant (GE Healthcare), of higher molecular weight pixel volume in SUMO1, 2 and 3 reactions vs total pixel volume of TDP1 in 'MT' lanes in the anti-TDP1 blot showed that TDP1 is primarily SUMOylated by SUMO1 *in vitro*.

4.2.1.2 TDP1 SUMOylation occurs within the N-terminus *in vitro*

To identify potential lysine residues for SUMO conjugation, I entered the TDP1 protein sequence into the SUMOsp 2.0 (<http://sumosp.biocuckoo.org/online.php>) online server. The server scans for consensus **Ψ-K-X-E** and non-consensus motifs. Non-consensus motifs are previously identified sites where SUMOylation can occur but does not conform to the expected consensus region (Ren *et al* 2009). Five potential sites in TDP1 were identified by *in silico* analysis (**Figure 4.2.3**). These were lysine residues at amino acid positions 111, 139, 231, 417 and 527. The first two residues conformed to the consensus motifs; both present in the N-terminal domain, whilst the last three were found in the catalytic domain and were type II non-consensus sites. It is important to note that covalent modification by SUMO does not automatically occur at these consensus sites and are used as a prediction to narrow down the most likely sites of SUMOylation

To isolate the region of TDP1 where SUMOylation may occur, I next introduced N- and C-terminal domain truncations into TDP1 by PCR-mutagenesis; these were expressed and purified, resulting in His tagged N- and C-terminal TDP1 domain fragments (see Methods for details) and concentrations of proteins were confirmed by Nanodrop (Thermo Scientific). Once purified, both proteins were used in the *in vitro* SUMOylation assay (as above) and were analysed via immunoblots. **Figure 4.2.4 B** shows that *in vitro* SUMOylation of N-terminal TDP1 does occur, as illustrated by anti-TDP1 blots. There was a prominent intense band just below 50 kDa that was not present when the mutant SUMO1 was used. The C-terminal TDP1-SUMO1 reaction resulted in a weaker band above the band corresponding to TDP1¹⁵¹⁻⁶⁰⁸ ~ 75 kDa, which was not present in the SUMO1 mutant reaction (**Figure 4.2.4 A**). It is important to point out that the Santa Cruz antibody used in the TDP1¹⁵¹⁻⁶⁰⁸ blot results in a very cloudy blot and therefore it is possible, although unlikely that SUMOylation can occur.

SUMOsp 2.0 was used as an online database to predict potential sites for SUMO modification .

Prediction of consensus sites, Ψ -K-X-E: K111 and 139

Prediction of non-consensus sites: K231, K417 and K527

This analysis highlighted 5 potential sites:

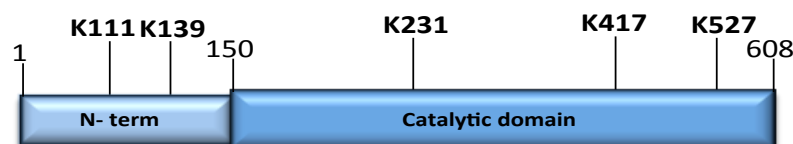


Figure 4.2.3 *In silico* analysis highlights 5 potential sites for SUMOylation

The amino acid sequence of full-length hTDP1 was entered into the SUMOsp2.0 database (Ren et al., 2009). *In silico* analysis predicted 2 consensus sites, K111 and K139, corresponding to the Ψ -K-X-E amino acid sequence normally required for SUMOylation. The scan also recognised 3 non-consensus sites, all in the catalytic domain at lysine positions 231, 417 and 527.

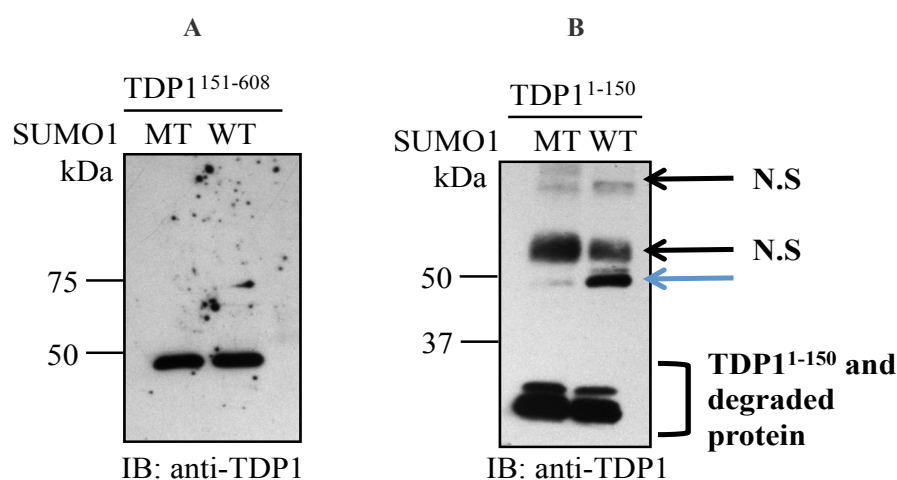


Figure 4.2.4 TDP1 is SUMOylated in the N-terminus by SUMO1 *in vitro*

500 nM recombinant His-TDP1¹⁵¹⁻⁶⁰⁸ (A), or His-TDP1¹⁻¹⁵⁰ (B) were incubated with SUMO1 wild-type 'WT', or SUMO1 harboring a di-alanine mutant 'MT' and the components of the SUMO pathway. TDP1¹⁵¹⁻⁶⁰⁸ probed with anti-TDP1 (Santa Cruz) is seen just below 50 kDa. There are no other prominent bands in the WT lane when compared to the MT lane. This indicates that TDP1¹⁵¹⁻⁶⁰⁸ is not a target for SUMOylation. (B) TDP1¹⁻¹⁵⁰ was probed with anti-TDP1 (Abcam) and reveals two bands at 20 kDa, which is probably due to degradation of the recombinant truncated protein. There is also a prominent band at 50 kDa in the WT lane that (denoted by a blue arrow) that is not found in the MT lane consistent with posttranslational modification by SUMO. There are several bands above 50 kDa found in both the SUMO MT and WT lanes that are due to non-specific binding of the antibody (N.S).

4.2.1.3 TDP1 SUMOylation occurs within the N-terminus at K111 *in vitro*

To study what effect SUMOylation might have on TDP1, I set out to identify at what site or sites the post-translational modification occurs. Mutation of the acceptor lysine to an arginine has been shown to prevent the covalent attachment of SUMO (Babic *et al* 2006, Bachant *et al* 2002). It is therefore possible to establish the site, or sites of SUMOylation. Due to the unlikelihood of modification occurring in the C-terminal domain, only lysines in the N-terminal region of TDP1 were subjected to mutational analysis, K111R, K139R and K111R+K139R mutations were introduced into TDP1 by site directed mutagenesis, confirmed by sequencing, expressed, purified and analysed by Coomassie stain to confirm purity. K111R, K139R and K111/139R mutants were then analysed to see if either or both mutations affected the extent of SUMOylation.

The wild-type TDP1¹⁻¹⁵⁰ displayed a SUMO1-modified band at ~50 kDa (indicated by the black arrow), consistent with the 25-30 kDa increase in size with SUMOylation previously seen with full-length TDP1 (**Figure 4.2.5**). Interestingly, the K111R mutant showed a significantly reduced level of SUMO-modification compared to the wild-type, while the double K111R + K139R mutant showed no detectable SUMO-modification at all. Unexpectedly, the K139R single mutant actually showed a small increase in overall SUMOylation relative to wild type. This data suggests that K111 is the primary SUMO-acceptor lysine in TDP1 (*in vitro*) and that K139, although peripherally influencing K111 SUMOylation is not in itself a primary site for modification.

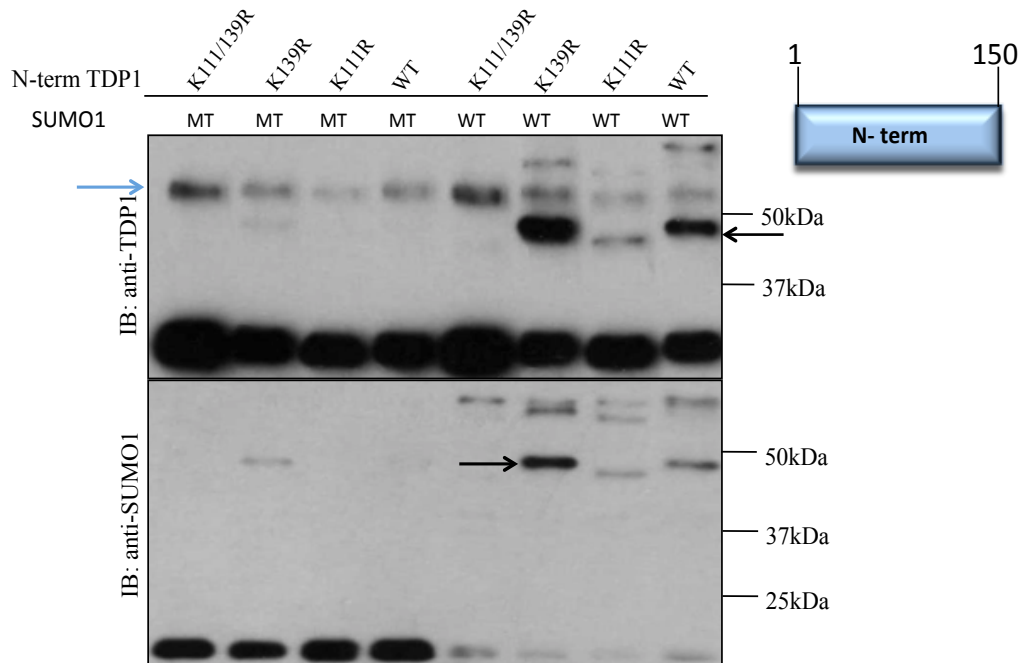


Figure 4.2.5 N-terminal TDP1 mutants show varying amounts of SUMOylation

500 nM of recombinant His-N-terminal TDP1 protein and mutants were added to the *in vitro* SUMOylation reaction. In the anti-TDP1 blot there was an unidentified band above 50 kDa in all lanes, indicated by a blue arrow. This was not picked up in the anti-SUMO1 blot and is probably due to non-specific binding of the protein. The K139R TDP1 mutant actually showed an increased in the amount of SUMOylated TDP1 at 48 kDa (indicated by a black arrow), whilst the K111R mutant shows almost complete abrogation of modified protein. This was picked up in both anti-TDP1 and SUMO1 blots and was not present in the SUMO1mt lanes 'MT' indicating that TDP1 K111R is the main acceptor lysine for SUMOylation.

4.2.1.4 Mutation of lysine to arginine at 111 abrogates SUMOylation *in vitro* and *in vivo*

In order to examine the hypothesis that K111 is the primary SUMOylation site, a K111R mutation in full-length TDP1 was generated, using the same methods described above for the truncated versions. The purified protein was then subject to the *in vitro* SUMOylation assay and immunoblot analysis confirmed that K111R mutation successfully prevents covalent modification by SUMO1 in the full-length TDP1 protein (**Figure 4.2.6 A**). In conjunction with this work, HEK293 cells (Jude) were co-transfected with Myc-TDP1 and GFP or GFP1-SUMO1, and Myc-TDP1 K111R with GFP or GFP1-SUMO1. Cells were lysed and protein was fractionated on an SDS-PAGE gel. Immunoblot analysis indicated a higher molecular weight band, consistent with SUMOylated TDP1, in the TDP1 + GFP-SUMO1 that was not detectable in the GFP + Myc-TDP1 fraction, or the TDP1 K111R fraction (**Figure 4.2.5 B**). Taken together this suggests that TDP1 is modified by SUMO1 at lysine 111 *in vitro* and *in vivo*. Blast sequence analysis of the N-terminal domain, using uniProt, demonstrates that the acceptor lysine residue and SUMO consensus was only conserved in higher eukaryotes (**Figure 4.2.7**). Combined with data from the first results chapter, one possible role for the N-terminal evolutionarily driven domain is to channel TDP1 driven reactions to deal with breaks associated with transcription, in a SUMOylation regulated process.

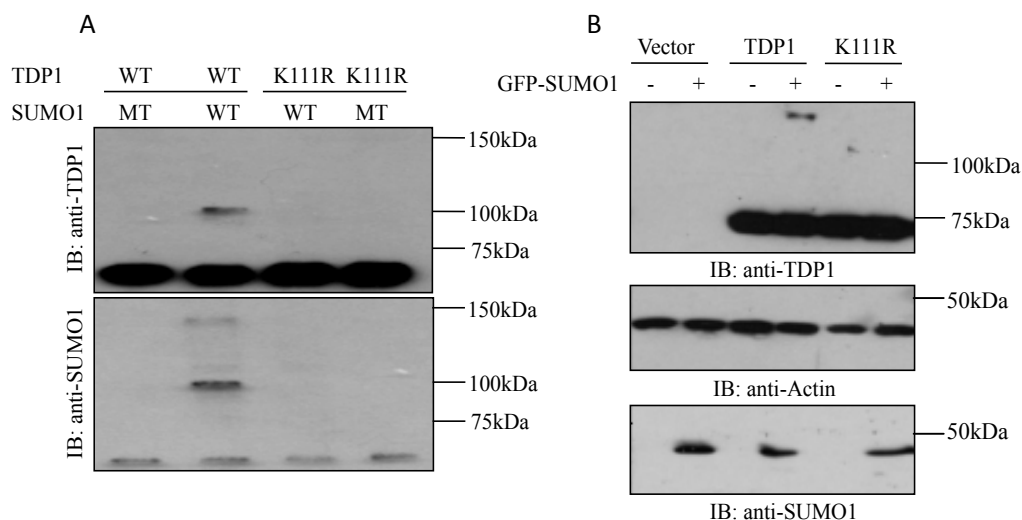


Figure 4.2.6 TDP1 is SUMOylated at lysine 111 *in vitro* and *in vivo*

(A) Recombinant His-TDP1 ‘WT’ or TDP1 K111R ‘K111R’ were added to the SUMOylation reaction and probed with anti-TDP1 (Abcam) and anti-SUMO1 (Activmotif) (B) HEK293 cells transfected with the indicated TDP1 constructs and an empty GFP ‘-’ or GFP1-SUMO1 ‘+’ vector were fractionated on an SDS-PAGE gel and analysed with anti-TDP1 (Abcam), anti-SUMO1 (Santa Cruz) and anti-Actin (Sigma). Both immunoblots show that mutation of lysine 111 to arginine in TDP1 is sufficient to prevent modification by SUMO1.

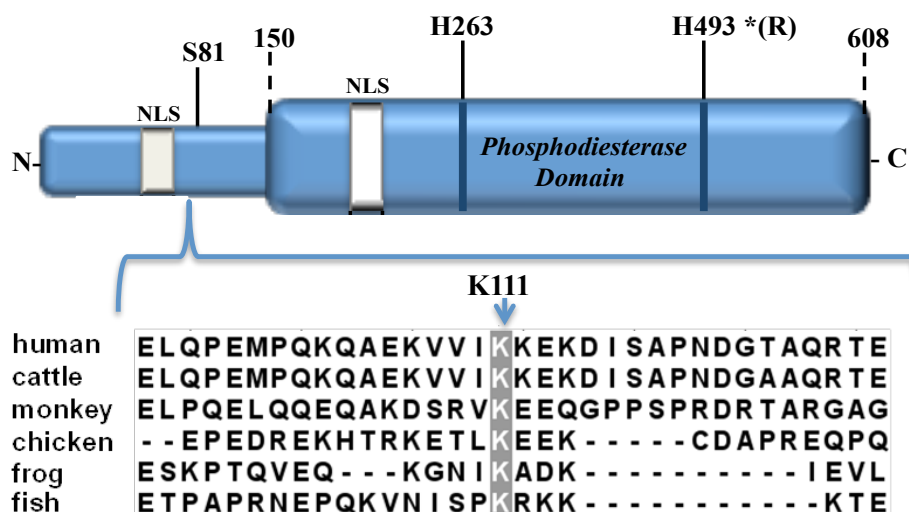


Figure 4.2.7 Lysine at 111 is conserved in several higher eukaryotes

BLAST sequence analysis of the TDP1 N-terminal domain identifies the SUMOylation consensus site with the acceptor K111 in humans (Homo sapiens; NP_001008744), is conserved in cattle (Bos taurus; NP_001180084; XP_874680), monkey (Pongo abelii; XP_002825063), chicken (Gallus gallus; XP_421313), frog (Xenopus tropicalis; NP_001039242), fish (Danio rerio; XP_700174) using uniProt software.

4.2.2 Mutation of lysine 111 to arginine does not alter protein structure, stability or catalytic activity

Before examining what role SUMOylation of TDP1 may have, in a cellular context, we wanted to examine if the lysine to arginine mutation may have an impact on TDP1, in terms of activity, stability and structure.

4.2.2.1 TDP1 and TDP1^{K111R} give similar melting temperatures and unfolding patterns

Recombinant wild-type and K111R TDP1 were subjected to thermal denaturation profiling; a technique used to determine the stability and the unfolding profile of proteins (Babic et al., 2006; Rehman et al., 2011). Fluorescence values were obtained using the lightcycler 480 (Roche) and melting points determined using a variation of the Boltzmann model (Ericsson et al., 2006). TDP1 and TDP1^{K111R} showed a similar unfolding pattern and resulted in T_m's of 55.5 and 54.3 °C for the mutant and wild-type respectively, indicating no detectable difference in the unfolding pattern between the two proteins (**Figure 4.2.8 A**).

4.2.2.2 TDP1^{K111R} does not result in a reduced activity compared to wild-type

To compare the activity of the wild-type and K111R TDP1, decreasing concentrations (30, 15, 7 and 3 nM) of purified recombinant TDP1 or TDP1 K111R were incubated for 1 hour at 37 °C with a ³²P-radiolabelled duplex-nicked substrate (50 nM) harbouring a 3' phosphotyrosine 'PY' at the nick (**Figure 4.2.8 B**). The 3'-phosphotyrosine mimics the abortive complex of Top1, thus is a substrate for TDP1. If TDP1 is added to the reaction then hydrolysis of the phosphotyrosine bond should occur, leaving a 3' phosphate which results in a faster migrating species termed 'repair products'. Repair products were analysed by denaturing PAGE and phosphorimaging and reaction products 'P' were quantified relative to total labelled substrate 'P + PY' and percentage conversion to 3'-P. The wild-type TDP1 and K111R both demonstrate the same enzymatic capacity to hydrolyse the 3' phosphotyrosine to a 3'phosphate (**Figure 4.2.8 C**). This suggests that the missense mutation does not change the activity of TDP1.

4.2.2.3 TDP1^{K111R} and TDP1 exhibit no significant structural differences

Comparison of the TDP1 and TDP1^{K111R} through circular dichroism spectroscopy, a well-known method for structural comparison (Lawton, 2002), resulted in two similar spectral graphs with negative bands at ~210 and 220 nm, and positive bands at ~195 nm (**Figure 4.2.9 A**). High tension voltage for both spectra was under 600 (**Figure 4.2.9 B**) at the lowest wavelength; a value required for accurate interpretation of data (Kelly et al., 2005) and subtraction of the K111R plot from the wild-type gave a spectra not associated with any structure (**Figure 4.2.9 C**). To obtain a quantitative read-out analyses of secondary structure was carried out using DichroWeb, by employing the variable selection algorithm (CDSSTR), which provides superior fits for globular proteins (Greenfield, 2006; Lobley et al., 2002; Perez-Iratxeta and Andrade-Navarro, 2008). This analysis revealed no significant difference in α -helical or β -sheet content ($P > 0.5$; t -test), suggesting no apparent change in structure (**Table 4.2.1**). This was also supported by the CONTIN and K2D algorithms (J. J. R. Hudson et al., 2012). From these analyses I concluded that there was no significant change in structure, unfolding, or catalytic activity of the mutant protein compared to the wild-type and that any cellular phenotype that might be observed is likely due to TDP1 K111R's inability to be modified by SUMO1.

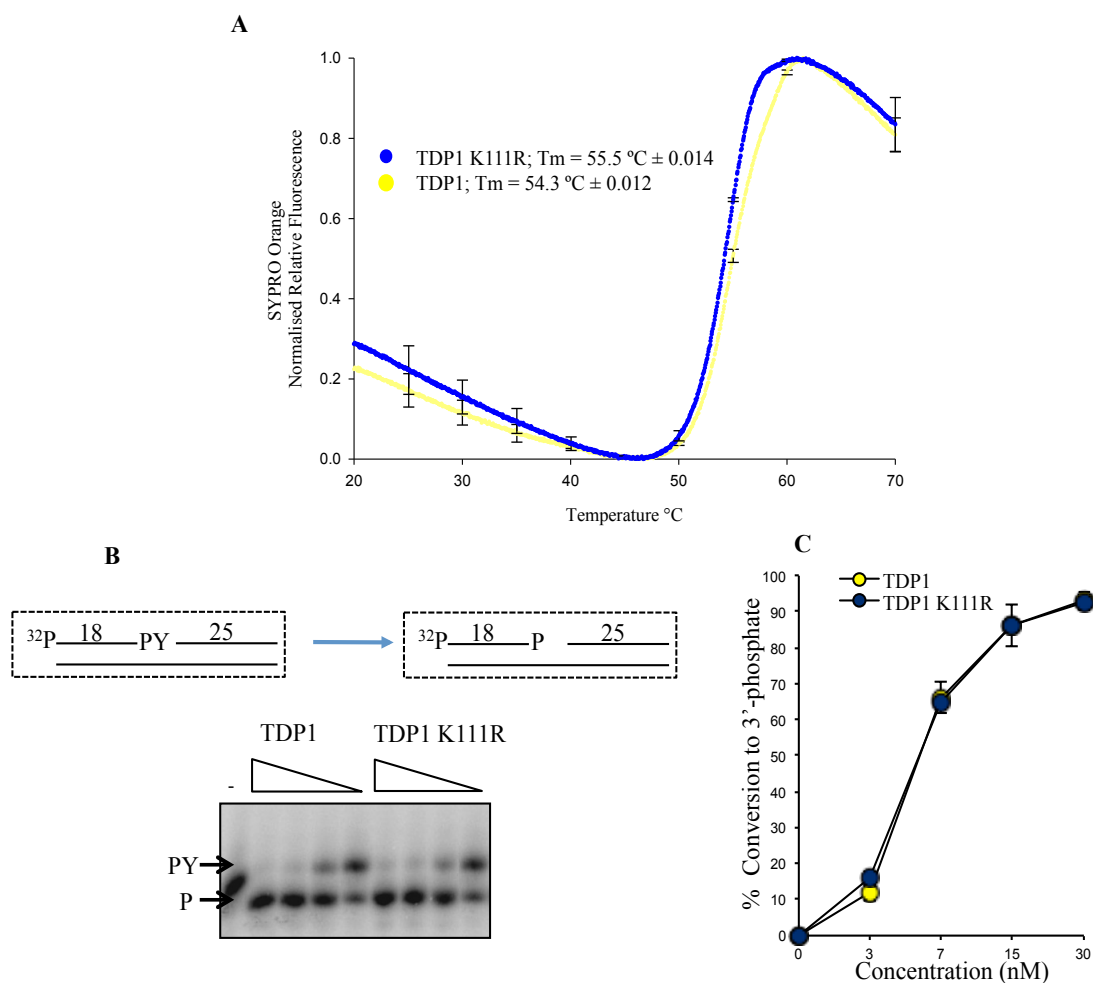


Figure 4.2.8 Removal of the N-terminal domain does not impact upon activity *in vivo*

(A) 0.75 μM of recombinant hTDP1 or hTDP1 K111R was mixed with SYPRO-orange, and melting profiles were obtained between 20 and 70 $^{\circ}\text{C}$. Mutation of the lysine to arginine does not diminish the stability, unfolding pattern or melting temperature of the protein. Error bars represent standard deviation from 3 independent replicates. (B) Decreasing concentrations (30, 15, 7 and 3 nm) of purified recombinant TDP1 or TDP1 K111R was incubated for 1 hour at 37 $^{\circ}\text{C}$ with a ^{32}P -radiolabelled duplex-nicked substrate (50 nM) harbouring a 3'-phosphotyrosine 'PY' at the nick (inset). Positions of the ^{32}P -radiolabelled substrate 'PY' and cleavage product 'P' are indicated by arrows and show that cleavage of the tyrosine is similar when TDP1 or TDP1 K111R are added to the substrate. (C) Reaction products 'P' were quantified relative to total labeled substrate 'P + PY' and percentage conversion to 3'-P from experiments conducted in (B) was quantified. This shows that mutation of TDP1 K111R does not impact on the cleavage of the tyrosine-phosphate.

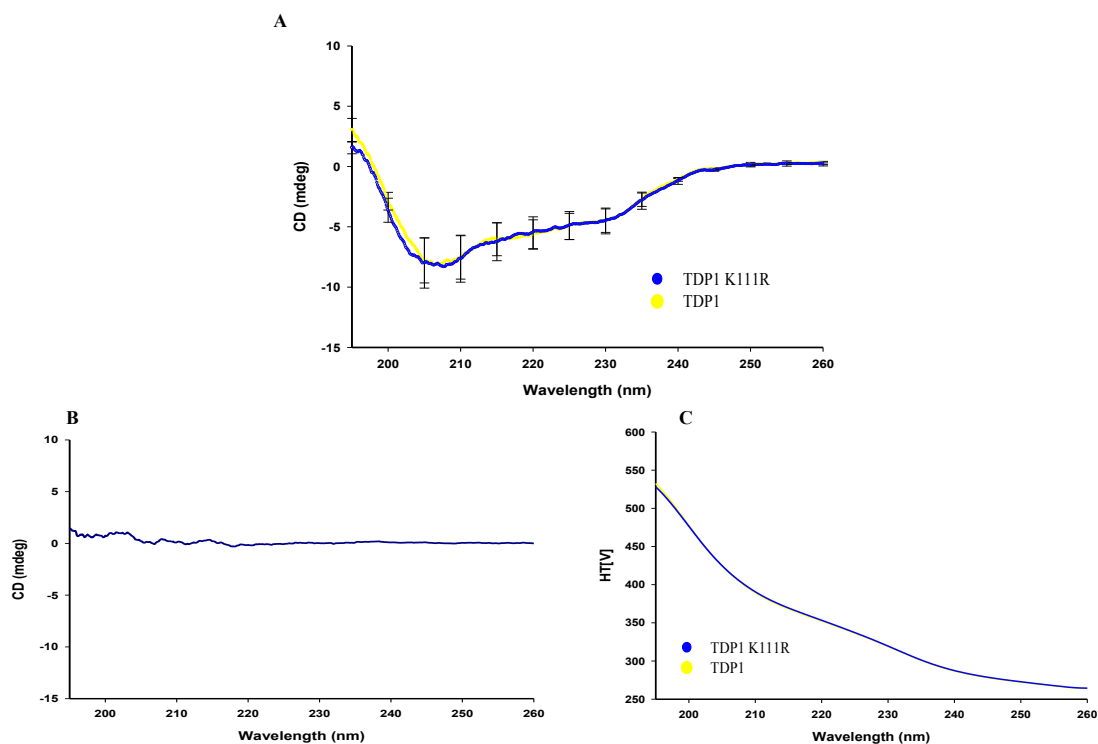


Figure 4.2.9 Comparison of hTDP1 and hTDP1¹⁵¹⁻⁶⁰⁸ using CD shows no gross structural difference between the proteins

Comparison of recombinant TDP1 and TDP1^{K111R} through circular dichroism spectroscopy resulted in two similar spectral graphs with negative bands at ~210 and 220 nm, and positive bands at ~195 nm (A). This suggests that a lysine to arginine mutation has no impact on secondary structure. Data represent the average of four independent experiments \pm s.e.m. Data was only recorded if the high-tension voltage (HT [V]) < 550 V (C). Subtraction of the hTDP1¹⁵¹⁻⁶⁰⁸ spectra from that of hTDP1 (B) results in a relatively flat line not indicative of any secondary structure.

	Helix		Sheet		Others*		NRMSD# (<0.25)	
CDSSTR Reference set 4	TDP1	TDP1 K111R	TDP1	TDP1 K111R	TDP1	TDP1 K111R	TDP1	TDP1 K111R
Secondary Structure (%)	16.3 \pm 1.2	16.3 \pm 0.3	29.0 \pm 1.2	28.0 \pm 0.6	55.3 \pm 0.3	55.7 \pm 0.9	0.04	0.05
P value (n=3)	1.00000		0.57792		0.74180			

* Loops, turns, and unstructured

Normalised root mean square deviation

Table 4.2.1 Bioinformatics analyses of core secondary structure using CDSSTR algorithm

Secondary structure analysis shows no significant difference in secondary structure between wild type TDP1 and TDP1^{K111R}. Analysis was conducted using the variable selection method (CDSSTR) with reference data set 4 from the Dichroweb server, which gave the best goodness-of-fits. NRMSD values were 0.04 and 0.05 for TDP1 and hTDP1^{K111R} respectively.

4.2.3 SUMOylation-deficient mutant accrues more DNA single-strand breaks that are in part transcription dependent

4.2.3.1 TDP1 K111R mutant shows reduced survival in the presence of CPT

To assess the impact the SUMO-deficient mutant may have *in vivo* another lab member infected Tdp1 ^{-/-} mouse embryonic fibroblasts (MEFs) with retrovirus particles containing vector alone (Tdp1 ^{-/-}; V) or encoding human TDP1 (Tdp1 ^{-/-}; hTDP1) or TDP1 K111R (Tdp1 ^{-/-}; hTDP1 K111R). To examine if hTDP1 K111R had an impact on cellular survival, we compared the three cell types for their ability to form macroscopic colonies following exposure to CPT. Tdp1 ^{-/-}; V cells were hypersensitive to CPT as expected and complementation with hTDP1 led to marked protection. Complementation with hTDP1^{K111R} led to an intermediate level of protection (**Figure 4.2.10 A**). These differences were not due to differences in protein expression, since immunoblots show a similar level of TDP1 and TDP1^{K111R} (**Figure 4.2.10 B**).

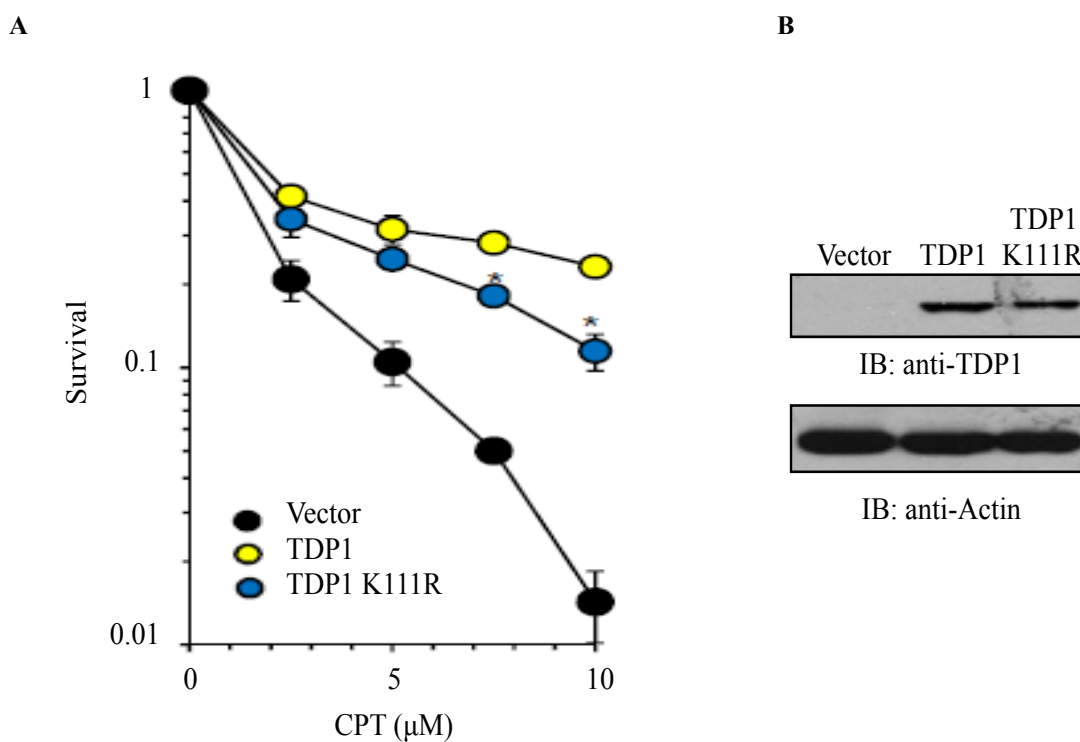


Figure 4.2.10 TDP1 K111R mutant shows reduced survival in the presence of CPT

(A) Tdp1 ^{-/-} MEFs complemented with empty vector or with vectors expressing wild-type TDP1 or TDP1 K111R. Survival was calculated by dividing the average number of colonies on treated plates by the average number of colonies on untreated plates and shows that hTDP1 helps protect Tdp1 ^{-/-} MEFs from CPT. Expression of hTDP1 K111R resulted in a better resistance to CPT than Vector alone but were more sensitive than cells expressing hTDP1. Error bars represent s.e.m for 3 independent experiments. A difference in survival was not due to levels of expression between TDP1 and TDP1 K111R (B).

4.2.3.2 TDP1 K111R accumulates more DNA single-strand breaks in the presence of CPT

Since TDP1 K111R had a reduced survival rate in the presence of CPT when compared to the wild type TDP1, we wanted to see if the K111R mutant resulted in an increase in single-strand breaks in the presence of CPT, when measured by the alkaline comet assay. HEK293 cells expressing endogenous levels of TDP1 were transfected with GFP-SUMO1 and empty vector, or vectors encoding Myc-TDP1 or Myc-TDP1 K111R. As expected, HEK293 cells expressing TDP1 accumulated four- to fivefold fewer SSBs compared with control cells (**Figure 4.2.11**). Cells expressing TDP1 K111R accumulated two- to three- fold more single-strand breaks when compared with TDP1-expressing cells ($P = 0.0169$; student's t-test). This defect was not due to differences in expression levels as confirmed by immunoblotting. With SUMOylation defective TDP1 mutant showing more CPT induced single-strand breaks and lower survival in comparison to the wild-type protein, the next step was to see if this defect was linked to transcriptionally associated breaks.

4.2.3.3 Accumulation of DNA single-strand breaks is in part transcription dependant

HEK293 cells were transfected with GFP-SUMO1 and empty vector, or vectors encoding Myc-TDP1 or Myc-TDP1 K111R. These cells were then pre-treated with DMSO '- DRB' or 50 μ M of the transcriptional inhibitor DRB '+ DRB' for 2 hrs. DMSO 'mock', or 50 μ M of CPT 'CPT' was then added to the cells for 1 hour at 37 °C. DNA single-strand breaks were then measured using the alkaline comet assay. Once again the TDP1 K111R mutant resulted in a 2-3fold higher average comet tail moment compared to the wild-type (**Figure 4.2.12 A**). Interestingly, pre-incubation with transcription inhibitor DRB resulted in no difference in comet tail moment ($P=0.87$; student's t-test) between wild-type and mutant protein. Differences seen, once again could not be due to differences in protein expression levels as determined by immunoblots (**Figure 4.2.12 B**). These data support the data presented in the previous chapter where the N-terminal domain of TDP1 is involved in the repair of transcription associated breaks, suggesting that this effect is, at least in part, mediated via SUMOylation at lysine 111.

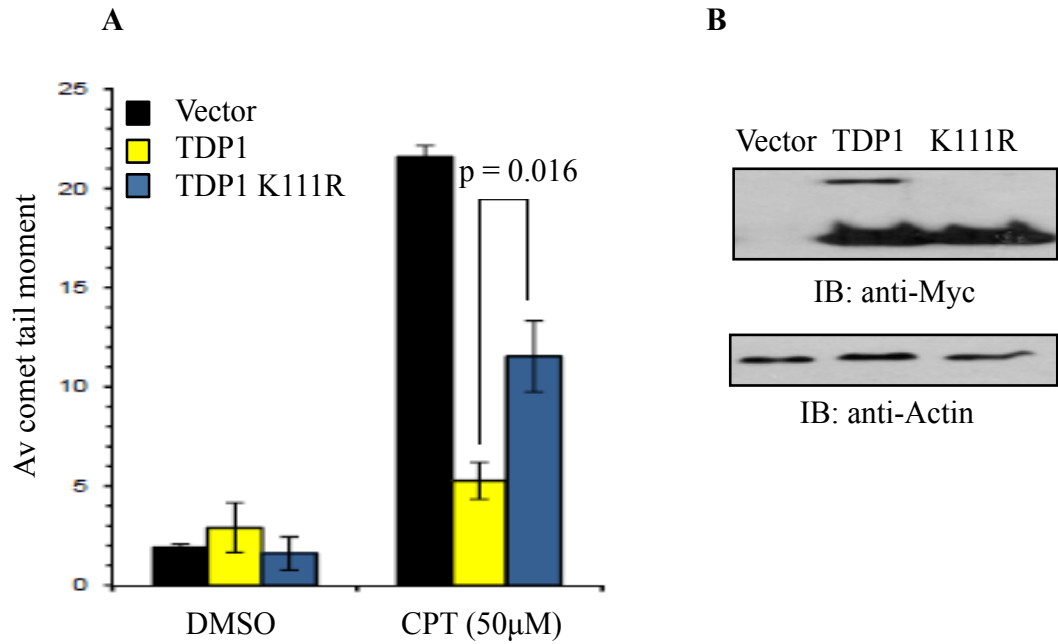


Figure 4.2.11 TDP1 K111R accumulates more DNA single-strand breaks in the presence of CPT
(A) HEK293 cells were transfected with GFP-SUMO1 and empty vector ‘vector’, or vectors encoding Myc-TDP1 ‘TDP1’ or Myc-TDP1 K111R ‘TDP1 K111R’ using conditions that produce > 90% transfection efficiency. Cells were incubated with DMSO ‘mock’ or 50 µM camptothecin ‘CPT’ and DNA breaks quantified by alkaline comet assays. Mean tail moments were quantified for 50 cells per sample per experiment, and data are the average of $n = 3$ biological replicates \pm s.e.m. **(B)** Lysate from cells used were analysed by immunoblotting.

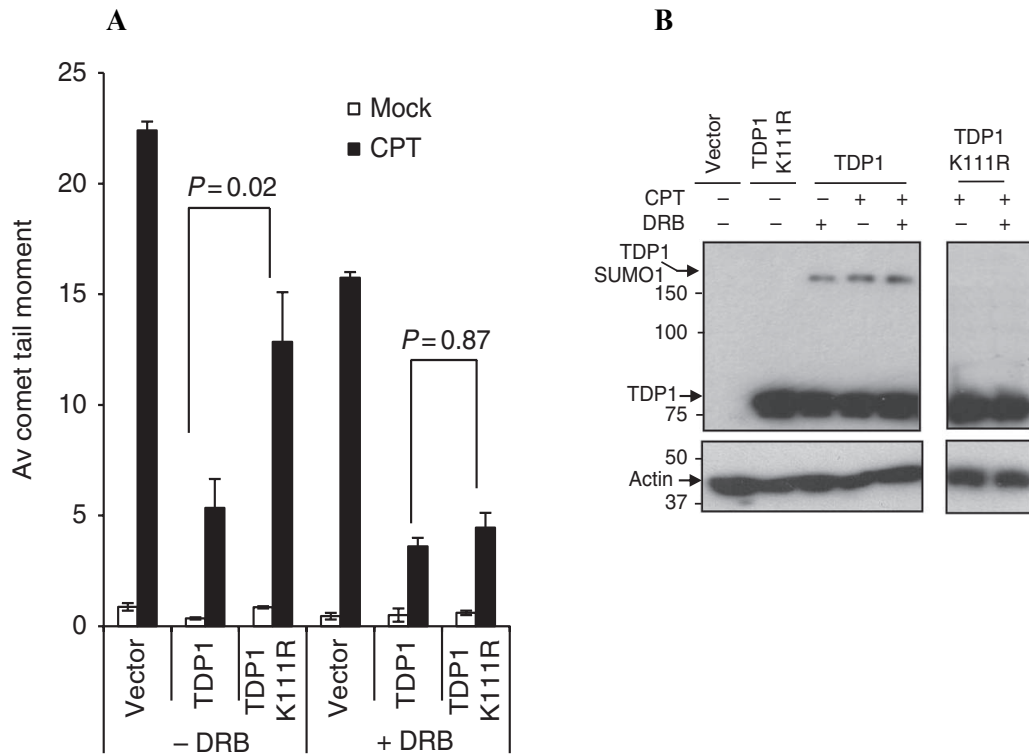


Figure 4.2.12 Accumulation of DNA single-strand breaks is in part transcription dependent

HEK293 cells were transfected with empty vector 'vector', myc-TDP1 'TDP1' or myc-TDP1K111R 'TDP1 K111R' and GFP-sumo1. Cells were incubated with DMSO 'mock' or treated with 50 μ M camptothecin 'CPT' with or without previous incubation with 50 μ M DRB for 2 h. DNA strand breakage was quantified by alkaline comet assays and presented as mean tail moment. Data are the average \pm s.e.m. From $n = 3$ biological replicates, where 50 cells per sample were blindly scored from each experiment. Statistical analyses (Student's t-test) were conducted to compare the difference between TDP1 and TDP1 K111R in the absence or presence of DRB, and the corresponding P-values are depicted. (A) Lysate from cells used for experiments in (B) were fractionated by SDS-PAGE and analysed by immunoblotting.

4.2.4 TDP1 interacts non-covalently with SUMO2/3

4.2.4.1 TDP1 interacts with SUMO2 *in vitro*

Since TDP1 is SUMOylated at K111 by SUMO1 and TDG has been reported to interact covalently and non-covalently with SUMO (Smet-Nocca et al., 2011), we wanted to investigate if TDP1 could also interact with the SUMO family in a non-covalent manner, via a SUMO interacting motif (SIM). Recombinant His-TDP1 (0.23 µg) was mixed with protein G beads (negative control) or SUMO1, 2, 3 conjugated beads (ENZO life sciences) for 1 hour at 4 °C. After incubation samples were washed 5 times before protein analysis by immunoblotting. TDP1 was found to interact non-covalently with SUMO2 and SUMO3, as intense bands were detected by anti-TDP1 (Eurogentec) in the SUMO2 and SUMO3 fractions 'S2', 'S3' when compared to the SUMO1 fraction 'S1' or protein G beads 'PG' (**Figure 4.2.13**).

4.2.4.2 TDP1 interacts with SUMO2 via the catalytic domain *in vitro*

Next we wanted to map the site of the SIM/s. To do this recombinant His-TDP1, His-TDP1¹⁻¹⁵⁰ or His-TDP1¹⁵¹⁻⁶⁰⁸ were mixed with SUMO1, or SUMO2 conjugated beads following the method described in section 2.6.2.1. The input in (**Figure 4.2.14 A**) shows that less TDP1¹⁵¹⁻⁶⁰⁸ was added to the beads, but the immunoprecipitate displays a band corresponding to TDP1¹⁵¹⁻⁶⁰⁸ with a similar intensity as the full-length protein. On the other hand TDP1¹⁻¹⁵⁰ gave a much weaker signal. None of the truncations, nor full-length protein showed an interaction with SUMO1 beads. These results could be interpreted in one of three ways: i) The N-terminal of TDP1 contains a SIM/s but has a much weaker interaction; ii) truncation of TDP1 moiety leads to the presentation of a SIM site/s that would otherwise be buried in the full-length protein, or finally iii) levels of protein used in this assay were much higher than would be found endogenously and therefore this weaker band seen when SUMO2 is mixed with TDP1¹⁻¹⁵⁰ is an artefact of the experiment. To dissect these three possibilities, TDP1¹⁵¹⁻⁶⁰⁸, TDP1¹⁻¹⁵⁰, or both TDP1¹⁵¹⁻⁶⁰⁸ and TDP1¹⁻¹⁵⁰, were mixed with conjugated SUMO2 beads (**Figure 4.2.14 B**). The immunoblot demonstrates that TDP1¹⁵¹⁻⁶⁰⁸ outcompetes the TDP1¹⁻¹⁵⁰ domain when the same levels of protein were in the same fraction. This implies that the SIM/s are in the phosphodiesterase domain and that the weak interaction shown in the N-terminal domain is probably an artifact of the assay.

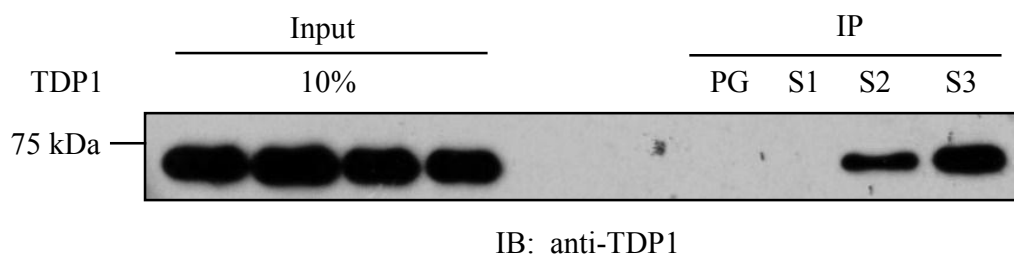


Figure 4.2.13 TDP1 binds to SUMO2 non-covalently

0.23 μ g of recombinant His-TDP1 was mixed with 20 μ l protein G 'PG', SUMO1 'S1', SUMO2 'S2', or SUMO3 conjugated beads (Activmotif). TDP1 was found at a much larger quantity in 'S2' and 'S3' fractions compared to 'S1' or PG negative control.

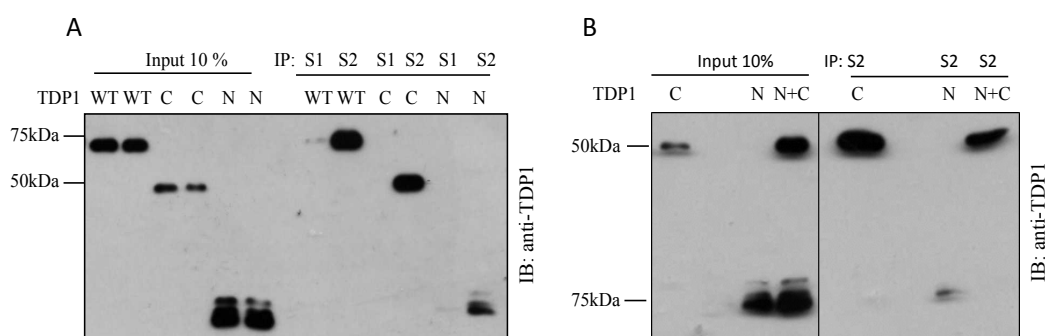


Figure 4.2.14 SUMO2 binds non-covalently to the catalytic domain of TDP1

0.23 μ g of recombinant His-TDP1 'WT', TDP1¹⁵¹⁻⁶⁰⁸ 'C' or TDP1¹⁻¹⁵⁰ 'N' were mixed with 20 μ l SUMO1 'S1' or SUMO2 'S2' conjugated beads (Activmotif). Immunoblots with anti-TDP1 (Eurogentec) detected TDP1 and TDP1¹⁵¹⁻⁶⁰⁸ in the SUMO2 fraction whilst TDP1¹⁻¹⁵⁰ was found at a much lower level (A). (B) Addition of TDP1¹⁵¹⁻⁶⁰⁸ and TDP1¹⁻¹⁵⁰ to the same tube containing SUMO2 'N+C' shows the catalytic domain out competes the N-terminal domain for binding to SUMO2.

4.2.4.3 Identification of SIM/s in TDP1

To identify the site/s of the SUMO2 - TDP1 interaction conserved regions across homologues were analysed to see if they contain the generic SIM recognition motif. This follows the sequence V/I-X-V/I-V/I where X is any amino acid (Hannich et al., 2005; Minty et al., 2000; J. Song et al., 2004). Sequence analysis identifies 6 potential sites: ILLV, LRVV, IDVI, VYLI and LGVL (**Figure 4.2.15**). SIM sites are hydrophobic regions that normally conform to a beta sheet structure (Kerscher, 2007). This would suggest that the two strongest candidates for the interacting motif were ILLV and LYLI. Intriguingly, TDP2, a newly discovered enzyme that possess 5'-TDP activity (Cortes Ledesma et al., 2009; Zeng et al., 2011), was found to have a SIM sequence of IDVI (Hecker et al., 2006) which is one of the candidates for TDP1. To determine if these sites were the location of the SIM the first two amino acids in these potential sites were mutated to harbor a di-alanine. These SIM mutations were then expressed and purified following the protocol for His-tagged proteins.

4.2.4.4 Mutation of the first two amino acids in potential SIM sites results in instability and co-expression with GroEL

Once purified SIM mutants: SIM1 (ILLV-AALV), SIM2 (LRVV-AAVV), SIM3 (IDVI-AAVI), SIM4 (VYLI-AALI) and SIM5 (LYLI-AALI) were mixed with the SUMO2 conjugated beads and analysed by immunoblots to test if these mutants could still bind to the SUMO2 protein (**Figure 4.2.16**). The input shows that SIM mutants all seem to cause a truncation resulting in the same size fragment at ~ 62 kDa. SIM 2, 3 and 5 gave faint bands consisting with the size of full-length TDP1 ~ 72 kDa. SIM4 shows a second truncated product ~ 58 kDa. The immunoprecipitate shows that the truncation found in all SIM mutants was present, whilst the higher molecular weight band in SIM5 'SM5' was no longer detectable. Both higher molecular weight signals were present in the SIM2 and 3 lanes and the truncation ~ 58 kDa found in the SIM4 mutant input is also present (**Figure 4.2.16 A**). The conclusion from this experiment could be that the higher molecular weight band present in the SIM5 input but not detectable in the immunoprecipitate ablates SUMO2 binding. The lower molecular weight bands, seen as "truncations", could be a result of an inherent instability of the di-alanine mutations, but this seems unlikely because they should not all result in the exact same molecular weight. It is more probable that the di-alanine mutation causes folding issues with all mutants during expression of the proteins. As a result chaperone proteins

are accrued to help stabilize the misfolded protein. One common chaperone protein, which is known to co-purify with misfolded proteins is that of GroEL (Katayama et al., 2009). GroEL is approximately 62 kDa in size and therefore the truncated band seen in all purifications on an SDS-PAGE gel (**Figure 4.2.16 B**) could in fact be GroEL.

To test this SIM4 and SIM5, which both show this lower molecular weight protein on the SDS-PAGE gel (**Figure 4.2.16 B**) were analysed by western and probed with anti-GroEL antibody (Pearl lab). Intense bands were seen at the weight corresponding to the “truncation” whilst no other bands were detected. This demonstrates that the truncated protein is in fact the heat shock protein GroEL. The ability of GroEL to co-purify with TDP1 SIM mutants and can bind to the Protein G beads means that these results could not be taken into account and an alternative method for SIM identification should be explored.

4.2.4.5 TDP1 fragments within the catalytic domain causes loss of expression *in vivo*

Circumvention of this problem could be obtained by creating truncations within the phosphodiesterase domain that would remove potential SIM sites. The crystal structure was studied and it was decided to create truncations within unfolded regions of the protein (**Figure 4.2.17**). This may help prevent folding issues and disturbances within the secondary structure. Initially four fragments were created: TDP1²⁸¹⁻⁶⁰⁸, TDP1³⁶⁸⁻⁶⁰⁸, TDP1¹⁵⁰⁻⁶⁰⁸ and TDP1¹⁵⁰⁻⁴¹⁶. These fragments were created in the PCI-Myc mammalian expression vector. Once sequenced TDP1 fragments were transfected into HEK293 cells by calcium phosphate transfection, lysed and analysed by immunoblots. Bands detected in the TDP1¹⁵¹⁻⁶⁰⁸ lanes show that transfections were successful (**Figure 4.2.18**). Unfortunately, none of the TDP1 fragments were detected, at levels detectable by anti-Myc. This means that SIM site determination could not be found using this method either.

SIM recognition motifs: V/I-X-V/I-V/I

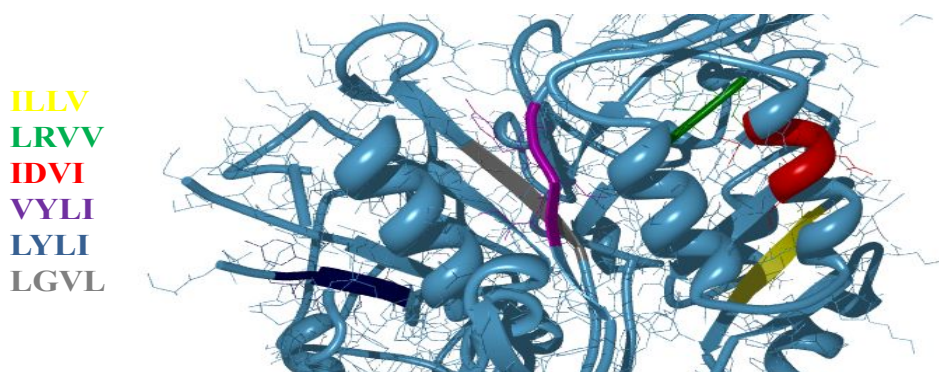


Figure 4.2.15 Sequence analysis predicts 6 potential SIM sites in TDP1

SIM binding motifs are recognized with the amino acid sequence as V/I-X-V/I-VI with X being any amino acid. BLAST sequence analysis of the TDP1 and conserved regions between species identifies 6 potential SIM binding sites. ILLV, LRVV, IDVI, VYLI, LYLI and LGVL.

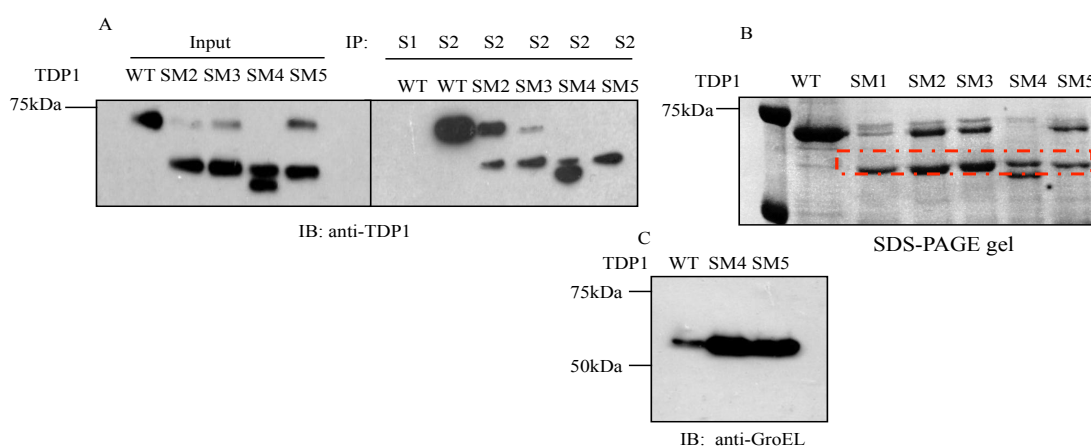
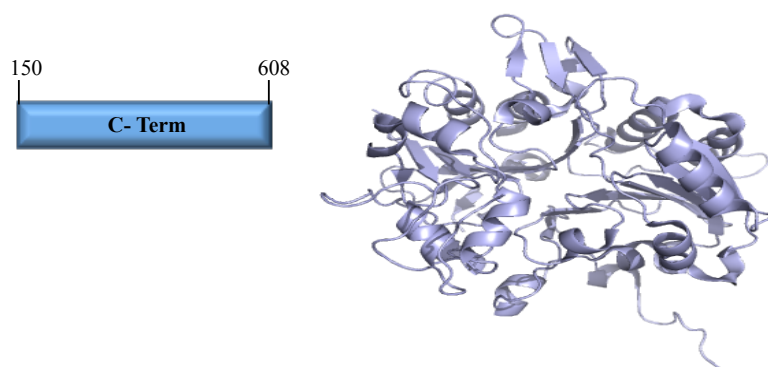


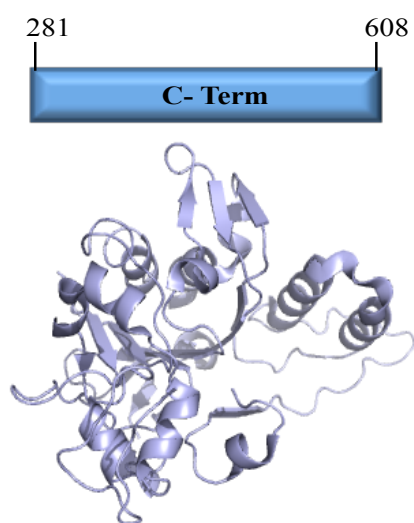
Figure 4.2.16 Mutation of SIM sites results in truncation and expression of GroEL

The first two amino acids of potential SIM sites were mutated to alanine. 0.23 μ g recombinant His-TDP1 'WT' or SIM mutants 'SM2, SM3, SM4, SM5' were mixed with SUMO2 conjugated beads. Immunoblot analysis (anti-TDP1; Eurogentec) was carried out to determine which TDP1 mutants would co-precipitate with SUMO2 (**A**). Due to a band being detected at the same molecular weight throughout all TDP1 mutants that does correspond to the size of TDP1 an SDS-PAGE gel was run to see if this band was present after purification or if the proteins had degraded (**B**). The SDS-PAGE gel shows that there are bands in all the lanes at 62 kDa (highlighted in red) that is not present in the wild type protein. This means that the SIM mutants all result in a truncation of a similar size or a contaminant is present. The band size seen is correct size for the chaperone protein GroEL, which can co-purify with misfolded protein. To test this recombinant protein of SIM mutant 4 and 5 was run on a fresh SDS-PAGE gel transferred to a nitrocellulose membrane and probed for heat shock protein GroEL, using anti-GroEL antibody (**C**). The prominent band found in (**C**) is of GroEL and means that the SIM mutants co-purify with this protein.

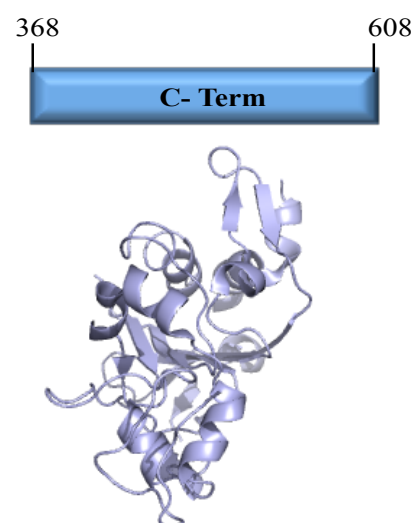
A



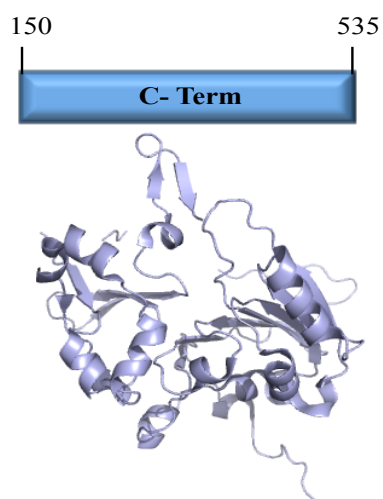
B



C



D



E

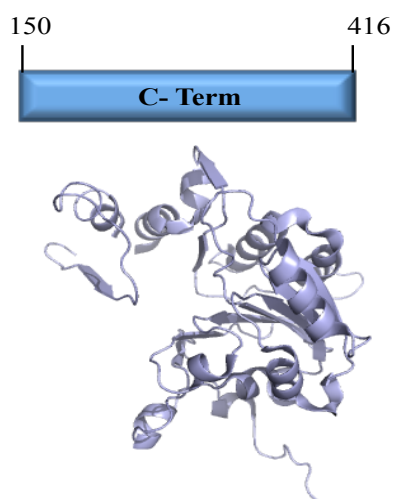


Figure 4.2.17 Determination of truncation points to map SIM motifs

The crystal structure of TDP1¹⁵¹⁻⁶⁰⁸ was analysed to determine locations within the protein where truncations could be created within unstructured regions to remove potential SIMs.

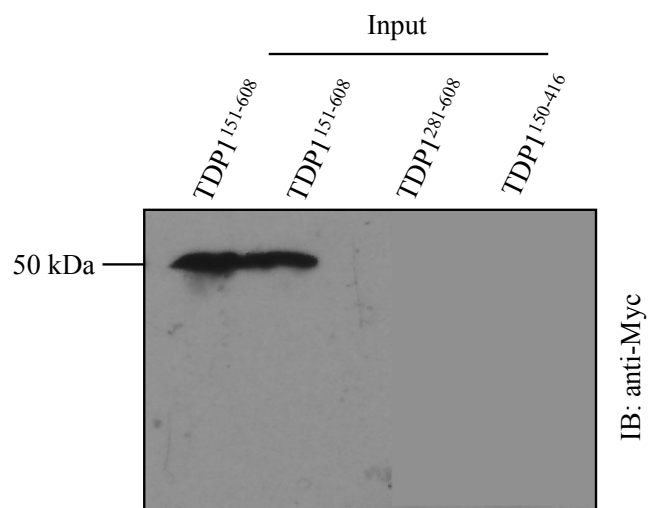


Figure 4.2.18 Truncations within the catalytic domain results in loss of protein expression

HEK293T cells were transfected with Myc-TDP1¹⁵¹⁻⁶⁰⁸, Myc-TDP1²⁸¹⁻⁶⁰⁸, or Myc-TDP1¹⁵⁰⁻⁴¹⁶ using calcium phosphate transfection. Lysates were prepared from transfected cells by re-suspension in lysis buffer (20 mM HEPES pH 7.4, 0.5 % NP40, 40 mM NaCl, 2 mM MgCl₂, 1 × protease inhibitor cocktail (Roche), 1 × phosphatase inhibitor (Roche), 20 mM NEM and 25 U/ml benzonase (Merck)) 10 % of total lysate was fractionated via SDS-PAGE and sample analysed by immunoblot using anti-Myc (9E11; cell signaling)

The last protocol to identify the location of SIM/s in TDP1 was to expose recombinant TDP1 to a tryptic digestion. Trypsin is known to cleave at the carboxy terminus of arginine or lysine, but not before proline (Olsen et al., 2004). If after a specific amount of time TDP1 is digested into multiple fragments then the SIM/s could be narrowed down to a particular fragment. **Figure 4.2.19** shows that incubation of TDP1 with trypsin at a ratio of 1:50 results in rapid fragmentation, even after one hour. This would not give us the desired fragments to determine the peptide sequence that interacts with SUMO2. Incubation at a ratio of 1:250 displays a varied digestion of TDP1 after one hour, as seen by multiple bands at different molecular weight sizes. This could be used in conjunction with SUMO2 conjugated to beads to “fish” out the compatible fragments. These stretches of fragmented protein could then be scrutinized using a mass spec to detect the peptide fragments of TDP1 that bind to SUMO2/3. Further single or double amino acid substitutions could then be used to obtain a stably expressing TDP1 that is unable to interact with SUMO2 or 3. This TDP1 SIM mutant could then be tested *in vivo* to determine what role the SIM interaction may have.

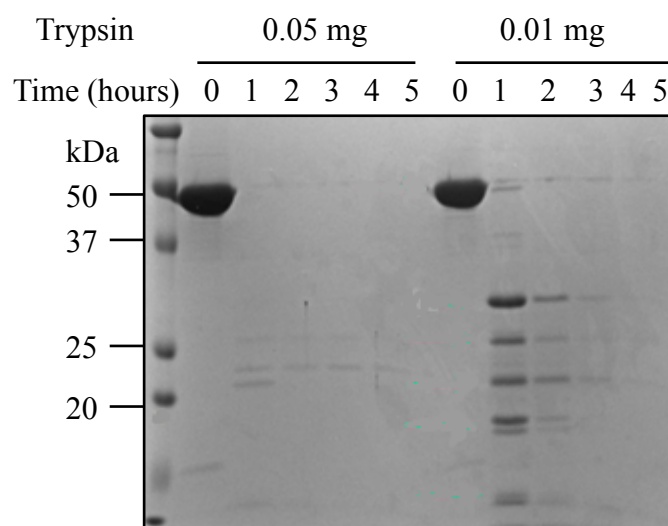


Figure 4.2.19 Fragmentation of TDP1¹⁵¹⁻⁶⁰⁸ by incubation with trypsin

2.5 mg of recombinant His-TDP1¹⁵¹⁻⁶⁰⁸ was incubated with 0.05 mg, or 0.01 mg of trypsin and samples were taken at 1, 2, 3, 4 and 5 hours. Protein from these time points were then fractionated on an SDS-PAGE gel and stained with Coomassie blue.

4.3 Discussion

This chapter identifies that TDP1 is post-translationally modified by SUMO1. This modification has been shown *in vitro* and *in vivo*. Furthermore, SUMOylation occurs via TDP1's evolutionarily driven N-terminal domain at lysine 111. *In silico* analyses of *S. cerevisiae* Tdp1 (Ren et al., 2009) also predicts no SUMOylation site when the amino acid sequence is scanned under high threshold conditions. Therefore, it appears that SUMOylation of TDP1 by SUMO1 is only found in higher eukaryotes. The lysine at position 111 in hTDP1 is conserved in several vertebrate species but does not seem to be induced by cellular damage (data not shown), as upon addition of CPT there is no increase in SUMOylation of TDP1. Instead this appears to be a house keeping modification where a basal amount of TDP1 is constantly SUMOylated. TDP1 SUMOylation is not a static process and is maintained by activities of sumo conjugation (SAE1/2, UBC9) and de-conjugation (SENPs). The observation that TDP1 has a basal amount of SUMOylated TDP1 is supported by the fact that the majority of SUMO1, unlike SUMO2 and 3, within the cell is not freely available. Therefore, proteins modified by SUMO1 generally have a small fraction constantly conjugated to the SUMO1 molecule (Minty et al., 2000; Saitoh and Hincley, 2000).

We also demonstrate that mutation of lysine 111 to an arginine abrogates the covalent modification and sensitizes cells to camptothecin when compared to wild-type TDP1. Ablation of SUMOylation also results in an increase in single-strand DNA breaks when measured by the alkaline comet assay. Because lysine residues also serve as attachment sites for other modifications, we also show a similar defect in wild-type TDP1 cells when UBC9 is knocked down, and that this has no effect on cells expressing TDP1^{K111R} (J. J. R. Hudson et al., 2012). This further consolidates the idea the defect observed is due to an inability for TDP1 to be covalently modified by SUMO1 at lysine 111.

We show that SUMOylation of TDP1 does not change the catalytic activity of the protein, nor does the amino acid substitution from lysine to arginine at 111 that is used to abolish SUMOylation, and identify the role that SUMO modification of TDP1 has in cells.

We identify that SUMO modification of TDP1 is required for optimal protection against the DNA damaging agent camptothecin, and that this is, in part, transcription dependent. Furthermore, SUMOylation of TDP1 at K111 led to an increase in cellular concentration at sites of damage. When cells expressing GFP-TDP1 or GFP-TDP1 K111R were subjected to ultraviolet A laser-induced DNA damage, GFP-TDP1 accumulated to sites of damage more rapidly than TDP1 K111R and also resulted in a higher concentration of TDP1 at these sites. Additionally, inclusion of transcription inhibitor DRB before ultraviolet A laser-induced DNA damage led to reduction in TDP1 accumulation at sites of laser damage, this could mean that recruitment of SUMOylated TDP1 is also transcriptionally dependent (J. J. R. Hudson et al., 2012). The defect seen when TDP1 cannot be SUMOylated is also not due to a change in core structure of the K111R mutant as alpha helical and beta sheet content remain the same when measured by circular dichroism. The unfolding pattern and melting temperatures of TDP1 and TDP1^{K111R} also remain similar and thought not to be the cause of the defect seen.

The fact that TDP1 is SUMOylated in the N-terminal domain and that this is required for optimal cellular protection, in which, part of this is transcription dependent re-enforces the conclusions of my first chapter and adds to the model for which the N-terminus of TDP1 is required. It shows that TDP1¹⁻¹⁵⁰ is needed to increase cellular concentrations of the enzyme to sites of transcriptionally associated Top1 damage, and that this process is facilitated by SUMOylation at lysine 111 by SUMO1. Furthermore, it establishes the importance the N-terminal has in post-mitotic cells as complementation of Tdp1 ^{-/-} quiescent cortical neural cells with human TDP1 or TDP1^{K111R} revealed a role for K111 to maintain cell viability following CPT (J. J. R. Hudson et al., 2012). Taken together, we conclude that mutating the acceptor lysine (K111) to a non-SUMOylatable arginine results in defects in repairing single-strand breaks and highlights the role for this domain, particularly in post-mitotic cells. This discovery builds on the model of the first chapter (**Figure 4.3.1**) and identifies a role for SUMOylation of TDP1 in the repair of abortive Top1 complexes during transcription (J. J. R. Hudson et al., 2012).

Hudson, J.J.R. *et al.* SUMO modification of the neuroprotective protein TDP1 facilitates chromosomal single-strand break repair. *Nat. Commun.* 3:733 doi: 10.1038/ncomms1739 (2012).

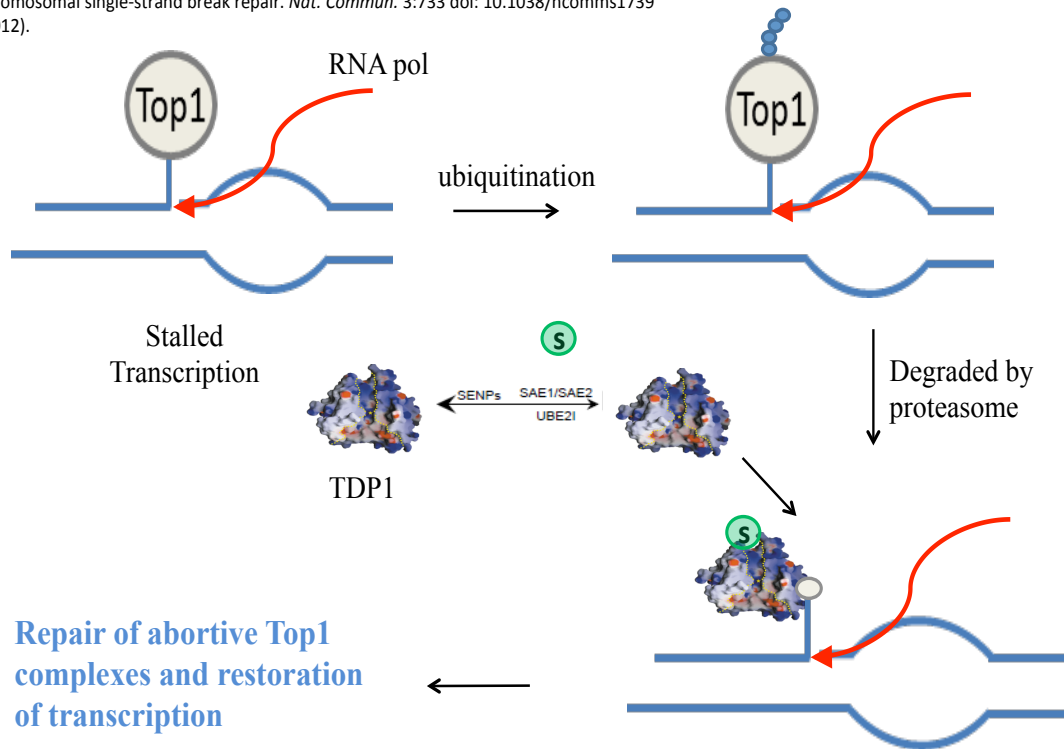


Figure 4.3.1 Model for the repair of abortive Top1 complexes at sites of transcription

Top1 can become trapped on the DNA and may collide with the transcription machinery. Once this occurs the Top1 enzyme undergoes ubiquitylation and is degraded by the 26S proteasome. SUMOylated TDP1 is then signalled to the site of damage and removes the 3' tyrosine. This allows other members of the DNA repair machinery to repair the single-stranded break and allow transcription to progress. Collision of elongating RNA polymerases, such as RNA Pol II with Top1 intermediates, leads to stalling of the polymerase with subsequent proteasomal degradation of Top1, and possibly the stalled RNA Pol. We propose that these collision events recruit TDP1 to sites of polymerase stalling. TDP1 exists in equilibrium between unmodified (the majority) and a SUMOylated version, and the balance is maintained by the opposing activities of SUMO conjugation (SAE1/2, UBC9 and possibly a sumo ligase) and deconjugation (SENPs). We show that TDP1 SUMOylation occurs primarily at K111 and propose that SUMOylated TDP1 is at least, in part, engaged in dealing with the transcription-blocking lesions.

I have also demonstrated that TDP1 interacts with SUMO2 or 3 via TDP1's catalytic domain. Unfortunately, due to problems with expression of TDP1 SIM truncations *in vivo*, or stability of TDP1 SIM mutants *in vitro*, I was unable to map the SIM/s that could allow us to determine what role this may have within cells. Interestingly, Chains of SUMO have been identified to bind covalently to Top1. Depending on how much of the Top1 peptide is left after degradation by the 26S proteasome the SUMO chains may act as a binding site for TDP1. This could increase the recruitment to the site of abortive Top1 complexes, or allow a specific orientation of TDP1 binding to increase the efficacy in which the enzyme hydrolyses the Top1 peptide from DNA. Ren *et al* have recently updated their database for post-translation modification (Ren et al., 2009) to now include SIM binding motifs. *In silico* analysis of the TDP1 protein identifies the sequence ILLVH at amino acid positions 224-228 as the only candidate to interact non-covalently with SUMO under high or medium threshold conditions. Further work must be carried out to obtain a TDP1 SIM mutant that can be expressed *in vivo* to determine what role this SUMO-TDP1 non-covalent interaction has in cells.

Chapter V

Results 3: The role of DNA ligases and CK2 during TDP1 mediated repair

5.1 Introduction

Tyrosyl DNA phosphodiesterase-1 (TDP1) protects cells from abortive Top1 activity by hydrolysing the 3'-phosphotyrosyl bond that links Top1 to a DNA single-strand break and is currently the only known human enzyme that displays this activity in cells. Following restoration of normal DNA termini, a DNA ligase is required to seal broken DNA strands and restore the integrity of the genetic material.

5.1.2 Ligase family and TDP1

TDP1 and Lig III α had been previously reported to interact via the N-terminus of TDP1 (Chiang et al., 2010; Das et al., 2009; El-Khamisy et al., 2005) and therefore was thought to be a prime candidate for TDP1 nuclear and mitochondrial mediated repair. Unlike the other members of the ligase family Lig III α contains a mitochondrial leader sequence and is the only Ligase shown to be involved in mitochondrial DNA repair (Bentley et al., 2002; Ellenberger and Tomkinson, 2008; Lakshminpathy and Campbell, 1999). A truncated version, splice variant, is nuclear bound and interaction with the XRCC1 scaffold protein is required for its stability (Caldecott et al., 1994; Nash et al., 1997). This nuclear Lig III α has been linked with several single-strand break repair pathways and is currently thought to be the only ligase involved in TDP1 nuclear mediated repair.

Surprisingly, depletion of Lig III α from human cells had no effect on the rapid repair of chromosomal Top1-linked DNA breaks (Jude, unpublished data). In striking contrast, DNA ligase I (Lig I), but not DNA ligase IV, deficient human cells exhibited marked defect in repairing chromosomal Top1-linked DNA breaks (Jude, unpublished data). These results prompted me to test if Lig I interacts with TDP1. Since homologues of Lig I are ubiquitous in Eukarya, whereas Lig III α appears to be restricted to vertebrates (Ellenberger and Tomkinson, 2008), I examined whether the association of Lig I and TDP1 occurs via the evolutionarily-conserved domain of TDP1. I show that whilst Lig III α interacts with the N-terminal vertebrate-conserved region of TDP1, Lig I interacts

with the C-terminal evolutionarily conserved domain. Our preliminary findings suggest a role for Lig I during TDP1-mediated nuclear DNA repair and suggest a distinct role for Lig III α likely through the repair of mitochondrial Top1-breaks.

Interestingly, Ligase III appears to be only found in vertebrae and interacts via S81 of the N-terminal domain of TDP1, another evolutionarily developed domain conserved only in vertebrae. This then gives rise to the question of how the TDP1 enzyme with only the C-terminal domain being ubiquitously conserved throughout organisms could lead to repair from this type of damage in lower organisms without the aid of Lig III α ?

If Lig III α is in fact the main ligase involved in the sealing of the processed DNA termini then treatment of Lig III α depleted cells with camptothecin, a topoisomerase 1 poison, should give a similar result to TDP1 knockout cells, or TDP1 defective SCAN 1 cells found in SCAN 1 patients (El-Khamisy et al., 2005; Hirano et al., 2007; Interthal et al., 2005b; Takashima et al., 2002).

5.1.3 Casein kinase II

Whilst investigating the interaction of TDP1 and Lig I, I also identified that TDP1 was phosphorylated *in vivo* and demonstrated that phosphorylation of TDP1 occurs within the N-terminal domain of TDP1. Since TDP1 is SUMOylated at K111, which is found in the N-terminal domain of TDP1 we examined the regulation of SUMOylation through a phospho-dependent SUMOylation motif.

TDP1 was already shown to be phosphorylated in response to damage by ATM and DNA-PK (Chiang et al., 2010; Das et al., 2009). These studies also rule out the possibility that TDP1 is phosphorylated by ATR. Thus, research into kinases that target protein interacting partners of TDP1 and that also possess the ability to constitutively phosphorylate proteins led us to further investigate casein kinase II (CK2) as a possible candidate.

CK2 is a hetero-tetramer serine/threonine and to some extent tyrosine kinase that is composed of two alpha and two beta sub units (Hanks and Hunter, 1995; Meggio et al., 1994; Vilks et al., 2008) Crystal structure analysis reveals that the catalytic kinase

activity is found within the two alpha subunits and has an extensive list of target substrates, with important roles in cell growth and survival (Meggio and Pinna, 2003; Pinna and Meggio, 1997).

5.1.4 Chapter objectives

In this chapter, I conducted a series of experiments to test if Lig I deficient cells are hypersensitive to the Top1 poison CPT and if Lig I interacts with TDP1, to support its role during nuclear repair of Top1-breaks. Since homologues of Lig I are ubiquitous in Eukarya, I then tested the hypothesis that the association of Lig I and TDP1 occurs via the evolutionarily conserved domain of TDP1. Investigations were also done to test if the interaction between Lig I and TDP1 was mediated by phosphorylation. The results of which lead to further experimental work into the phosphorylation of TDP1, the potential impact phosphorylation may have on TDP1 and if TDP1 is limited to phosphorylation by ATM and DNA-PK.

5.2 Results

5.2.1 Ligase I and not ligase III α is implicated in TDP1-mediated nuclear repair

5.2.1.1 Ligase I deficient cells results in hypersensitivity to camptothecin

Our lab previously reported that TDP1 interacts with DNA Lig III α (Chiang et al., 2010) (**Figure 5.2.1**), suggesting that it is the main ligase involved in the repair of nuclear Top1-mediated DNA damage. To test this hypothesis, Sherif El-Khamisy depleted Lig III α from human cells and quantified the extent of DNA breaks induced by the Top1 poison camptothecin using alkaline comet assays. Surprisingly, whilst TDP1 deficient cells accumulated 5-fold more DNA breaks compared to control cells, breaks did not increase above background level in Lig III α -deficient cells (Jude, unpublished observations). This led to two possible conclusions; either other members of the ligase family were contributing to the repair in the absence of Lig III α , or Lig III α is not the main contributor to nuclear repair of abortive Top I complexes. To test the latter conclusion we examined whether DNA ligase IV, or Lig I could fulfil this function *in vivo*. Our lab compared Lig IV deficient human cells for their ability to repair Top1-breaks. Similarly, CPT induced similar levels of breaks in both control and Lig IV deficient cells (data not shown), suggesting that Lig IV does not participate in this repair process. In striking contrast, initial experiments show that DNA Lig I deficient human cells (46BR), containing a hypomorphic mutation, accumulated more CPT-induced DNA breaks compared to control cells (**Figure 5.2.2 A**). Taken together, these results suggest one of two things. That Lig I and not Lig III α facilitates the repair of nuclear Top1-linked DNA breaks, or that there is a redundancy between ligase I and III α in nuclear repair of Top-linked DNA breaks.

The later scenario seems most likely, as TDP1 has been shown to interact with the Lig III α -XRCC1 complex, a well-known component of single-strand break repair pathway (Caldecott, 2007; Reynolds et al., 2009). It is possible that the hypomorphic mutation in Lig I cells (Barnes et al., 1992; Ellenberger and Tomkinson, 2008; Prigent et al., 1994; Webster et al., 1992) leads to interaction with the DNA break, but has an inability to seal the phosphodiester backbone, thus acting as a competitive inhibitor to Lig III α .

Knockdown of Lig III α , Lig I and both ligases in cells should be carried out to answer which, or if both these ligases can play a role in TDP1-mediated repair. Single-strand DNA breaks observed could also be due to okazaki fragments, and should be ruled out by repeat experiments with replication inhibitor aphidicolin.

5.2.1.2 The conserved catalytic domain of TDP1 is sufficient to maintain interaction with Lig I

The results reported above prompted me to test whether Lig I does interact with TDP1, in a manner similar to that of Lig III α . For these experiments we ectopically expressed Flag-tagged Lig I and Myc-tagged TDP1 in HEK 293 cells. Immunoblots of the anti-Myc immunoprecipitate (IP) revealed Flag-Lig I in IPs conducted from cells co-expressing both Lig I and TDP1, but not in control cells expressing Flag-Lig I and Myc (**Figure 5.2.2 B**). We next examined whether Myc-TDP1 would be pulled down in a reverse IP. Flag-Lig I was pulled down using anti-Flag antibody (**Figure 5.2.2 C**) the immunoblots of the anti-Flag immunoprecipitate show that Myc-TDP1 is present in the Flag-Lig I IP but not in control cells. These results suggest that Lig I interacts with TDP1 in human cells.

Homologues of Lig I are ubiquitous in Eukarya, whereas Lig III α appears to be restricted to vertebrates. Consistent with an evolutionarily driven requirement for Lig III α , we previously reported that it interacts with a domain of TDP1 that was only found in vertebrates (Chiang et al., 2010). To examine whether Lig I interacts with the more conserved domain of TDP1, we examined if the association of Lig I and TDP1 occurs via the evolutionarily conserved C-terminal domain. For these experiments we ectopically expressed Myc-TDP1 and Myc-TDP1¹⁵¹⁻⁶⁰⁸ and examined their association with Lig I by Co-IPs (**Figure 5.2.3**). The results establish that the interaction between Lig I and TDP1 occurs within the conserved C-terminal portion of TDP1, suggesting a similar mechanism of repair in lower eukarya to that found in vertebrates.

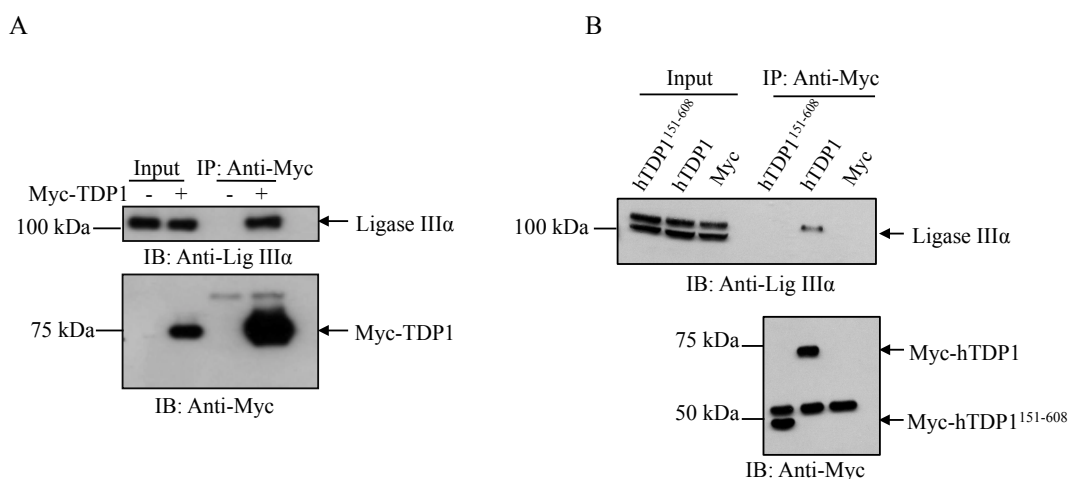


Figure 5.2.1 Interaction with hTDP1 and Ligase III occurs within the N-terminal domain of TDP1 in DT40 cells

(A) Myc-TDP1 or Myc empty vector were ectopically expressed in HEK293 cells. Proteins were immunopurified using anti-Myc antibody (9B11; cell signaling) and immunoblots were probed with anti-Myc (9B11; 1:2000) or anti-Lig IIIα (TL25; 1:2000). Lig IIIα is found to co-purify with Myc-TDP1 but not Myc alone, suggesting an interaction between TDP1 and Lig IIIα. (B) DT40 cells expressing Myc-hTDP1, Myc or Myc-TDP1¹⁵¹⁻⁶⁰⁸ were lysed and proteins immunopurified. Immunoblots were probed with anti-Myc or anti-Lig IIIα antibody and blots show an interaction with Lig IIIα and hTDP1 but not with hTDP1¹⁵¹⁻⁶⁰⁸ indicating that IIIα interacts with the N-terminus of hTDP1.

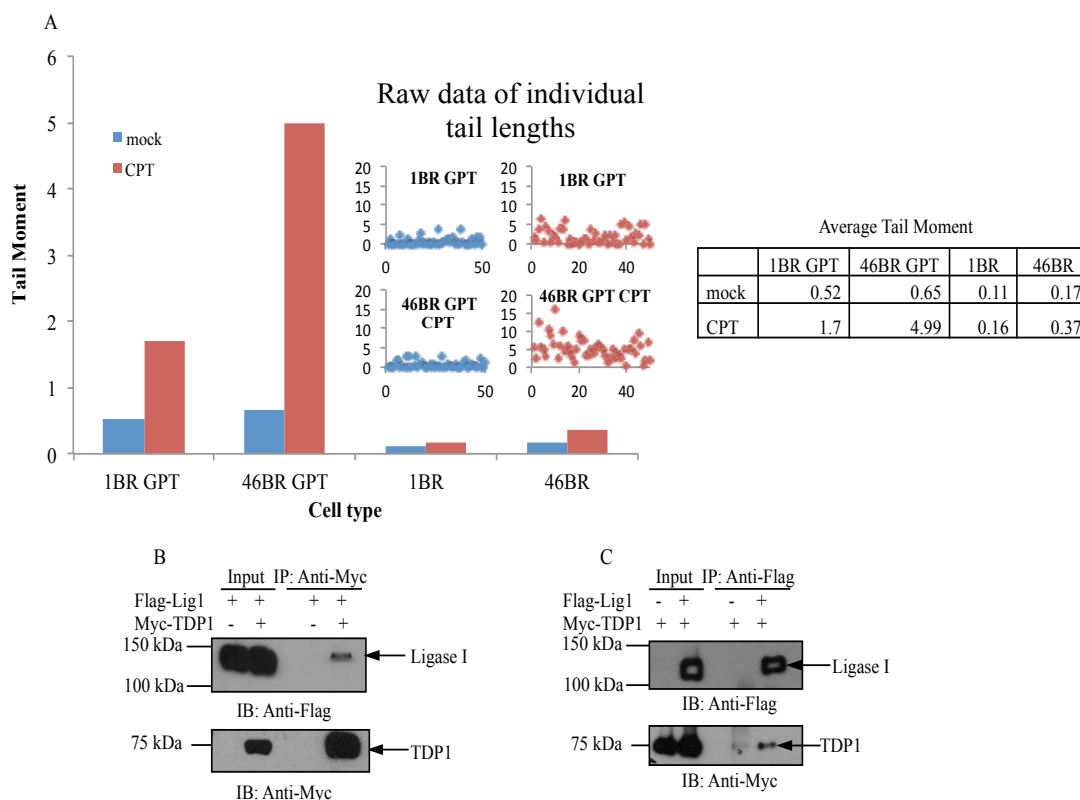


Figure 5.2.2 Ligase I interacts with hTDP1 and ligase I deficient cells are hypersensitive to camptothecin

(A) Wild type immortalized fibroblasts (1BR GPT), Wild type (1BR), Lig I deficient fibroblasts (46BR) and Lig I deficient immortalized fibroblasts (46BR GPT) cells were mock treated with DMSO or 15 μ M camptothecin for 40 minutes and breaks quantified by the alkaline comet assay. The average tail length of 50 cells were calculated and raw data of individual tail lengths is shown. Ligase I deficient cells have a larger tail moment when treated with CPT compared to the wild type cells. **(B)** Myc-TDP1 or Myc empty vector were ectopically expressed with Flag-Ligase I in HEK293 cells. TDP1 was immunopurified using anti-Myc antibody (9B11; cell signaling). Immunoblots were probed with anti-Myc and anti-Flag antibodies and showed that TDP1 co-purifies with Ligase I. **(C)** Flag-Ligase I or Flag empty vector were ectopically expressed with Myc-TDP1 in HEK293 cells. Flag-Ligase I was immunopurified using anti-Flag antibody (F3165; Sigma). Immunoblots were probed with anti-Myc (9B11; 1:2000) and anti-Flag (F3165; 1:2000) antibodies and shows that Ligase I co-purifies with TDP1.

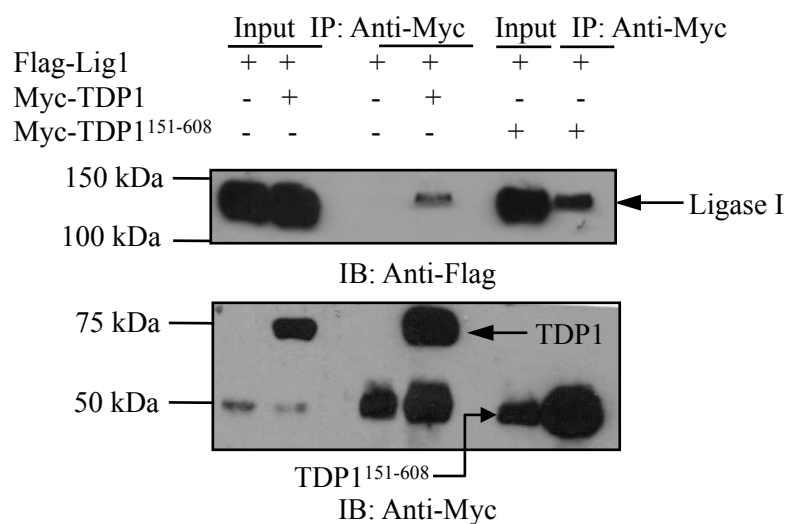


Figure 5.2.3 Ligase I interacts with hTDP1¹⁵¹⁻⁶⁰⁸

Myc-TDP1, Myc empty vector or Myc-TDP1¹⁵¹⁻⁶⁰⁸ was ectopically expressed with Flag-Ligase I in HEK293 cells. TDP1 was immunopurified using anti-Myc antibody (9B11; cell signaling). Immunoblots were probed with anti-Myc (9B11; 1:2000) and anti-Flag (F3165; 1:2000) antibodies and demonstrates that TDP1 co-purifies with Ligase I via its' catalytic domain.

5.2.2 TDP1 is a novel substrate for Casein Kinase II

5.2.2.1 TDP1 is phosphorylated *in vivo*

Because Lig I is hyper-phosphorylated by CK2 in a cell cycle dependent manner (Prigent et al., 1992) we wanted to test if the phosphorylation could play a role in the interaction with TDP1. To test this we ectopically expressed Myc-TDP1 and Flag-Lig I in HEK 293 cells. After proteins were purified using Myc antibody, lambda phosphatase was added to one fraction and samples were left for 1 hour at 30 ° C. Samples were then washed 3 times in 25 mM HEPES and 150 mM NaCl, pH 7.4 followed by fractionation by SDS-PAGE gel, and analysis. Immunoblots showed loss of the upper band corresponding to phosphorylated Lig I (**Figure 5.2.4**) but it did not seem to ablate interaction with TDP1. Unexpectedly, and perhaps more interesting was the faster migrating band of TDP1 in the lambda phosphatase treated fraction. Whilst TDP1 was shown to be phosphorylated by ATM and DNA-PK at serine 81 (Das et al., 2009), a single phosphorylation site at S81 is unlikely to correspond to such a significant band shift. It is plausible that this is not the only phosphorylation event that occurs, and this shift in molecular weight could be due to a cluster of phosphorylation events.

5.2.2.2 Phosphorylation of TDP1 occurs in the N-terminal domain

Since TDP1 is phosphorylated at serine 81 we decided to investigate if the reduction in TDP1 band size when treated with lambda phosphatase was limited to either the catalytic domain or the N-terminal region of TDP1. Myc-TDP1, Myc-TDP1¹⁻¹⁵⁰ and Myc-TDP1¹⁵¹⁻⁶⁰⁸ were transfected and expressed in HEK 293 cells. Proteins were immuno-purified using Myc antibody (9B11; Cell signalling) and subsequently treated with or without lambda phosphatase for 1 hr at 30 °C. Treatment with lambda phosphatase resulted in faster migrating bands for Myc-TDP1 and Myc-TDP1¹⁻¹⁵⁰ when compared to control lanes (**Figure 5.2.5 A and C**) but no shift was seen in the catalytic domain (**Figure 5.2.5 B**). This led us to believe that there could be several phosphorylation events occurring within the N-terminal region of TDP1.

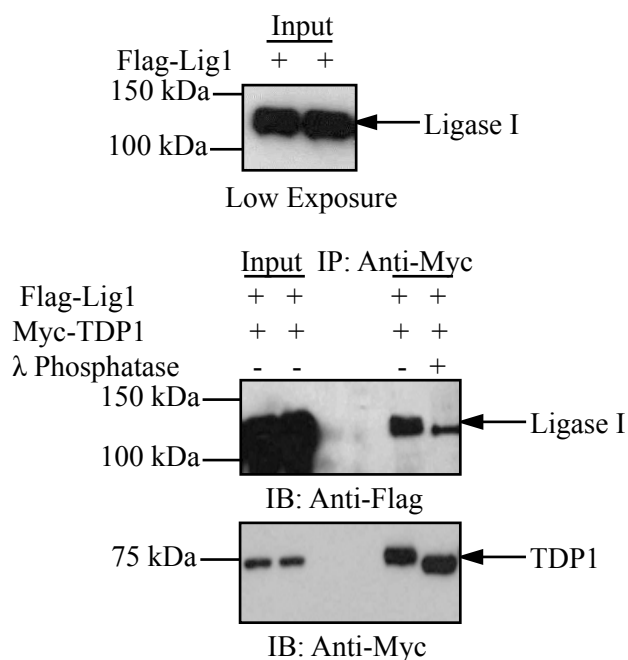


Figure 5.2.4 TDP1 is constitutively phosphorylated as treatment of TDP1 with Lambda phosphatase results in a faster migrating band, indicative of dephosphorylation

Myc-TDP1 and Flag-Ligase I were ectopically expressed in HEK293 cells. TDP1 was immunopurified using anti-Myc antibody (9B11; cell signaling). Immunoblots were probed with anti-Myc and anti-Flag antibodies. Once purified buffer and Lambda phosphatase was added to one reaction '+', whilst buffer alone was added to the control '-'. Addition of lambda phosphatase resulted in loss of the upper band of Ligase I, corresponding to faster migration due to dephosphorylation by the phosphatase. Unexpectedly, TDP1 is also seen to migrate at a faster rate upon treatment of lambda phosphatase.

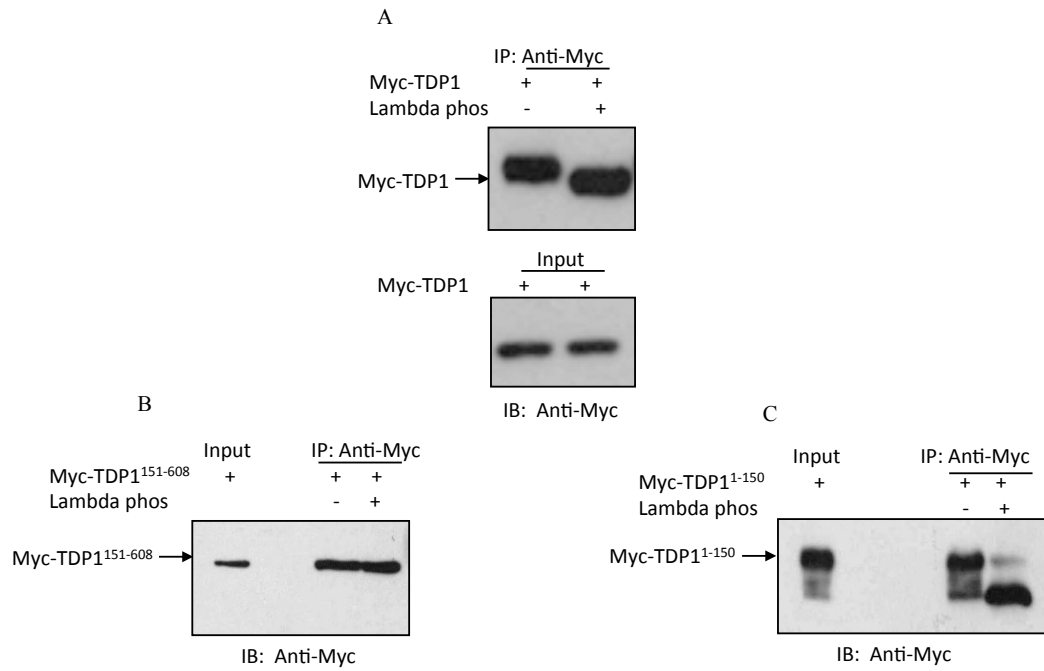


Figure 5.2.5 *In vivo* phosphorylation of TDP1 occurs within the N-terminal domain

Myc-TDP1 (A), Myc-TDP1¹⁵¹⁻⁶⁰⁸ (B), or Myc-TDP1¹⁻¹⁵⁰ (C) were ectopically expressed in HEK293 cells. TDP1 was immunopurified using anti-Myc antibody (9B11; cell signaling) and immunoblots were probed with anti-Myc antibody (9B11; 1:2000). Once purified, buffer and Lambda phosphatase was added to one reaction '+', whilst buffer alone was added to the control '-'. Addition of lambda phosphatase results in lower molecular weight proteins in the full-length TDP1 and TDP1¹⁻¹⁵⁰ fractions, but no change is observed in the catalytic domain of TDP1.

5.2.3 Phosphorylation of TDP1 by CK2 inhibits SUMOylation, *in vitro*

What might be the role of TDP1 phosphorylation? Phospho-dependent SUMOylation motifs are an intriguing way to add a level of control by acting as a molecular switch for SUMOylation. I first examined the role of S81 phosphorylation. Since the phosphorylation observed was found in the N-terminus of TDP1 and S81 is phosphorylated in response to damage we wanted to test if this could impact upon SUMOylation of TDP1, especially since we had previously identified TDP1 SUMOylation at lysine 111 and that this site was in close proximity to phosphorylation at S81. In order to do this recombinant His-TDP1, His-TDP1 S81A (a phosphomutant) and His-TDP1 S81E (phosphomimetic) were purified using a nickel charged column. Protein concentrations were determined by the nanodrop and subjected to the *in vitro* SUMOylation assay (Activmotif) previously described in the second results chapter.

Immunoblots for 3 independent experiments were analysed using the image quant software and percentage SUMOylation was calculated for TDP1 wild-type and mutants against loading controls (**Figure 5.2.6**). These repeats showed a reduced level of SUMOylation in the phosphomutant when compared to the wild-type protein. However, an increase in levels of SUMOylation was not observed in the phosphomimetic protein, although mutation of serine to glutamic acid in order to mimic serine phosphorylation does not always work. This probably means that an interaction with UBC9 (SUMO E2) and TDP1 occurs in the vicinity and mutation of the serine residue interferes with the binding. Since a slight decrease in SUMOylation was observed in the serine 81 to alanine mutation it is possible that this region is important for binding of UBC9 and therefore has a slight impact upon SUMOylation.

Whilst SUMOylation of TDP1 is slightly affected by mutation of S81, this may not be the only phosphorylation site within the N-terminus. In 2004, the Caldecott group, identified a role for CK2 in maintaining the genetic integrity of mammalian cells by facilitating the rapid repair of chromosomal DNA single-strand breaks through phosphorylation of the molecular scaffold protein XRCC1 (Loizou et al., 2004; Parsons et al., 2010; Ström et al., 2011). Their results match several target criteria seen in the phosphorylation of TDP1. They found that XRCC1, like TDP1, was constitutively

phosphorylated *in vivo*, that upon treatment with calf intestinal phosphatase a large shift in gel mobility was seen, and that mass spectrometry revealed tri- and tetra-phosphorylated species. The fact that XRCC1, a scaffold protein, is also in a complex with TDP1 during single-strand break repair makes CK2 an interesting candidate.

I next examined the possibility that TDP1 phosphorylation by CK2 modulates SUMOylation. To test if TDP1 is a substrate for CK2 recombinant His-TDP1 was incubated at 30 °C for 2 hrs in the presence of 5 mM ATP with or without CK2 (NEB). Samples were fractionated on an SDS-PAGE gel and probed with anti-TDP1 (Abcam; 1:2000). **Figure 5.2.7 A** shows when TDP1 is in the presence of CK2 a slower migrating band corresponding to an increase in molecular weight is observed. This shift was not seen in the control fraction, which led us to conclude that TDP1 is a substrate for CK2, at least, *in vitro*. To determine possible sites for phosphorylation the amino acid sequence of TDP1 was subject to *in silico* analysis. The table in **Figure 5.2.7 B** shows a number of potential sites for phosphorylation, interestingly all of which are found in the N-terminal domain of TDP1.

The next step was to see if these phosphorylation events could affect the *in vitro* SUMOylation of TDP1. Therefore, TDP1 was phosphorylated by CK2 before addition to the SUMOylation reaction previously described in chapter III. A control sample without the addition of CK2 was also subject to analysis. Results show that phosphorylation of TDP1 by CK2 severely inhibits SUMOylation *in vitro* (**Figure 5.2.7 C and D**). While this result appears to be intriguing, suggesting that CK2 could act as a molecular switch for TDP1 modification by SUMO1, we cannot rule out that the kinase activity of CK2 may act on the components of the SUMO pathway, inhibiting their activity. To test this, we first established conditions that would inhibit CK2 catalytic activity. Several concentrations of TBB (a known inhibitor of CK2) (Pagano et al., 2008; Sarno et al., 2001) were pre-incubated with CK2 before addition of the kinase to TDP1. SDS-PAGE fractionation and Coomassie blue staining of these reactions showed that the lower concentration of TBB (400 µM) was enough to ablate the phosphorylation of TDP1 (**Figure 5.2.8 A**).

Once the concentration had been identified, we next subjected recombinant TDP1 to CK2 that had, or had not, been pre-treated with the appropriate dose of TBB. Following

incubation of TDP1 with CK2 for 2 hours TBB was then added to both reactions for an additional hour, to prevent any impact CK2 might have on the efficiency of SUMOylation independently of TDP1. Reactions were then fractionated and analysed by immunoblotting. Whilst both reactions had CK2 and recombinant His-TDP1, pre-incubation with TBB was enough to prevent phosphorylation of TDP1, as shown by the faster migrating band on the left hand side at ~ 75 kDa (**Figure 5.2.8 B**). The un-phosphorylated TDP1 also shows a higher molecular weight band at ~ 100 kDa that corresponds to SUMOylated TDP1. This band was not present in the lane that contains phosphorylated TDP1. Because TBB was added after the addition of CK2 but before the SUMOylation reaction we were able to conclude that it is the phosphorylation of TDP1 by CK2 that negatively impacts SUMOylation and not phosphorylation of proteins in the SUMO pathway.

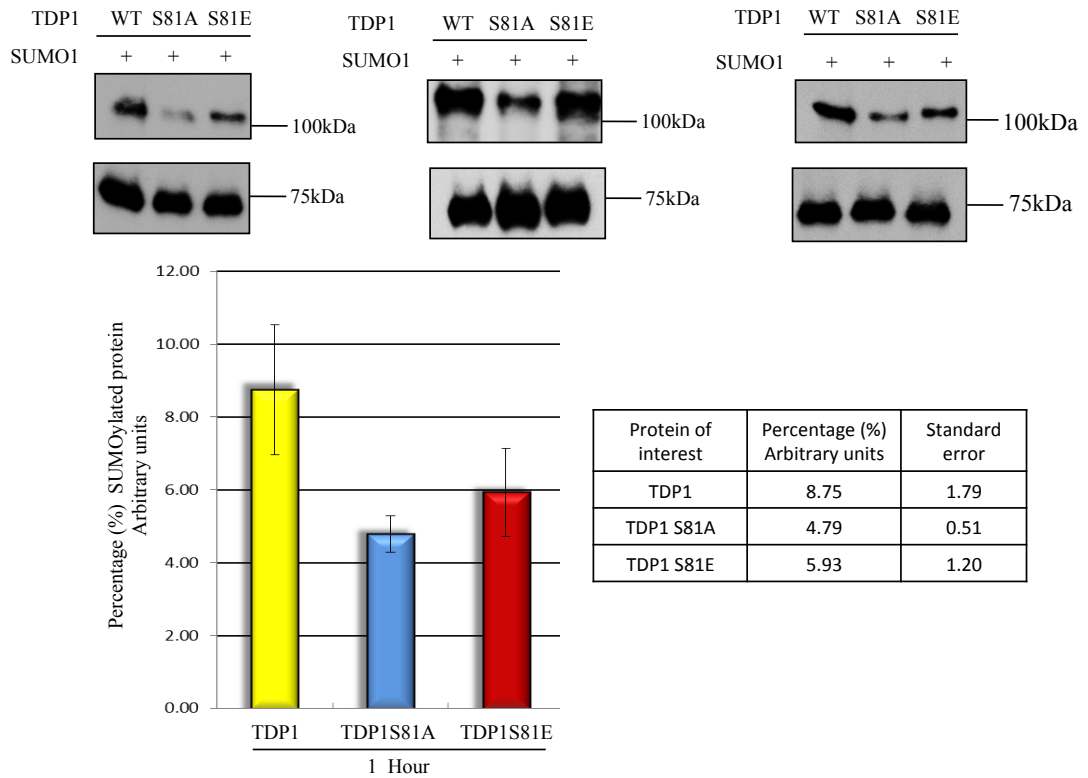


Figure 5.2.6 Mutation of TDP1 S81A reduces SUMOylation *in vitro*

Purified His-TDP1, His-TDP1 S81A or His-TDP1 S81E (500 nM) were incubated with purified SAE1/SAE2 (50 nM), UBC9 (500 nM) and SUMO1 (30 μ M) in 50 mM Tris-HCl, pH 8.0; 50 mM NaCl, 5 mM MgCl₂, 10% glycerol, 0.5 mM DTT and 5 mM ATP. Immunoblots for 3 independent experiments (top left middle and right panels) were analysed by image quant and the percentage SUMOylation was calculated for TDP1 wild-type and mutants (bar chart). SUMO blots were exposed 17 times longer than TDP1 blots and this was taken into account when calculating percentage of SUMOylated TDP1. Error bars represent standard error of the mean of 3 independent experiments.

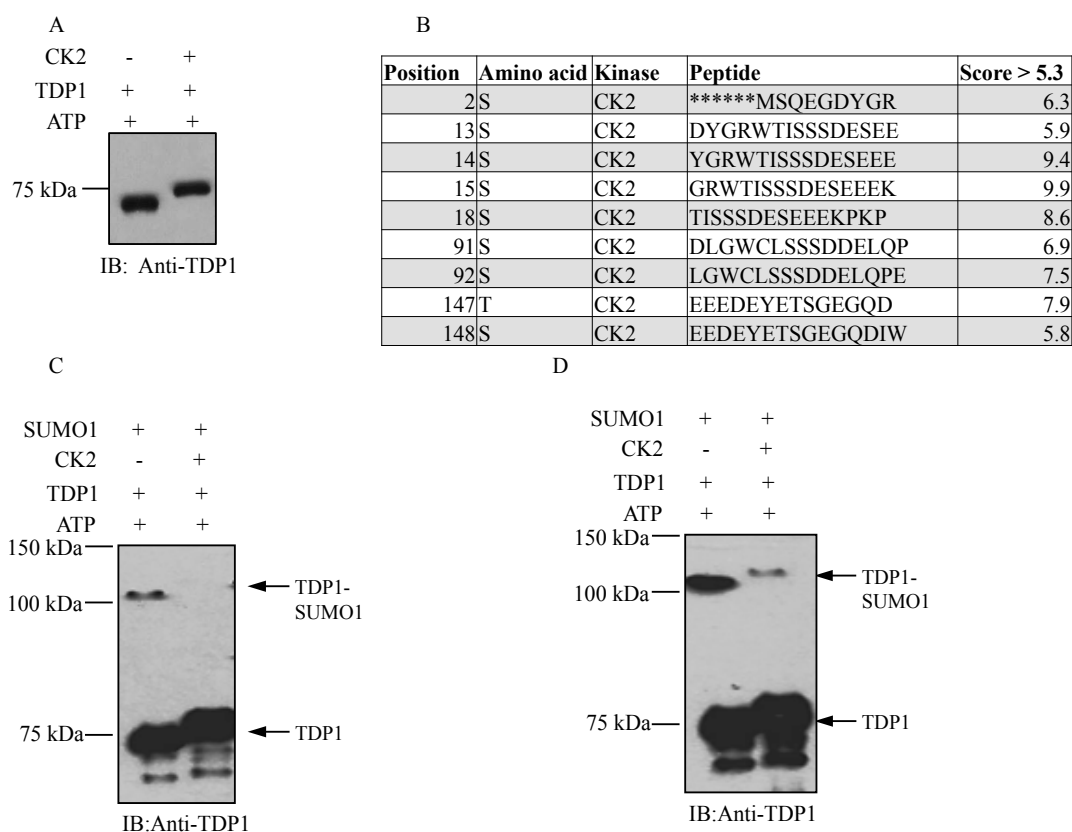


Figure 5.2.7 TDP1 is phosphorylated by CK2 *in vitro*, reducing levels of SUMOylation

(A) Purified His-TDP1, (500 nM) was incubated with ATP (400 μ M), with '+' or without '-' 2 μ l CK2 (NEB). The immunoblot shows that in the presence of CK2 TDP1 runs at a higher molecular than TDP1 - CK2. (B) *In silico* analysis reveals several potential sites for phosphorylation by CK2. (C and D) 2 μ l of the 10 μ l CK2-TDP1 reaction was subjected to SUMOylation followed by fractionation on SDS-PAGE and analyses by immunoblotting using anti-TDP1 antibodies (1:2,000; Abcam). Addition of CK2 to TDP1 decreases the levels of SUMOylation observed.

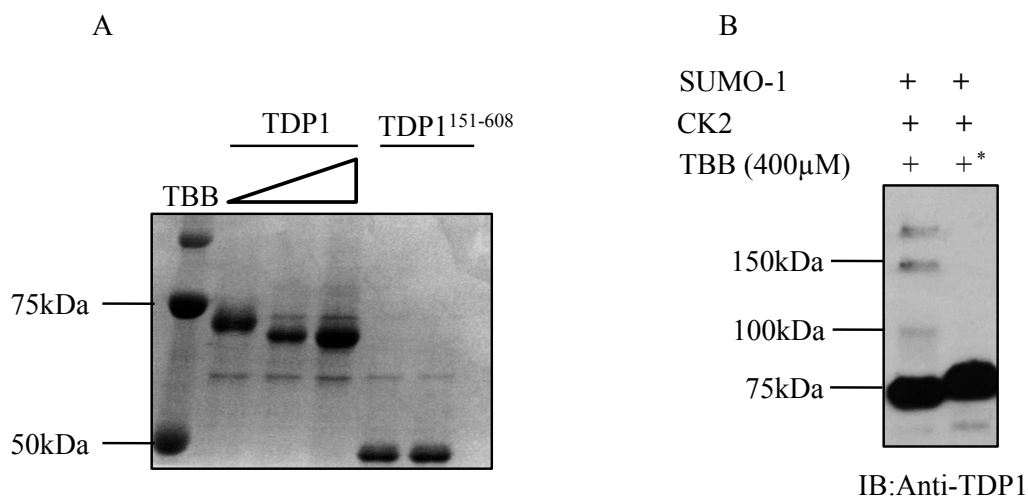


Figure 5.2.8 Phosphorylation of TDP1 and not components of the SUMO pathway ablate SUMOylation of TDP1

(A) CK2 was pre-incubated with 0, 400 or 800 μM TBB for 1 hour at 30 °C before addition of recombinant His-TDP1. TDP1¹⁵¹⁻⁶⁰⁸ (last two lanes) was also incubated with CK2. Samples were fractionated by SDS-PAGE gel and stained with Coomassie blue. This shows that 400 μM TBB is able to prevent CK2 from phosphorylating TDP1 and that TDP1¹⁵¹⁻⁶⁰⁸ is not a substrate for CK2. (B) CK2 was pre-incubated with 400 μM TBB before addition of His-TDP1 (left lane) whilst added after the TDP1-CK2 reaction (right lane asterisk). Both reactions were subject to SUMOylation (Activmotif), followed by fractionation and analysis by immunoblots using anti-TDP1 antibodies (1:2,000; Abcam). Results show that TDP1 phosphorylation and not phosphorylation of components of the SUMO pathway affect TDP1 SUMOylation.

5.2.4 TDP1 is phosphorylated by CK2 in whole cell extracts

Since phosphorylation of TDP1 by CK2 regulates SUMOylation of TDP1 *in vitro*, we next tested if this would be recapitulated with whole cell extracts. One obstacle to overcome was that if CK2 phosphorylation of TDP1 was in fact constitutive it would be difficult to obtain an un-phosphorylated version of TDP1 without knocking down CK2. To circumvent this, recombinant TDP1 (500 nM) was incubated with or without whole cell extract (W.C.E) from HEK 293 cells for 30 mins at 30 °C before fractionation and immunoblot analysis with anti-TDP1 antibody. Analysis initially showed no change in molecular weight of TDP1 in the presence or absence of W.C.E (**Figure 5.2.9 A**). Before ruling out phosphorylation of TDP1 we decided to repeat the experiment with the addition of ATP (5 mM) and Magnesium Chloride (2 mM). **Figure 5.2.9 B** shows that with the addition of ATP and Magnesium Chloride when TDP1 is in the presence of whole cell extract a shift in molecular weight occurs, as seen by the slower migrating band when compared to TDP1 without W.C.E.

Since CK2 is found in both the nucleus and cytoplasm (Faust and Montenarh, 2000; Litchfield, 2003) we next decided to incubate recombinant TDP1 or TDP1 S81A with cytoplasmic or nuclear extract from HEK293 cells. In both cases TDP1 (+) and TDP1 S81A (mt) resulted in slower migrating bands in the presence of nuclear or cytoplasmic extract as compared to incubation with buffer alone (**Figure 5.2.9 C**). This suggests that TDP1 can be phosphorylated by the nuclear or cytoplasmic extract, further indicating that TDP1 could be phosphorylated by CK2 and that mutation of serine to alanine at amino acid 81 does not have a major impact on phosphorylation of recombinant TDP1.

We now had an established protocol for phosphorylation of recombinant TDP1 by cell extract and could use this with selective commercial inhibitors for specific kinases to identify if CK2 is the main kinase for the constitutive phosphorylation of TDP1. Whole cell extract was treated with DMSO control, ATMi, ATRi, DNA-PKi or all three inhibitors for 1 hour at 30 °C. Treated cell extract, or buffer control, was then added to recombinant TDP1 in the presence of ATP (5 mM) and Magnesium Chloride (2 mM). Reactions were left for 30 mins at 30 °C before SDS-PAGE gel fractionation and immunoblotting with anti-TDP1 antibody (Abcam; 1:2000). When in the presence of

W.C.E, TDP1 was seen migrating at a higher molecular weight as previously shown. Interestingly, pre-treatment of W.C.E with the kinase inhibitors did not prevent phosphorylation of TDP1, not even when all three inhibitors were combined (**Figure 5.2.10 A**). This indicates that the higher molecular weight of TDP1 when in the presence of cell extract was not due to ATM, ATR or DNA-PK. The next step was to see if incubation of cell extract with TBB could prevent the slower migration of TDP1 in the presence of cellular extract.

Whole cell extracts were pre-incubated with or without 400 μ M TBB for 1 hour at 30 °C before incubation with recombinant TDP1 for an additional 30 mins. Samples were fractionated and analysed by immunoblotting with anti-TDP1 antibody (Abcam; 1:2000). Incubation of TDP1 with cell extract again resulted in a retardation of the band corresponding to TDP1 when compared to the control lane. Intriguingly, when cellular extract was pre-treated with TBB, a known CK2 inhibitor, TDP1 migration was similar to that of the control sample (**Figure 5.2.10 B**). This observation shows for the first time that TDP1 is phosphorylated by CK2 *in vitro* and by whole cell extracts, although TBB inhibition of other kinases cannot be ruled out.

To discover the role of TDP1 phosphorylation by CK2 we used mass spectrometry to try to pinpoint the sites of phosphorylation. 14 mgs of DT40 whole cell extract containing Myc-hTDP1 was split into two fractions. One fraction was be treated with lambda phosphatase, as a negative control, and one fraction to be washed with buffer alone. TDP1 was purified using anti-Myc antibody (9B11; cell signaling). Fractions were washed 5 times in 20 mM HEPES, 150 mM NaCl at pH 7.4. 2 μ l of lambda phosphatase, or buffer alone was then added to the purified fractions and samples were left for one hour. Samples were then fractionated on an SDS-PAGE gel and stained with instant blue (Expedeon). Bands corresponding to TDP1 and dephosphorylated TDP1 were then excised, digested with trypsin and gel purified to a final volume of 5 μ l before mass spec analysis. Although peptide identification of fragments showed a good coverage of the catalytic domain no fragments corresponding to the N-terminus of TDP1 could be found. The probable cause for the lack of N-terminal fragments captured by mass spec was the large stretches of acidic amino acids within this domain. This combined with an increase in negative charge due to the clusters of phosphorylation meant that these fragments could not fly.

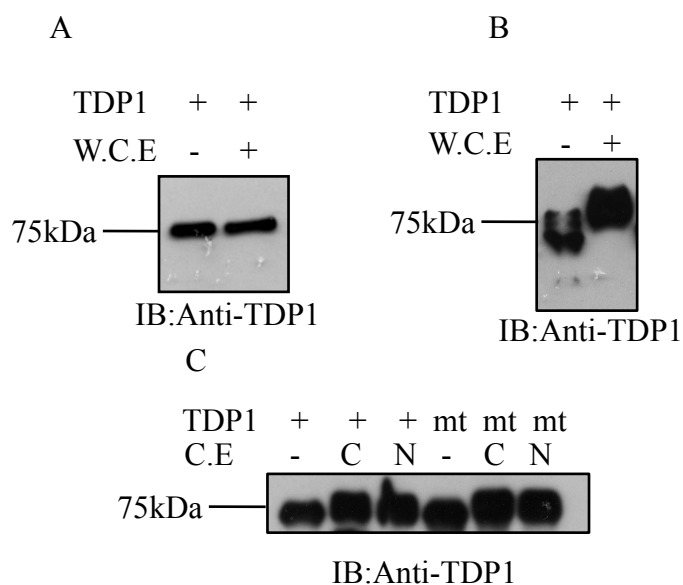


Figure 5.2.9 Purified His-TDP1 and TDP1 S81A are phosphorylated in the presence of cell extract, in an ATP dependent manner

(A) His-TDP1 (500 nM) was incubated with or without whole cell extract (W.C.E) from HEK 293 cells before fractionation and immunoblot analysis with anti-TDP1 antibody (Abcam; 1:2000). No shift in molecular weight could be seen between recombinant TDP1 and TDP1 treated with W.C.E. (B) His-TDP1 (500 nM) was incubated with or without whole cell extract (W.C.E) from HEK 293 cells, in the presence of ATP (5 mM) and Magnesium Chloride (2 mM). Samples were fractionated and analysed by immunoblotting with anti-TDP1 antibody (Abcam; 1:2000) and shows that when in the presence of ATP and W.C.E recombinant TDP1 can be seen at a higher molecular weight. (C) His-TDP1 '+', or His-TDP1 S81A 'mt' (500 nM) was incubated in the presence or absence of with or without nuclear 'N' or cytosolic 'C' extract from HEK 293 cells with ATP (5 mM) and Magnesium Chloride (2 mM). Immunoblot analysis shows that both TDP1 and TDP1 S81A in the presence of either nuclear or cytoplasmic extract results in a higher molecular weight band than TDP1, or TDP1 S81A control.

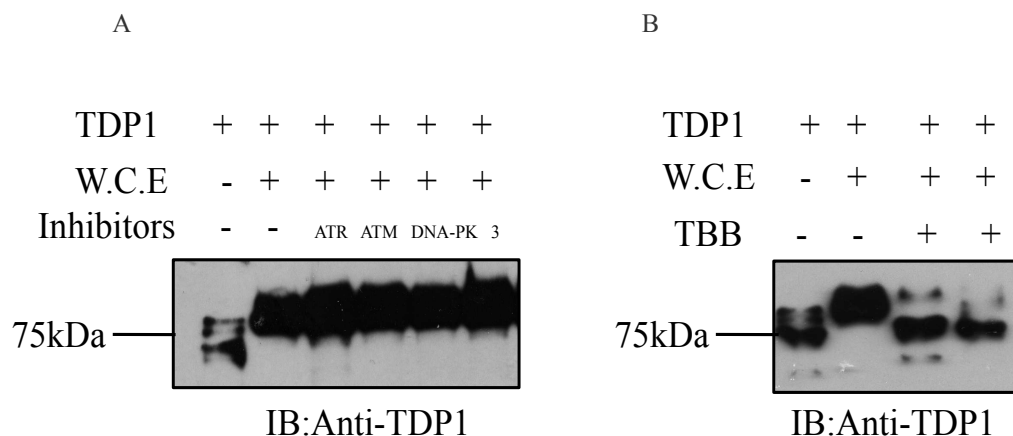


Figure 5.2.10 Purified His-TDP1 is phosphorylated by CK2

(A) His-TDP1 (500 nM) was incubated in the presence '+' or absence '-' of whole cell extract (W.C.E) from HEK 293 cells. Whole cell extracts were pre-incubated with or without 10 μ M ATR, ATM, DNA-PK inhibitors, or all three '3' for 1 hour at 30 °C. Immunoblot analysis reveals that none of the kinase inhibitors are able to prevent phosphorylation of TDP1. (B) His-TDP1 (500 nM) was incubated in the presence '+' or absence '-' of whole cell extract (W.C.E) from HEK 293 cells. Whole cell extracts were pre-incubated with or without 400 μ M TBB. The presence of TBB had a significant affect upon the phosphorylation of TDP1 in the presence of W.C.E. Immunoblot analysis revealed that in the presence of TBB, TDP1 is observed to run at the same molecular weight as recombinant TDP1 without W.C.E.

5.3 Discussion

Previous work has shown that TDP1 interacts with ligase III α and is phosphorylated in response to DNA damage by ATM or DNA-PK. Our work reveals that while TDP1 interacts with Lig III α this interaction is not critical for the repair of TDP1-mediated nuclear repair. Our results confirm that whilst Lig III α interacts with the N-terminal vertebrate-conserved region of TDP1, the C-terminal evolutionarily conserved domain of TDP1 is sufficient to maintain the interaction with Lig I. Our results establish a role for Lig I during TDP1-mediated nuclear DNA repair.

We show for the first time that TDP1 can interact with another member of the ligase family, Lig I and that this is mediated via TDP1's catalytic domain. Preliminary results show that cells lacking an active Lig I show an increase in DNA single-strand breaks when measured using the alkaline comet assay. Further experiments must be done to see if these results can be repeated and that the damage seen is not due to okazaki fragments that are not sealed by Lig I during replication. Whilst the theory that Lig I accompanies TDP1-mediated nuclear repair and that the Lig III-TDP1 interaction is necessary for the efficient repair of mitochondrial DNA is an appealing one, further work must be done to rule out a redundancy between these ligases in the context of nuclear DNA repair. Mckinnons' lab have already shown a redundancy between ligases I and III α , in terms of nuclear repair, when subjected to IR and hydrogen peroxide (Gao et al., 2011) and that Lig III α is critical for mitochondrial repair (Simsek et al., 2011) and this may well be the case in the context of chromosomal Top1-linked breaks brought on by CPT.

Our studies also reveal that when Myc-TDP1, expressed in mammalian cells, is treated with lambda phosphatase we see a moderate shift in gel mobility when compared to the untreated sample. We further identify that this shift in gel mobility is only seen within the N-terminus of TDP1 and, moreover, with the use of inhibitors demonstrate that TDP1 is a novel substrate for CK2 *in vitro*. We also establish a possible role for CK2, *in vitro*, where phosphorylation of TDP1 by CK2 negatively regulates SUMOylation of recombinant TDP1.

With regards to phosphorylation, we have discovered a constitutive phosphorylation of TDP1 by CK2 and mapped this modification within TDP1's evolutionarily driven N-terminal domain. *In vitro* experiments point to a role in regulation of SUMOylation of TDP1, and that de-phosphorylation of TDP1 allows post-translational modification of TDP1 by SUMO1. This leads to the possibility that SUMOylation of TDP1 is in fact damage dependent. Whilst SUMO1 pools are predominantly in low supply (Minty et al., 2000; Saitoh and Hinchey, 2000), several proteins have been identified to be SUMOylated in response to damage (Mao et al., 2000; Wang et al., 2005).

The reason that previous experiments did not uncover SUMOylation of TDP1 in response to damage could be due to the fact that ectopic expression of TDP1 when looking at SUMOylation *in vivo* gave rise to over an 8 fold increase in levels of TDP1 when compared to endogenous levels. It is therefore possible that this overexpression of TDP1 in cells meant that not all of the protein could be constitutively phosphorylated due to limiting supply of ATP or CK2. This could mean that the un-phosphorylated protein promoted SUMOylation of TDP1 independently of damage.

A possible scenario is that at endogenous levels of TDP1, all of the protein is constitutively phosphorylated by CK2. This in turn prevents the binding of UBC9 and hence modification by SUMO1. Upon response to DNA damage a phosphatase could de-phosphorylate a proportion of TDP1, which in turn could then be SUMOylated allowing an increase in levels of TDP1 at the sites of damage. To support this theory and to identify what, if any, other roles phosphorylation of TDP1 by CK2 may have it is vital to identify the sites of phosphorylation. In order to do this mutation of serine and tyrosine residues, identified by *in silico* analysis, to alanines should be considered. Alternatively, mutation of these residues to glutamic acid (phosphomimics) could be carried out. If phosphorylation does indeed negatively regulate SUMOylation, then phosphomimic mutations should prevent SUMOylation from occurring *in vivo*, resulting in a phenotype similar to that of the TDP1 K111R mutant described in the previous chapter.

One other possibility would be to overexpress CK2, if this is the rate-limiting step, in order to phosphorylate all ectopically expressed TDP1. Thus, upon overexpression of

CK2 SUMOylation of TDP1, if the theory holds true, should now be activated in response to DNA damage.

Chapter VI

Discussion: Cellular and biochemical analyses of TDP1 mediated chromosomal break repair

6.1 Discussion

This thesis demonstrates novel insights into the TDP1 repair pathway, elucidates new protein interacting partners and details the role of post-translational modifications of TDP1 at the cellular level.

A hypomorphic missense mutation in TDP1 has been linked to the autosomal recessive ataxia spinocerebellar ataxia with axonal neuropathy 1 (SCAN1) and mutations in other end-processing enzymes gives rise to overlapping attributes seen in SCAN1 patients (El-Khamisy, 2011; Fogel and Perlman, 2007; Reynolds et al., 2012; Takashima et al., 2002). Whilst Tdp1 depletion in mice shows a hypersensitivity to the Top1 inhibitor camptothecin, phenotypic attributes associated with SCAN1 patients, such as degeneration of the cerebellum are not fully recapitulated in the animal model (Katyal et al., 2007). Discussion between groups leads to different conclusions on why this might be the case. These range from the normal longevity of the mice when compared to humans, the differences in requirement for TDP1, and the postulations that TDP1 H493R found in the human disease could lead to more toxic protein-DNA linked lesions than when TDP1 is fully absent (Hirano et al., 2007), although the latter possibility has been dismissed by the recent results of the Boerkoel lab (personal communication with Sherif El-Khamisy).

Advancement in understanding and manipulation of technology has meant we can now investigate defective components of the DNA repair pathway that once would have caused embryonic lethality. Several groups are now using promoter-driven expression of Cre recombinase to knock out proteins in specific tissues. For example, the McKinnon laboratory has used this method to gain an insight into the requirement for specific DNA ligases within the repair pathways, and has shown that XRCC1 repairs neural DNA damage and is important in the genesis of cerebellar interneurons in mice. (Gao et al., 2011; Y. Lee et al., 2009; Simsek et al., 2011).

More recently this methodology has, for the first time, enabled the production of Cul4b deficient mice (Chen et al., 2012; Zhao and Y. Sun, 2012). These mice have demonstrated X-linked mental retardation-like symptoms and not just at the cellular

level. Mouse models for SCAN1 and AOA1 have given insights into the biochemistry behind the defects in repair, and have exhibited sensitivity to the same damaging agents found in patient cells lines. Unfortunately, for these recessive ataxias, they have not resulted in the neuropathological phenotypes found in humans. It is possible by creating knock-in mutants to increase ROS within the mouse brain; we could observe the cerebellar degeneration seen in patients. However, this may lead to further issues of whether the degeneration is due to the defect of the protein in question, or if the increase in ROS may saturate another component of the repair pathway. If this cannot be accounted for then the need for better model organisms will continue to grow. Model organisms that recapitulate the phenotypic pathology of patients will help to test therapies aimed at slowing the progression; through the help of ROS scavengers, by preventing the degeneration from occurring or by introducing the functional protein through lentiviral technology. Suitable models will also help us understand the neurotoxicity seen in several treatments for other diseases, such as cancer.

So, what novel findings and conclusions can we draw from this thesis?

In the first chapter, the data show for the first time that the N-terminal domain (only found in higher eukaryotes) is required for optimal protection against a variety of damaging agents. *In vitro* assays for recombinant full-length and the C-terminal domain of TDP1 show a similar catalytic activity, but *in vivo* viability and single-strand break repair assays elucidate a role for the N-terminal domain of TDP1 in conjunction with transcription. This is of particular interest for post-mitotic cells as they are not easily replaced and do not have the extra repair pathways found in the G2/S phase of the cell cycle in replicating cells.

Our research supports previously published work by the Pommier group that hTDP1 can sufficiently replace Tdp1 that is knocked out in chicken DT40 cells, even in the presence of genotoxic damaging agents such as CPT and MMS, but not IR (Alagoz et al., 2013; Murai et al., 2012). Furthermore, removal of the first 150 amino acids, corresponding to the N-terminal of TDP1, results in an increased sensitivity to CPT and MMS in terms of viability and at the level of DNA single-strand breaks when compared to full-length TDP1. Addition of transcription inhibitor DRB decreased the level of single-strand breaks, measured by the alkaline comet assay, found in the C-terminal

TDP1 stable cell line to levels comparable to the full-length protein. This implies a role for the N-terminal domain of TDP1 in the nuclear repair of transcriptionally associated breaks.

The creation of stable full-length hTDP1 and TDP1¹⁵¹⁻⁶⁰⁸ in DT40 Tdp1 -/- cell lines in the first results chapter also gave us the opportunity to look at the role of PARP in TDP1-mediated repair and contributed to our most recent publication (Alagoz et al., 2013). I found that pre-treatment of DT40 Tdp1 -/-; hTDP1 and DT40 Tdp1 -/-; hTDP1¹⁵¹⁻⁶⁰⁸ cells with the PARP inhibitor Olaparib prior to addition of CPT, resulted in cell viabilities comparable to Tdp1 -/- cells in the presence of CPT. Interestingly, addition of Olaparib to Tdp1 -/- cells prior to treatment with CPT showed no additive sensitivity (Alagoz et al., 2013). This indicated that TDP1 and PARP had a synergistic relationship in the repair of Top1-abortive complexes, a hypothesis confirmed by a concomitant publication (Das et al., 2014).

Whilst an interesting observation in its own right upon treatment with a different genotoxic agent (MMS) hTDP1 cells pre-treated with Olaparib resulted in a better survival compared to Tdp1 -/- cells under the same conditions. If PARP is required in the repair of Top1-DNA abortive complexes by TDP1 then this means that TDP1 could possibly repair Top1 independent lesions. Further investigations including the knock down of Top1 by another member of the lab concluded that TDP1 could indeed repair abasic sites without the requirement for PARP (Alagoz et al., 2013).

In summary the first results chapter showed:

- The N-terminal domain of TDP1 is unstructured.
- Loss of the N-terminal domain does not impact upon the structure, activity or stability TDP1.
- The N-terminal domain of TDP1 is required for optimal protection against DNA damage.
- The catalytic domain of TDP1 accrues more DNA single-strand breaks than full-length TDP1.
- The increase in DNA single-strand breaks due to removal of the N-terminus of TDP1 appears to be, in part, transcriptionally related.

The second results chapter builds on the discoveries of the first results chapter. We demonstrate a novel interaction between TDP1 and UBC9, the E2 enzyme of the SUMO pathway. Moreover, for the first time we show that TDP1 is covalently modified by SUMO1 and map the location of attachment to a lysine residue at position 111 in TDP1. This second results chapter looks at the effect SUMOylation has at the cellular level. Studies demonstrate that this modification occurs constitutively in cells, supporting the idea that the majority of SUMO1 is already conjugated to its target substrates, and that SUMOylation of TDP1 is required for the efficient repair of transcriptionally associated chromosomal single-strand breaks (J. J. R. Hudson et al., 2012). This is particularly important in post-mitotic cells as high levels of oxidative stress and transcription can lead to stalled Top1-DNA complexes and abasic sites, both of which are repaired by TDP1. These cells are not easily replaced, are required to be long-lived and therefore repair of these lesions in a timely fashion is essential for survival. This importance is highlighted by the missense hypomorphic mutation in TDP1 leading to SCAN1, an autosomal recessive ataxia resulting in cerebellar atrophy (Takashima et al., 2002).

Whilst not a complete story, the second chapter also provides evidence for a novel non-covalent interaction between TDP1 and SUMO2, or 3. We further map the location to the catalytic domain of TDP1 but issues with the stability of TDP1 containing missense mutations in the potential SIM sites and problems with expression of the various TDP1 fragments prevented us from pinpointing the motif/s with this domain. Promising initial experiments using tryptic digests to obtain soluble fragments of TDP1 has meant that mapping of the motifs is still a possibility. Furthermore, new *in silico* analysis strongly points to a particular interacting motif within TDP1 where multiple mutations in this short hydrophobic stretch maybe feasible to obtain a stable protein that can no longer physically interact with SUMO2 or 3.

Important findings from the second results chapter include:

- A novel covalent attachment of TDP1 at lysine 111 by SUMO1.
- SUMOylation of TDP1 is constitutive and is required for optimal function.
- SUMOylation localises TDP1 to sites of DNA damage that are in part transcriptionally dependent.

- TDP1 interacts non-covalently with SUMO2 and 3 via a SUMO interacting motif (SIM)
- The SIM is located within the catalytic domain of TDP1.

These findings have led to publication in Nature communications (J. J. R. Hudson et al., 2012) and resulted in a reasonable foundation for a continued project investigating the role of the SUMO-TDP1 non-covalent interaction within cells.

Finally the third results chapter establishes an interaction between TDP1 and Lig I, that has previously been unreported and suggests a role, whether a primary or redundant one in the repair of abortive nuclear Top1-DNA complexes. We also show that the catalytic domain of TDP1 is sufficient for the interaction with Lig I and suggest a possible role for repair in lower eukarya.

Further investigations must be done to solidify the evidence that Lig I deficiency leads to an increase in single-strand DNA breaks in the presence of CPT and that the defect seen is not due to an increase in unrepaired okazaki fragments or through competitive inhibition of alternative ligases due to the hypomorphic nature of the Lig I in these cell lines (Ellenberger and Tomkinson, 2008).

Cell lines comprising of Lig I, Lig III α , Lig IV single knock outs and Lig I Lig III α double knock outs will go a long way to answering these questions.

Perhaps more intriguing was the discovery that TDP1 is constitutively phosphorylated by CK2, in a similar fashion to that of the previously published XRCC1 protein. The determination of the phosphorylation being limited to the N-terminal domain gives rise to several questions. Does ablation of SUMOylation at K111 in TDP1 constitute the sole reason we see sub-optimal protection against various damaging agents when we remove the N-terminal domain of TDP1, or is this only one part of many roles the N-terminal domain has?

Initial investigations allude to a role for CK2 phosphorylation of TDP1 in terms of SUMOylation. Previous studies allude to phosphodependent motifs proximal to the site of SUMOylation that can act as a molecular switch for inhibition or activation of

SUMOylation on a particular protein (Hietakangas et al., 2006). If this is the case then initial findings lean towards the idea that TDP1 is constitutively phosphorylated by CK2 preventing SUMOylation at lysine 111. This means that a phosphatase dephosphorylates a small proportion of TDP1 and this fraction becomes SUMOylated in response to damage.

Locating the sites of phosphorylation is key to determining the role it has on the cellular activity of TDP1.

In summary this chapter presents several important findings:

- Ligase III α is not required for repair of nuclear abortive Top1-DNA complexes.
- Ligase I interacts with the catalytic domain of TDP1.
- TDP1 is phosphorylated by CK2.
- CK2 phosphorylation occurs within the N-terminal domain of TDP1.

Overall this thesis identifies several novel TDP1 protein interacting partners; including the sealing of nicks through interaction with Lig I or Lig III α . The role the N-terminal domain has in the repair of transcriptionally associated chromosomal DNA single-strand breaks through SUMOylation at K111R (J. J. R. Hudson et al., 2012), and that TDP1 is phosphorylated by CK2, within the N-terminus and that this may act as a molecular switch for TDP1 modification by SUMO1 in response to damage.

During this thesis, the role for the evolutionarily driven N-terminal domain of TDP1 for optimal protection against DNA damage was further investigated by the Pommier group. Their studies showed that the N-terminus of TDP1 is required for the interaction with PARP (Das et al., 2014) and that PARylation is important for recruitment of XRCC1 and TDP1's stability. This raises several interesting questions.

Firstly, previous publications have shown a direct interaction with the N-terminal domain of TDP1 and Lig III α (Chiang et al., 2010), since Lig III α directly interacts with XRCC1 (Ellenberger and Tomkinson, 2008) why does TDP1 also need to interact with XRCC1 (Das et al., 2014)?

It is possible that the interaction between XRCC1 and TDP1 is required for nuclear repair whereas the direct action between Lig III α and TDP1 is part of the mitochondrial repair pathway. If this is the case then knock down or knock out of nuclear Lig III α would not affect the interaction between TDP1 and XRCC1 through a co-IP.

Secondly, inhibition of PARP leads to instability of TDP1 as seen previously by mutation of S81 to an alanine, which prevents phosphorylation by ATM and DNA-PK. But, inhibition of PARP also leads to an increase in phosphorylation of TDP1 at S81. So which is more important for the stability of TDP1 - Parylation or phosphorylation? It would be interesting to look at the degradation rate of a TDP1 S81A mutant in the presence of PARP inhibitor and to see if this would further increase the instability of TDP1.

Finally, why does inhibition of PARP that has an epistatic relationship with TDP1 (Alagoz et al., 2013) result in an increase in phosphorylation at S81? It may well be a caveat within the cell programming that even though TDP1 can no longer repair Top1-DNA breaks, because PARP is inhibited, it still leads to the phosphorylation of TDP1 as the cell does not recognize that TDP1 is functionally incapacitated.

Since PARYlation, phosphorylation of TDP1 by ATM and CK2, SUMOylation and protein interaction with XRCC1, Lig III α and PARP1, all occur within the N-terminal domain (the first 185 amino acids), does removal of this seemingly regulatory domain result in an incapacitation of TDP1 against genotoxic agents. The TDP1¹⁵¹⁻⁶⁰⁸ gave an intermediate level of protection against DNA damaging agents but initial studies show that this truncation may still interact with PARP. This leads us to believe that the PARP1 interaction occurs between amino acid 151 and 185 of TDP1. Therefore, does the removal of these 34 amino acids prevent TDP1 from reaching the site of damage either through PARYlation? Cell viability assays using the DT40 Tdp1^{-/-}; hTDP1¹⁵¹⁻⁶⁰⁸ stable cell line from the first results chapter compared to DT40 Tdp1^{-/-} expressing hTDP1¹⁸⁶⁻⁶⁰⁸ and DT40 Tdp1^{-/-} cells would be particularly interesting.

Data accumulated from this thesis and other labs show several novel mechanisms for TDP1 mediated repair and demonstrate a newer model for TDP1's mode of action

(**Figure 6.1**). However, with new discoveries we also provide new questions and further work must be done in order to fully establish the role of TDP1 phosphorylation by CK2, the non-covalent interaction between SUMO2 and TDP1 and how SUMOylation of TDP1 results in localization to damage.

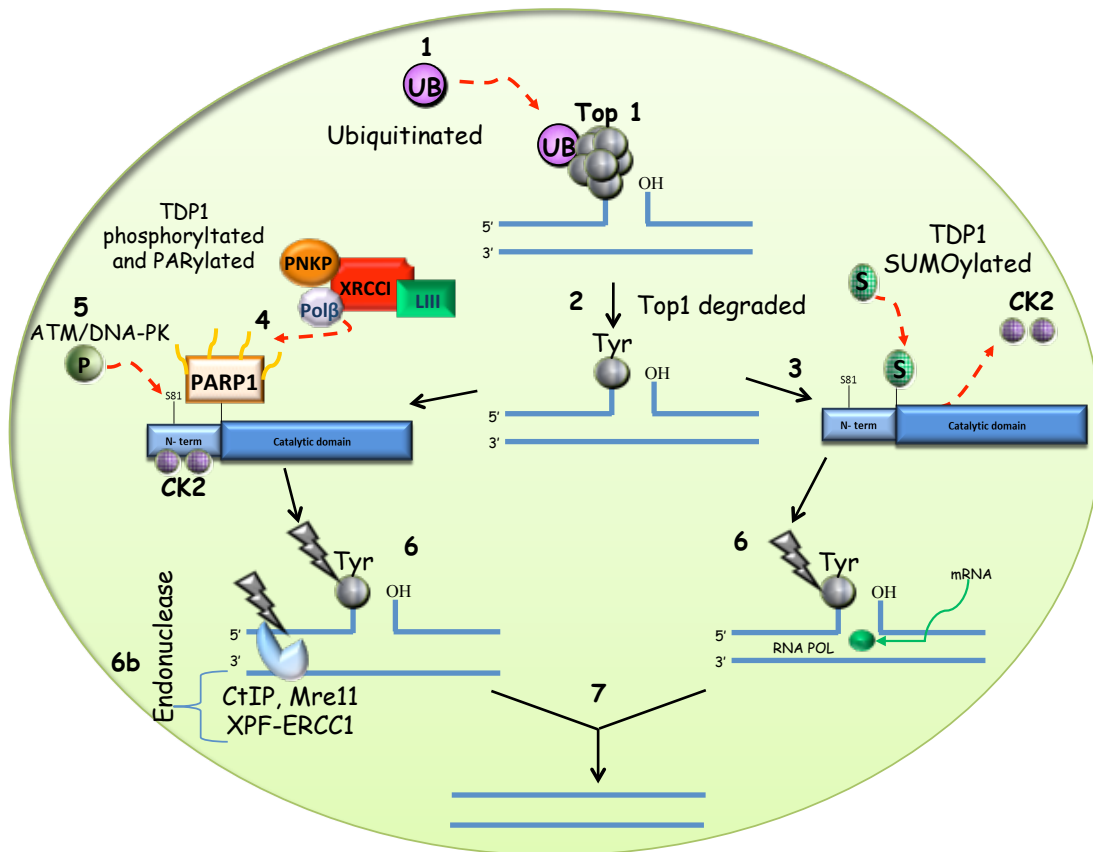


Figure 6.1 Schematic of TDP1 mediated nuclear repair

Upon collision with a DNA or RNA polymerase, or if in close proximity to a DNA lesion, Top1 becomes trapped forming a protein-DNA conjugate, termed an abortive Top1 complex. Top1 then becomes ubiquitylated (1) and degraded by the 26 S proteasome. Post degradation a tyrosine, and possibly small peptide fragment remains linked to the 3' phosphate of the nicked DNA (2). This 3' phosphotyrosine is the primary substrate of TDP1. If this abortive complex is associated due to transcription then either TDP1 becomes SUMOylated at K111 due to removal of constitutively phosphorylated residues within TDP1's N-terminal, or a small proportion of TDP1 that is constitutively SUMOylated (3) localizes to the site of damage (6). If not associated with blockage of RNA polymerases or DNA double-strand breaks arise then TDP1 becomes phosphorylated at S81 and PARylated within the N-terminus due to the PARP1-TDP1 complex (4, 5). PARylation recruits the molecular chaperone protein XRCC1 and both PARylation and phosphorylation lead to an increase in TDP1 stability. TDP1 is then recruited to the site of DNA damage (6). Once TDP1 localizes to the abortive Top1-DNA complex it catalyzes the removal of the tyrosine residue, leaving a 3' phosphate and 5' hydroxyl. Alternatively, replicating cells deficient in TDP1 can remove the lesion through endonuclease activity by creating a nick upstream of the break (6b). Once the lesion is removed the DNA is replaced and the backbone is sealed by Lig I or Lig III α . This restores the integrity of DNA and allows for replication restart and, or restoration of transcription.

Publications

Co-first author

Hudson, J.J.R., Chiang, S.-C., Wells, O.S., Rookyard, C., El-Khamisy, S.F., 2012. SUMO modification of the neuroprotective protein TDP1 facilitates chromosomal single-strand break repair. *Nat Commun* 3, 733.

Second author

Alagoz, M., Wells, O.S., El-Khamisy, S.F., 2013. TDP1 deficiency sensitizes human cells to base damage via distinct topoisomerase I and PARP mechanisms with potential applications for cancer therapy. *Nucleic Acids Research*.

Book chapter

Textbook: Animal models and movement disorders 2nd edition. Book chapter Autosomal Recessive Ataxias Due to Defects in DNA Repair. Wells, O.S., El-Khamisy, S.F. Submitted and due for release September 2014.

References

- Ahel, I., Rass, U., El-Khamisy, S.F., Katyal, S., Clements, P.M., McKinnon, P.J., Caldecott, K.W., West, S.C., 2006. The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. *Nature* 443, 713–716.
- Ahnesorg, P., Smith, P., Jackson, S.P., 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124, 301–313.
- Alagoz, M., Wells, O.S., El-Khamisy, S.F., 2013. TDP1 deficiency sensitizes human cells to base damage via distinct topoisomerase I and PARP mechanisms with potential applications for cancer therapy. *Nucleic Acids Research*.
- Alano, C.C., Garnier, P., Ying, W., Higashi, Y., Kauppinen, T.M., Swanson, R.A., 2010. NAD⁺ depletion is necessary and sufficient for poly(ADP-ribose) polymerase-1-mediated neuronal death. *Journal of Neuroscience* 30, 2967–2978.
- Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., Liberi, G., 2012. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* 151, 835–846.
- Anckar, J., Sistonen, L., 2007. SUMO: getting it on. *Biochem. Soc. Trans* 35, 1409–1413.
- Apps, R., Garwicz, M., 2005. Anatomical and physiological foundations of cerebellar information processing. *Nat. Rev. Neurosci.* 6, 297–311.
- Arakawa, H., Bednar, T., Wang, M., Paul, K., Mladenov, E., Bencsik-Theilen, A.A., Iliakis, G., 2012. Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells. *Nucleic Acids Research* 40, 2599–2610.
- Baba, D., Maita, N., Jee, J.-G., Uchimura, Y., Saitoh, H., Sugasawa, K., Hanaoka, F., Tochio, H., Hiroaki, H., Shirakawa, M., 2005. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* 435, 979–982.
- Babic, I., Cherry, E., Fujita, D.J., 2006. SUMO modification of Sam68 enhances its ability to repress cyclin D1 expression and inhibits its ability to induce apoptosis. *Oncogene* 25, 4955–4964.
- Barnes, D.E., Stamp, G., Rosewell, I., Denzel, A., Lindahl, T., 1998. Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr. Biol.* 8, 1395–1398.
- Barnes, D.E., Tomkinson, A.E., Lehmann, A.R., Webster, A.D., Lindahl, T., 1992. Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* 69, 495–503.

- Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., Becker, J., 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. *Journal of Molecular Biology* 280, 275–286.
- Beneke, S., 2012. Regulation of chromatin structure by poly(ADP-ribosyl)ation. *Front Genet* 3, 169.
- Bentley, D.J., Harrison, C., Ketchen, A.-M., Redhead, N.J., Samuel, K., Waterfall, M., Ansell, J.D., Melton, D.W., 2002. DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability. *Journal of Cell Science* 115, 1551–1561.
- Bergink, S., Jentsch, S., 2009. Principles of ubiquitin and SUMO modifications in DNA repair. *Nature* 458, 461–467.
- Bermejo, R., Doksan, Y., Capra, T., Katou, Y.M., Tanaka, H., Shirahige, K., Foiani, M., 2007. Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes & Development* 21, 1921–1936.
- Bernier-Villamor, V., Sampson, D.A., Matunis, M.J., Lima, C.D., 2002. Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108, 345–356.
- Bluml, S., Wisnowski, J.L., Nelson, M.D., Paquette, L., Gilles, F.H., Kinney, H.C., Panigrahy, A., 2012. Metabolic Maturation of the Human Brain From Birth Through Adolescence: Insights From In Vivo Magnetic Resonance Spectroscopy. *Cereb Cortex*.
- Boddy, M.N., Howe, K., Etkin, L.D., Solomon, E., Freemont, P.S., 1996. PIC 1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukaemia. *Oncogene* 13, 971–982.
- Bossis, G., Melchior, F., 2006. SUMO: regulating the regulator. *Cell Div* 1, 13.
- Boulton, J., 2001. Ataxia telangiectasia gene mutations in leukaemia and lymphoma. *J. Clin. Pathol.* 54, 512–516.
- Boyer, A.-S., Grgurevic, S., Cazaux, C., Hoffmann, J.-S., 2013. The human specialized DNA polymerases and non-B DNA: vital relationships to preserve genome integrity. *Journal of Molecular Biology* 425, 4767–4781.
- Brooks, S.C., Adhikary, S., Robinson, E.H., Eichman, B.F., 2013. Recent advances in the structural mechanisms of DNA glycosylases. *Biochim. Biophys. Acta* 1834, 247–271.
- Buerstedde, J.M., Takeda, S., 1991. Increased ratio of targeted to random integration after transfection of chicken B cell lines. *Cell* 67, 179–188.

- Calcerrada, P., Peluffo, G., Radi, R., 2011. Nitric oxide-derived oxidants with a focus on peroxynitrite: molecular targets, cellular responses and therapeutic implications. *Curr. Pharm. Des.* 17, 3905–3932.
- Caldecott, K.W., 2003. DNA single-strand break repair and spinocerebellar ataxia. *Cell* 112, 7–10.
- Caldecott, K.W., 2007. Mammalian single-strand break repair: mechanisms and links with chromatin. *DNA Repair* 6, 443–453.
- Caldecott, K.W., 2008. Single-strand break repair and genetic disease. *Nat Rev Genet* 9, 493–493.
- Caldecott, K.W., Aoufouchi, S., Johnson, P., Shall, S., 1996. XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular “nick-sensor” in vitro. *Nucleic Acids Research* 24, 4387–4394.
- Caldecott, K.W., McKeown, C.K., Tucker, J.D., Ljungquist, S., Thompson, L.H., 1994. An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Molecular and Cellular Biology* 14, 68–76.
- Caldecott, K.W., Tucker, J.D., Stanker, L.H., Thompson, L.H., 1995. Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells. *Nucleic Acids Research* 23, 4836–4843.
- Caviness, V.S., Takahashi, T., Nowakowski, R.S., 1995. Numbers, time and neocortical neurogenesis: a general developmental and evolutionary model. *Trends Neurosci* 18, 379–383.
- Champoux, J.J., 2001. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* 70, 369–413.
- Chance, B., Sies, H., Boveris, A., 1979. Hydroperoxide metabolism in mammalian organs. *Physiol. Rev.* 59, 527–605.
- Chen, C.-Y., Tsai, M.-S., Lin, C.-Y., Yu, I.-S., Chen, Y.-T., Lin, S.-R., Juan, L.-W., Chen, Y.-T., Hsu, H.-M., Lee, L.-J., Lin, S.-W., 2012. Rescue of the genetically engineered Cul4b mutant mouse as a potential model for human X-linked mental retardation. *Hum. Mol. Genet.* 21, 4270–4285.
- Chiang, S.-C., Carroll, J., El-Khamisy, S.F., 2010. TDP1 serine 81 promotes interaction with DNA ligase IIIalpha and facilitates cell survival following DNA damage. *Cell Cycle* 9, 588–595.
- Connelly, J.C., Leach, D.R.F., 2004. Repair of DNA covalently linked to protein. *Molecular Cell* 13, 307–316.

- Cortes Ledesma, F., El-Khamisy, S.F., Zuma, M.C., Osborn, K., Caldecott, K.W., 2009. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* 461, 674–678.
- Cox, J., Jackson, A.P., Bond, J., Woods, C.G., 2006. What primary microcephaly can tell us about brain growth. *Trends Mol Med* 12, 358–366.
- Das, B.B., Antony, S., Gupta, S., Dexheimer, T.S., Redon, C.E., Garfield, S., Shiloh, Y., Pommier, Y., 2009. Optimal function of the DNA repair enzyme TDP1 requires its phosphorylation by ATM and/or DNA-PK. *The EMBO Journal* 28, 3667–3680.
- Das, B.B., Dexheimer, T.S., Maddali, K., Pommier, Y., 2010. Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria. *Proceedings of the National Academy of Sciences* 107, 19790–19795.
- Das, B.B., Huang, S.-Y.N., Murai, J., Rehman, I., Amé, J.-C., Sengupta, S., Das, S.K., Majumdar, P., Zhang, H., Biard, D., Majumder, H.K., Schreiber, V., Pommier, Y., 2014. PARP1-TDP1 coupling for the repair of topoisomerase I-induced DNA damage. *Nucleic Acids Research*.
- Date, H., Onodera, O., Tanaka, H., Iwabuchi, K., Uekawa, K., Igarashi, S., Koike, R., Hiroi, T., Yuasa, T., Awaya, Y., Sakai, T., Takahashi, T., Nagatomo, H., Sekijima, Y., Kawachi, I., Takiyama, Y., Nishizawa, M., Fukuhara, N., Saito, K., Sugano, S., Tsuji, S., 2001. Early-onset ataxia with ocular motor apraxia and hypoalbuminemia is caused by mutations in a new HIT superfamily gene. *Nat. Genet.* 29, 184–188.
- Davies, D.R., Interthal, H., Champoux, J.J., Hol, W.G.J., 2002a. Insights into Substrate Binding and Catalytic Mechanism of Human Tyrosyl-DNA Phosphodiesterase (Tdp1) from Vanadate and Tungstate-inhibited Structures. *Journal of Molecular Biology* 324, 917–932.
- Davies, D.R., Interthal, H., Champoux, J.J., Hol, W.G.J., 2002b. The crystal structure of human tyrosyl-DNA phosphodiesterase, Tdp1. *Structure* 10, 237–248.
- Debéthune, L., Kohlhagen, G., Grandas, A., Pommier, Y., 2002. Processing of nucleopeptides mimicking the topoisomerase I-DNA covalent complex by tyrosyl-DNA phosphodiesterase. *Nucleic Acids Research* 30, 1198–1204.
- Desai, S.D., Liu, L.F., Vazquez-Abad, D., D'Arpa, P., 1997. Ubiquitin-dependent destruction of topoisomerase I is stimulated by the antitumor drug camptothecin. *J. Biol. Chem.* 272, 24159–24164.
- Desai, S.D., Zhang, H., Rodriguez-Bauman, A., Yang, J.-M., Wu, X., Gounder, M.K., Rubin, E.H., Liu, L.F., 2003. Transcription-dependent degradation of topoisomerase I-DNA covalent complexes. *Molecular and Cellular Biology* 23, 2341–2350.
- Desterro, J.M., Rodriguez, M.S., Hay, R.T., 1998. SUMO-1 modification of I κ B α inhibits NF- κ B activation. *Molecular Cell* 2, 233–239.

- Desterro, J.M., Rodriguez, M.S., Kemp, G.D., Hay, R.T., 1999. Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *J. Biol. Chem.* 274, 10618–10624.
- Deweese, J.E., Osheroff, M.A., Osheroff, N., 2008. DNA Topology and Topoisomerases: Teaching a “Knotty” Subject. *Biochem Mol Biol Educ* 37, 2–10.
- Dianov, G.L., Parsons, J.L., 2007. Co-ordination of DNA single strand break repair. *DNA Repair* 6, 454–460.
- Dohmen, R.J., 2004. SUMO protein modification. *Biochim. Biophys. Acta* 1695, 113–131.
- Eccles, L.J., O'Neill, P., Lomax, M.E., 2011. Delayed repair of radiation induced clustered DNA damage: friend or foe? *Mutat. Res.* 711, 134–141.
- Eid, W., Steger, M., El-Shemerly, M., Ferretti, L.P., Peña-Diaz, J., König, C., Valtorta, E., Sartori, A.A., Ferrari, S., 2010. DNA end resection by CtIP and exonuclease 1 prevents genomic instability. *EMBO reports* 11, 962–968.
- El-Khamisy, S.F., 2006. TDP1-dependent DNA single-strand break repair and neurodegeneration. *Mutagenesis* 21, 219–224.
- El-Khamisy, S.F., 2011. To live or to die: a matter of processing damaged DNA termini in neurons. *EMBO Mol Med* 3, 78–88.
- El-Khamisy, S.F., Caldecott, K.W., 2007. DNA single-strand break repair and spinocerebellar ataxia with axonal neuropathy-1. *Neuroscience* 145, 1260–1266.
- El-Khamisy, S.F., Hartsuiker, E., Caldecott, K.W., 2007. TDP1 facilitates repair of ionizing radiation-induced DNA single-strand breaks. *DNA Repair* 6, 1485–1495.
- El-Khamisy, S.F., Katyal, S., Patel, P., Ju, L., McKinnon, P.J., Caldecott, K.W., 2009. Synergistic decrease of DNA single-strand break repair rates in mouse neural cells lacking both Tdp1 and aprataxin. *DNA Repair* 8, 760–766.
- El-Khamisy, S.F., Saifi, G.M., Weinfeld, M., Johansson, F., Helleday, T., Lupski, J.R., Caldecott, K.W., 2005. Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature* 434, 108–113.
- Ellenberger, T., Tomkinson, A.E., 2008. Eukaryotic DNA Ligases: Structural and Functional Insights. *Annu. Rev. Biochem.* 77, 313–338.
- Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., Nordlund, P., 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical Biochemistry* 357, 289–298.
- Espinós-Armero, C., González-Cabo, P., Palau-Martínez, F., 2005. [Autosomal recessive cerebellar ataxias. Their classification, genetic features and pathophysiology]. *Rev Neurol* 41, 409–422.

- Fam, H.K., Chowdhury, M.K., Boerkoel, C.F., 2012. Spinocerebellar Ataxia with Axonal Neuropathy (SCAN1): A disorder of nuclear and mitochondrial DNA Repair. *Spinocerebellar Ataxia*.
- Fam, H.K., Chowdhury, M.K., Walton, C., Choi, K., Boerkoel, C.F., Hendson, G., 2013. Expression profile and mitochondrial colocalization of Tdp1 in peripheral human tissues. *J. Mol. Histol.* 44, 481–494.
- Faust, M., Montenarh, M., 2000. Subcellular localization of protein kinase CK2. A key to its function? *Cell Tissue Res* 301, 329–340.
- Ferrari, G., 2003. Cell Cycle-dependent Phosphorylation of Human DNA Ligase I at the Cyclin-dependent Kinase Sites. *Journal of Biological Chemistry* 278, 37761–37767.
- Fishel, M.L., Vasko, M.R., Kelley, M.R., 2007. DNA repair in neurons: So if they don't divide what's to repair? *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 614, 24–36.
- Fisher, A.E.O., Hocheegger, H., Takeda, S., Caldecott, K.W., 2007. Poly(ADP-ribose) polymerase 1 accelerates single-strand break repair in concert with poly(ADP-ribose) glycohydrolase. *Molecular and Cellular Biology* 27, 5597–5605.
- Floyd, R.A., West, M., Hensley, K., 2001. Oxidative biochemical markers; clues to understanding aging in long-lived species. *Experimental Gerontology* 36, 619–640.
- Fogel, B.L., Perlman, S., 2007. Clinical features and molecular genetics of autosomal recessive cerebellar ataxias. *Lancet Neurol* 6, 245–257.
- Frank, K.M., Sekiguchi, J.M., Seidl, K.J., Swat, W., Rathbun, G.A., Cheng, H.L., Davidson, L., Kangaloo, L., Alt, F.W., 1998. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* 396, 173–177.
- Frank, K.M., Sharpless, N.E., Gao, Y., Sekiguchi, J.M., Ferguson, D.O., Zhu, C., Manis, J.P., Horner, J., DePinho, R.A., Alt, F.W., 2000. DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Molecular Cell* 5, 993–1002.
- Gao, Y., Katyal, S., Lee, Y., Zhao, J., Rehg, J.E., Russell, H.R., McKinnon, P.J., 2011. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. *Nature* 471, 240–244.
- Geiss-Friedlander, R., Melchior, F., 2007. Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* 8, 947–956.
- Genschel, J., Bazemore, L.R., Modrich, P., 2002. Human exonuclease I is required for 5' and 3' mismatch repair. *J. Biol. Chem.* 277, 13302–13311.

- Geoffroy-Perez, B., Janin, N., Ossian, K., Laugé, A., Croquette, M.F., Griscelli, C., Debré, M., Bressac-de-Paillerets, B., Aurias, A., Stoppa-Lyonnet, D., Andrieu, N., 2001. Cancer risk in heterozygotes for ataxia-telangiectasia. *Int. J. Cancer* 93, 288–293.
- Goellner, G.M., Tester, D., Thibodeau, S., Almquist, E., Goldberg, Y.P., Hayden, M.R., McMurray, C.T., 1997. Different mechanisms underlie DNA instability in Huntington disease and colorectal cancer. *Am J Hum Genet* 60, 879–890.
- Goodarzi, A.A., Jeggo, P.A., 2012. The heterochromatic barrier to DNA double strand break repair: how to get the entry visa. *Int J Mol Sci* 13, 11844–11860.
- Goodarzi, A.A., Yu, Y., Riballo, E., Douglas, P., Walker, S.A., Ye, R., Härer, C., Marchetti, C., Morrice, N., Jeggo, P.A., Lees-Miller, S.P., 2006. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *The EMBO Journal* 25, 3880–3889.
- Gostissa, M., Hengstermann, A., Fogal, V., Sandy, P., Schwarz, S.E., Scheffner, M., Del Sal, G., 1999. Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *The EMBO Journal* 18, 6462–6471.
- Gough, D.R., Cotter, T.G., 2011. Hydrogen peroxide: a Jekyll and Hyde signalling molecule. *Cell Death and Disease* 2, e213–8.
- Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1, 2876–2890.
- Gupta, R.C., Lutz, W.K., 1999. Background DNA damage for endogenous and unavoidable exogenous carcinogens: a basis for spontaneous cancer incidence? *Mutat. Res.* 424, 1–8.
- Hada, M., Georgakilas, A.G., 2008. Formation of clustered DNA damage after high-LET irradiation: a review. *J Radiat Res* 49, 203–210.
- Hamilton, N.K., Maizels, N., 2010. MRE11 Function in Response to Topoisomerase Poisons Is Independent of its Function in Double-Strand Break Repair in *Saccharomyces cerevisiae*. *PLoS ONE* 5, e15387.
- Hanks, S.K., Hunter, T., 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* 9, 576–596.
- Hannich, J.T., Lewis, A., Kroetz, M.B., Li, S.-J., Heide, H., Emili, A., Hochstrasser, M., 2005. Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 280, 4102–4110.
- Hardeland, U., Steinacher, R., Jiricny, J., Schär, P., 2002. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *The EMBO Journal* 21, 1456–1464.

- Hartsuiker, E., Neale, M.J., Carr, A.M., 2009. Distinct requirements for the Rad32(Mre11) nuclease and Ctp1(CtIP) in the removal of covalently bound topoisomerase I and II from DNA. *Molecular Cell* 33, 117–123.
- Hawkins, A.J., Subler, M.A., Akopiants, K., Wiley, J.L., Taylor, S.M., Rice, A.C., Windle, J.J., Valerie, K., Povirk, L.F., 2009. In vitro complementation of Tdp1 deficiency indicates a stabilized enzyme-DNA adduct from tyrosyl but not glycolate lesions as a consequence of the SCAN1 mutation. *DNA Repair* 8, 654–663.
- Hecker, C.-M., Rabiller, M., Haglund, K., Bayer, P., Dikic, I., 2006. Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* 281, 16117–16127.
- Hegde, M.L., Hazra, T.K., Mitra, S., 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res* 18, 27–47.
- Heigener, D.F., 2011. Non-small cell lung cancer in never-smokers: a new disease entity? *Onkologie* 34, 202–207.
- Henderson, L.M., Arlett, C.F., Harcourt, S.A., Lehmann, A.R., Broughton, B.C., 1985. Cells from an immunodeficient patient (46BR) with a defect in DNA ligation are hypomutable but hypersensitive to the induction of sister chromatid exchanges. *Proceedings of the National Academy of Sciences* 82, 2044–2048.
- Heyer, W.-D., Ehmsen, K.T., Liu, J., 2010. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* 44, 113–139.
- Hibi, M., Shimizu, T., 2012. Development of the cerebellum and cerebellar neural circuits. *Dev Neurobiol* 72, 282–301.
- Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., Sistonen, L., 2006. PDSM, a motif for phosphorylation-dependent SUMO modification. *Proceedings of the National Academy of Sciences* 103, 45–50.
- Hirano, R., Interthal, H., Huang, C., Nakamura, T., Deguchi, K., Choi, K., Bhattacharjee, M.B., Arimura, K., Umehara, F., Izumo, S., Northrop, J.L., Salih, M.A.M., Inoue, K., Armstrong, D.L., Champoux, J.J., Takashima, H., Boerkoel, C.F., 2007. Spinocerebellar ataxia with axonal neuropathy: consequence of a Tdp1 recessive neomorphic mutation? *The EMBO Journal* 26, 4732–4743.
- Holliday, R., Ho, T., 1998a. Evidence for gene silencing by endogenous DNA methylation. *Proceedings of the National Academy of Sciences* 95, 8727–8732.
- Holliday, R., Ho, T., 1998b. Gene silencing and endogenous DNA methylation in mammalian cells. *Mutat. Res.* 400, 361–368.
- Hombauer, H., Campbell, C.S., Smith, C.E., Desai, A., Kolodner, R.D., 2011. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147, 1040–1053.

- Hsiang, Y.H., Hertzberg, R., Hecht, S., Liu, L.F., 1985. Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. *J. Biol. Chem.* 260, 14873–14878.
- Huang, S.-Y.N., Murai, J., Dalla Rosa, I., Dexheimer, T.S., Naumova, A., Gmeiner, W.H., Pommier, Y., 2013. TDP1 repairs nuclear and mitochondrial DNA damage induced by chain-terminating anticancer and antiviral nucleoside analogs. *Nucleic Acids Research* 41, 7793–7803.
- Hudson, E.K., Hogue, B.A., Souza-Pinto, N.C., Croteau, D.L., Anson, R.M., Bohr, V.A., Hansford, R.G., 1998. Age-associated change in mitochondrial DNA damage. *Free Radic. Res.* 29, 573–579.
- Hudson, J.J.R., Chiang, S.-C., Wells, O.S., Rookyard, C., El-Khamisy, S.F., 2012. SUMO modification of the neuroprotective protein TDP1 facilitates chromosomal single-strand break repair. *Nat Commun* 3, 733.
- Inamdar, K.V., Pouliot, J.J., Zhou, T., Lees-Miller, S.P., Rasouli-Nia, A., Povirk, L.F., 2002. Conversion of phosphoglycolate to phosphate termini on 3' overhangs of DNA double strand breaks by the human tyrosyl-DNA phosphodiesterase hTdp1. *J. Biol. Chem.* 277, 27162–27168.
- Interthal, H., Chen, H.J., Champoux, J.J., 2005a. Human Tdp1 cleaves a broad spectrum of substrates, including phosphoamide linkages. *J. Biol. Chem.* 280, 36518–36528.
- Interthal, H., Chen, H.J., Kehl-Fie, T.E., Zotzmann, J., Leppard, J.B., Champoux, J.J., 2005b. SCAN1 mutant Tdp1 accumulates the enzyme--DNA intermediate and causes camptothecin hypersensitivity. *The EMBO Journal* 24, 2224–2233.
- Interthal, H., Pouliot, J.J., Champoux, J.J., 2001. The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily. *Proceedings of the National Academy of Sciences* 98, 12009–12014.
- Isidor, B., Pichon, O., Baron, S., David, A., Le Caignec, C., 2010. Deletion of the CUL4B gene in a boy with mental retardation, minor facial anomalies, short stature, hypogonadism, and ataxia. *Am. J. Med. Genet. A* 152A, 175–180.
- Jackson, A.L., Chen, R., Loeb, L.A., 1998. Induction of microsatellite instability by oxidative DNA damage. *Proceedings of the National Academy of Sciences* 95, 12468–12473.
- Jackson, A.L., Loeb, L.A., 2001. The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutat. Res.* 477, 7–21.
- Jacobs, A.L., Schär, P., 2012. DNA glycosylases: in DNA repair and beyond. *Chromosoma* 121, 1–20.
- Jeppesen, D.K., Bohr, V.A., Stevnsner, T., 2011. DNA repair deficiency in neurodegeneration. *Prog. Neurobiol.* 94, 166–200.

- Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X.M., Scherer, S.W., Lasko, D.D., 1999. Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *J. Biol. Chem.* 274, 24176–24186.
- Johnson, E.S., Gupta, A.A., 2001. An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* 106, 735–744.
- Johnson, E.S., Schwienhorst, I., Dohmen, R.J., Blobel, G., 1997. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *The EMBO Journal* 16, 5509–5519.
- Kagey, M.H., Melhuish, T.A., Wotton, D., 2003. The polycomb protein Pc2 is a SUMO E3. *Cell* 113, 127–137.
- Kamitani, T., Kito, K., Nguyen, H.P., Fukuda-Kamitani, T., Yeh, E.T., 1998. Characterization of a second member of the sentrin family of ubiquitin-like proteins. *J. Biol. Chem.* 273, 11349–11353.
- Kasperek, T.R., Humphrey, T.C., 2011. DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Semin. Cell Dev. Biol.* 22, 886–897.
- Kass, E.M., Jasin, M., 2010. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Letters* 584, 3703–3708.
- Katayama, H., McGill, M., Kearns, A., Brzozowski, M., Degner, N., Harnett, B., Kornilayev, B., Matković-Calogović, D., Holyoak, T., Calvet, J.P., Gogol, E.P., Seed, J., Fisher, M.T., 2009. Strategies for folding of affinity tagged proteins using GroEL and osmolytes. *J. Struct. Funct. Genomics* 10, 57–66.
- Katyal, S., El-Khamisy, S.F., Russell, H.R., Li, Y., Ju, L., Caldecott, K.W., McKinnon, P.J., 2007. TDP1 facilitates chromosomal single-strand break repair in neurons and is neuroprotective in vivo. *The EMBO Journal* 26, 4720–4731.
- Katyal, S.S., McKinnon, P.J.P., 2007. DNA repair deficiency and neurodegeneration. Audio, Transactions of the IRE Professional Group on 6, 2360–2365.
- Katyal, S.S., McKinnon, P.J.P., 2008. DNA strand breaks, neurodegeneration and aging in the brain. *Mech Ageing Dev* 129, 9–9.
- Kedar, P.S., Stefanick, D.F., Horton, J.K., Wilson, S.H., 2012. Increased PARP-1 association with DNA in alkylation damaged, PARP-inhibited mouse fibroblasts. *Mol. Cancer Res.* 10, 360–368.
- Kelley, L.A., Sternberg, M.J.E., 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4, 363–371.
- Kelly, S.M., Jess, T.J., Price, N.C., 2005. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 1751, 119–

139.

- Kerscher, O., 2007. SUMO junction—what's your function? New insights through SUMO-interacting motifs. *EMBO reports* 8, 550–555.
- Kerzendorfer, C., Whibley, A., Carpenter, G., Outwin, E., Chiang, S.-C., Turner, G., Schwartz, C., El-Khamisy, S., Raymond, F.L., O'Driscoll, M., 2010. Mutations in Cullin 4B result in a human syndrome associated with increased camptothecin-induced topoisomerase I-dependent DNA breaks. *Hum. Mol. Genet.* 19, 1324–1334.
- Kirsh, O., Seeler, J.S., Pichler, A., Gast, A., Müller, S., Miska, E., Mathieu, M., Harel-Bellan, A., Kouzarides, T., Melchior, F., Dejean, A., 2002. The SUMO E3 ligase RanBP2 promotes modification of the HDAC4 deacetylase. *The EMBO Journal* 21, 2682–2691.
- Kulkarni, A.A., Wilson, D.M.D., 2008. The Involvement of DNA-Damage and -Repair Defects in Neurological Dysfunction. *Am J Hum Genet* 82, 28–28.
- Kunkel, T.A., 2003. Considering the cancer consequences of altered DNA polymerase function. *Cancer Cell* 3, 105–110.
- Kunkel, T.A., Erie, D.A., 2005. DNA MISMATCH REPAIR*. *Annu. Rev. Biochem.* 74, 681–710.
- Lakin, N.D., Weber, P., Stankovic, T., Rottinghaus, S.T., Taylor, A.M., Jackson, S.P., 1996. Analysis of the ATM protein in wild-type and ataxia telangiectasia cells. *Oncogene* 13, 2707–2716.
- Lakshmipathy, U., Campbell, C., 1999. The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Molecular and Cellular Biology* 19, 3869–3876.
- Lakshmipathy, U., Campbell, C., 2001. Antisense-mediated decrease in DNA ligase III expression results in reduced mitochondrial DNA integrity. *Nucleic Acids Research* 29, 668–676.
- Lane, N., 2006. Mitochondrial disease: powerhouse of disease. *Nature* 440, 600–602.
- Lange, S.S., Takata, K.-I., Wood, R.D., 2011. DNA polymerases and cancer. *Nature Publishing Group* 11, 96–110.
- Lavin, M.F., Gueven, N., Grattan-Smith, P., 2008. Defective responses to DNA single- and double-strand breaks in spinocerebellar ataxia. *DNA Repair* 7, 1061–1076.
- Lawrence, C.W., Borden, A., Banerjee, S.K., LeClerc, J.E., 1990. Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Research* 18, 2153–2157.
- Lawton, D.G., 2002. Interactions of the Type III Secretion Pathway Proteins LcrV and LcrG from *Yersinia pestis* Are Mediated by Coiled-Coil Domains. *Journal of Biological Chemistry* 277, 38714–38722.

- Le Ber, I., Moreira, M.-C., Rivaud-Péchoux, S., Chamayou, C., Ochsner, F., Kuntzer, T., Tardieu, M., Saïd, G., Habert, M.-O., Demarquay, G., Tannier, C., Beis, J.-M., Brice, A., Koenig, M., Dürr, A., 2003. Cerebellar ataxia with oculomotor apraxia type 1: clinical and genetic studies. *Brain* 126, 2761–2772.
- Lebedeva, N.A., Rechkunova, N.I., El-Khamisy, S.F., Lavrik, O.I., 2012. Tyrosyl-DNA phosphodiesterase 1 initiates repair of apurinic/apyrimidinic sites. *Biochimie* 94, 1749–1753.
- Lee, J.M., Niles, J.C., Wishnok, J.S., Tannenbaum, S.R., 2002. Peroxynitrite reacts with 8-nitropurines to yield 8-oxopurines. *Chem. Res. Toxicol.* 15, 7–14.
- Lee, Y., Chong, M.J., McKinnon, P.J., 2001. Ataxia telangiectasia mutated-dependent apoptosis after genotoxic stress in the developing nervous system is determined by cellular differentiation status. *Journal of Neuroscience* 21, 6687–6693.
- Lee, Y., Katyal, S., Li, Y., El-Khamisy, S.F., Russell, H.R., Caldecott, K.W., McKinnon, P.J., 2009. The genesis of cerebellar interneurons and the prevention of neural DNA damage require XRCC1. *Nat. Neurosci.* 12, 973–980.
- Lee, Y., McKinnon, P.J., 2007. Responding to DNA double strand breaks in the nervous system. *Neuroscience* 145, 1365–1374.
- Lees-Miller, S.P., Meek, K., 2003. Repair of DNA double strand breaks by non-homologous end joining. *Biochimie* 85, 1161–1173.
- Lehmann, A.R., McGibbon, D., Stefanini, M., 2011. Xeroderma pigmentosum. *Orphanet J Rare Dis* 6, 70.
- Leppard, J.B., Champoux, J.J., 2005. Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma* 114, 75–85.
- Levin, D.S., Bai, W., Yao, N., O'Donnell, M., Tomkinson, A.E., 1997. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. *Proceedings of the National Academy of Sciences* 94, 12863–12868.
- Liao, Z., 2006. Inhibition of human Tyrosyl-DNA Phosphodiesterase (Tdp1) by aminoglycoside antibiotics and ribosome inhibitors. *Molecular Pharmacology*.
- Lieber, M.R., Lu, H., Gu, J., Schwarz, K., 2008. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res* 18, 125–133.
- Lindahl, T., 1993. Instability and decay of the primary structure of DNA. *Nature* 362, 709–715.

- Lindahl, T., Karlström, O., 1973. Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution. *Biochemistry* 12, 5151–5154.
- Lindahl, T., Wood, R.D., 1999. Quality control by DNA repair. *Science* 286, 1897–1905.
- Litchfield, D.W., 2003. Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem. J.* 369, 1–15.
- Liu, B., Nicolaides, N.C., Markowitz, S., Willson, J., 1995. Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nature*.
- Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., 1996. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nature medicine*.
- Liu, H.C., Enikolopov, G., Chen, Y., 2012. Cul4B regulates neural progenitor cell growth. *BMC Neurosci* 13, 112.
- Liu, L., Lee, S., Zhang, J., Peters, S.B., Hannah, J., Zhang, Y., Yin, Y., Koff, A., Ma, L., Zhou, P., 2009. CUL4A abrogation augments DNA damage response and protection against skin carcinogenesis. *Molecular Cell* 34, 451–460.
- Liu, L., Yin, Y., Li, Y., Prevedel, L., Lacy, E.H., Ma, L., Zhou, P., 2012. Essential role of the CUL4B ubiquitin ligase in extra-embryonic tissue development during mouse embryogenesis. *Cell Res* 22, 1258–1269.
- Ljungman, M., Zhang, F., Chen, F., Rainbow, A.J., McKay, B.C., 1999. Inhibition of RNA polymerase II as a trigger for the p53 response. *Oncogene* 18, 583–592.
- Lobley, A., Whitmore, L., Wallace, B.A., 2002. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18, 211–212.
- Loizou, J.I., El-Khamisy, S.F., Zlatanou, A., Moore, D.J., Chan, D.W., Qin, J., Sarno, S., Meggio, F., Pinna, L.A., Caldecott, K.W., 2004. The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. *Cell* 117, 17–28.
- Mao, Y., Sun, M., Desai, S.D., Liu, L.F., 2000. SUMO-1 conjugation to topoisomerase I: A possible repair response to topoisomerase-mediated DNA damage. *Proceedings of the National Academy of Sciences* 97, 4046–4051.
- Marnett, L.J., 2003. Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein. *Journal of Clinical Investigation* 111, 583–593.
- Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J., de Murcia, G., 1998. XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Molecular and Cellular Biology* 18, 3563–3571.

- Matunis, M.J., Coutavas, E., Blobel, G., 1996. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *The Journal of Cell Biology* 135, 1457–1470.
- McArt, D.G., McKerr, G., Howard, C.V., Saetzler, K., Wasson, G.R., 2009. Modelling the comet assay. *Biochem. Soc. Trans* 37, 914.
- McKinnon, P.J., 2009. DNA repair deficiency and neurological disease. *Nat. Rev. Neurosci.* 10, 100–112.
- Meggio, F., Marin, O., Pinna, L.A., 1994. Substrate specificity of protein kinase CK2. *Cell. Mol. Biol. Res.* 40, 401–409.
- Meggio, F., Pinna, L.A., 2003. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* 17, 349–368.
- Melchior, F., 2000. SUMO--nonclassical ubiquitin. *Annu. Rev. Cell Dev. Biol.* 16, 591–626.
- Mimitou, E.P., Symington, L.S., 2009. Nucleases and helicases take center stage in homologous recombination. *Trends in Biochemical Sciences* 34, 264–272.
- Minty, A., Dumont, X., Kaghad, M., Caput, D., 2000. Covalent modification of p73alpha by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *J. Biol. Chem.* 275, 36316–36323.
- Modrich, P., 1994. Mismatch repair, genetic stability, and cancer. *Science*.
- Moore, S., Stanley, F.K.T., Goodarzi, A.A., 2014. The repair of environmentally relevant DNA double strand breaks caused by high linear energy transfer irradiation - No simple task. *DNA Repair*.
- Moreira, M.-C., Klur, S., Watanabe, M., Németh, A.H., Le Ber, I., Moniz, J.-C., Tranchant, C., Aubourg, P., Tazir, M., Schöls, L., Pandolfo, M., Schulz, J.B., Pouget, J., Calvas, P., Shizuka-Ikeda, M., Shoji, M., Tanaka, M., Izatt, L., Shaw, C.E., M'Zahem, A., Dunne, E., Bomont, P., Benhassine, T., Bouslam, N., Stevanin, G., Brice, A., Guimarães, J., Mendonça, P., Barbot, C., Coutinho, P., Sequeiros, J., Dürr, A., Warter, J.-M., Koenig, M., 2004. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.* 36, 225–227.
- Moreira, M.C., Barbot, C., Tachi, N., Kozuka, N., Mendonça, P., Barros, J., Coutinho, P., Sequeiros, J., Koenig, M., 2001a. Homozygosity mapping of Portuguese and Japanese forms of ataxia-oculomotor apraxia to 9p13, and evidence for genetic heterogeneity. *Am J Hum Genet* 68, 501–508.

- Moreira, M.C., Barbot, C., Tachi, N., Kozuka, N., Uchida, E., Gibson, T., Mendonça, P., Costa, M., Barros, J., Yanagisawa, T., Watanabe, M., Ikeda, Y., Aoki, M., Nagata, T., Coutinho, P., Sequeiros, J., Koenig, M., 2001b. The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. *Nat. Genet.* 29, 189–193.
- Murai, J., Huang, S.Y.N., Das, B.B., Dexheimer, T.S., Takeda, S., Pommier, Y., 2012. Tyrosyl-DNA Phosphodiesterase 1 (TDP1) Repairs DNA Damage Induced by Topoisomerases I and II and Base Alkylation in Vertebrate Cells. *Journal of Biological Chemistry* 287, 12848–12857.
- Nakamura, K., Kogame, T., Oshiumi, H., Shinohara, A., Sumitomo, Y., Agama, K., Pommier, Y., Tsutsui, K.M., Tsutsui, K., Hartsuiker, E., Ogi, T., Takeda, S., Taniguchi, Y., 2010. Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. *PLoS Genet* 6, e1000828.
- Nash, R.A., Caldecott, K.W., Barnes, D.E., Lindahl, T., 1997. XRCC1 protein interacts with one of two distinct forms of DNA ligase III. *Biochemistry* 36, 5207–5211.
- Nilsen, L., Forstrøm, R.J., Bjørås, M., Alseth, I., 2012. AP endonuclease independent repair of abasic sites in *Schizosaccharomyces pombe*. *Nucleic Acids Research* 40, 2000–2009.
- O'Driscoll, M., Cerosaletti, K.M., Girard, P.M., Dai, Y., Stumm, M., Kysela, B., Hirsch, B., Gennery, A., Palmer, S.E., Seidel, J., Gatti, R.A., Varon, R., Oettinger, M.A., Neitzel, H., Jeggo, P.A., Concannon, P., 2001. DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. *Molecular Cell* 8, 1175–1185.
- O'Driscoll, M., Jeggo, P.A., 2006. The role of double-strand break repair - insights from human genetics. *Nat Rev Genet* 7, 45–54.
- Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N., Yasuda, H., 1999. In vitro SUMO-1 modification requires two enzymatic steps, E1 and E2. *Biochemical and Biophysical Research Communications* 254, 693–698.
- Okura, T., Gong, L., Kamitani, T., Wada, T., Okura, I., Wei, C.F., Chang, H.M., Yeh, E.T., 1996. Protection against Fas/APO-1- and tumor necrosis factor-mediated cell death by a novel protein, sentrin. *J. Immunol.* 157, 4277–4281.
- Olsen, J.V., Ong, S.-E., Mann, M., 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell Proteomics* 3, 608–614.
- Otoshi, E., Yagi, T., Mori, T., Matsunaga, T., Nikaido, O., Kim, S.T., Hitomi, K., Ikenaga, M., Todo, T., 2000. Respective roles of cyclobutane pyrimidine dimers, (6-4)photoproducts, and minor photoproducts in ultraviolet mutagenesis of repair-deficient xeroderma pigmentosum A cells. *Cancer Res.* 60, 1729–1735.

- Pagano, M.A., Bain, J., Kazimierczuk, Z., Sarno, S., Ruzzene, M., Di Maira, G., Elliott, M., Orzeszko, A., Cozza, G., Meggio, F., Pinna, L.A., 2008. The selectivity of inhibitors of protein kinase CK2: an update. *Biochem. J.* 415, 353–365.
- Panier, S., Boulton, S.J., 2014. Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol* 15, 7–18.
- Parsons, J.L., Dianova, I.I., Finch, D., Tait, P.S., m, C.E.S., Helleday, T., Dianov, G.L., 2010. XRCC1 phosphorylation by CK2 is required for its stability and efficient DNA repair. *DNA Repair* 9, 835–841.
- Pascal, J.M., O'Brien, P.J., Tomkinson, A.E., Ellenberger, T., 2004. Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature* 432, 473–478.
- Pegg, A.E., 1984. Methylation of the O6 position of guanine in DNA is the most likely initiating event in carcinogenesis by methylating agents. *Cancer Invest.* 2, 223–231.
- Perez-Iratxeta, C., Andrade-Navarro, M.A., 2008. K2D2: estimation of protein secondary structure from circular dichroism spectra. *BMC Struct. Biol.* 8, 25.
- Pichler, A., Gast, A., Seeler, J.S., Dejean, A., Melchior, F., 2002. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* 108, 109–120.
- Pinna, L.A., Meggio, F., 1997. Protein kinase CK2 (“casein kinase-2”) and its implication in cell division and proliferation. *Prog Cell Cycle Res* 3, 77–97.
- Plo, I., 2003. Association of XRCC1 and tyrosyl DNA phosphodiesterase (Tdp1) for the repair of topoisomerase I-mediated DNA lesions. *DNA Repair* 2, 1087–1100.
- Pommier, Y., 2003. Repair of and checkpoint response to topoisomerase I-mediated DNA damage. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 532, 173–203.
- Pommier, Y., Barcelo, J.M., Rao, V.A., Sordet, O., Jobson, A.G., Thibaut, L., Miao, Z.H., Seiler, J.A., Zhang, H., Marchand, C., Agama, K., Nitiss, J.L., Redon, C., 2006. *Progress in Nucleic Acid Research and Molecular Biology*. Elsevier.
- Pourquier, P., Pommier, Y., 2001. Topoisomerase I-mediated DNA damage. *Adv. Cancer Res.* 80, 189–216.
- Prigent, C., Lasko, D.D., Kodama, K., Woodgett, J.R., Lindahl, T., 1992. Activation of mammalian DNA ligase I through phosphorylation by casein kinase II. *The EMBO Journal* 11, 2925–2933.
- Prigent, C., Satoh, M.S., Daly, G., Barnes, D.E., Lindahl, T., 1994. Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Molecular and Cellular Biology* 14, 310–317.

- Ramilo, C., Gu, L., Guo, S., Zhang, X., Patrick, S.M., Turchi, J.J., Li, G.-M., 2002. Partial reconstitution of human DNA mismatch repair in vitro: characterization of the role of human replication protein A. *Molecular and Cellular Biology* 22, 2037–2046.
- Rass, U., Ahel, I., West, S.C., 2007. Defective DNA Repair and Neurodegenerative Disease. *Cell* 130, 991–1004.
- Raymond, A.C., Rideout, M.C., Staker, B., Hjerrild, K., Burgin, A.B., Jr, 2004. Analysis of Human Tyrosyl-DNA Phosphodiesterase I Catalytic Residues. *Journal of Molecular Biology* 338, 895–906.
- Redinbo, M.R., Stewart, L., Kuhn, P., Champoux, J.J., Hol, W.G., 1998. Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* 279, 1504–1513.
- Rehman, M.T., Dey, P., Hassan, M.I., Ahmad, F., Batra, J.K., 2011. Functional Role of Glutamine 28 and Arginine 39 in Double Stranded RNA Cleavage by Human Pancreatic Ribonuclease. *PLoS ONE* 6, e17159.
- Ren, J., Gao, X., Jin, C., Zhu, M., Wang, X., Shaw, A., Wen, L., Yao, X., Xue, Y., 2009. Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. *Proteomics* 9, 3409–3412.
- Reynolds, J.J., El-Khamisy, S.F., Katyal, S., Clements, P., McKinnon, P.J., Caldecott, K.W., 2009. Defective DNA ligation during short-patch single-strand break repair in ataxia oculomotor apraxia 1. *Molecular and Cellular Biology* 29, 1354–1362.
- Reynolds, J.J., Walker, A.K., Gilmore, E.C., Walsh, C.A., Caldecott, K.W., 2012. Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair. *Nucleic Acids Research* 40, 6608–6619.
- Riballo, E., Woodbine, L., Stiff, T., Walker, S.A., Goodarzi, A.A., Jeggo, P.A., 2009. XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. *Nucleic Acids Research* 37, 482–492.
- Rivera-Calzada, A., Spagnolo, L., Pearl, L.H., Llorca, O., 2007. Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO reports* 8, 56–62.
- Rodriguez, M.S., Desterro, J.M., Lain, S., Midgley, C.A., Lane, D.P., Hay, R.T., 1999. SUMO-1 modification activates the transcriptional response of p53. *The EMBO Journal* 18, 6455–6461.
- Rydberg, B., Lindahl, T., 1982. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *The EMBO Journal* 1, 211–216.

- Sailer, A., Houlden, H., 2012. Recent advances in the genetics of cerebellar ataxias. *Curr Neurol Neurosci Rep* 12, 227–236.
- Saitoh, H., Hinchey, J., 2000. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J. Biol. Chem.* 275, 6252–6258.
- Sale, J.E., Lehmann, A.R., Woodgate, R., 2012. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Biol* 13, 141–152.
- San Filippo, J., Sung, P., Klein, H., 2008. Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* 77, 229–257.
- Sarikas, A., Hartmann, T., Pan, Z.-Q., 2011. The cullin protein family. *Genome Biol.* 12, 220.
- Sarno, S., Reddy, H., Meggio, F., Ruzzene, M., Davies, S.P., Donella-Deana, A., Shugar, D., Pinna, L.A., 2001. Selectivity of 4,5,6,7-tetrabromobenzotriazole, an ATP site-directed inhibitor of protein kinase CK2 (“casein kinase-2”). *FEBS Letters* 496, 44–48.
- Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., Jackson, S.P., 2007. Human CtIP promotes DNA end resection. *Nature* 450, 509–514.
- Schmutz, J., Grimwood, J., 2004. Genomes: fowl sequence. *Nature* 432, 679–680.
- Schreiber, V., Amé, J.-C., Dollé, P., Schultz, I., Rinaldi, B., Fraulob, V., Ménissier-de Murcia, J., de Murcia, G., 2002. Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J. Biol. Chem.* 277, 23028–23036.
- Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., Ganesh, V.S., Chang, B.S., Grix, A., Hill, R.S., Topcu, M., Caldecott, K.W., Barkovich, A.J., Walsh, C.A., 2010. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nat. Genet.* 42, 245–249.
- Shen, Z., Pardington-Purtymun, P.E., Comeaux, J.C., Moyzis, R.K., Chen, D.J., 1996. Associations of UBE2I with RAD52, UBL1, p53, and RAD51 proteins in a yeast two-hybrid system. *Genomics* 37, 183–186.
- Shiio, Y., Eisenman, R.N., 2003. Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences* 100, 13225–13230.
- Shrivastav, M., De Haro, L.P., Nickoloff, J.A., 2008. Regulation of DNA double-strand break repair pathway choice. *Cell Res* 18, 134–147.
- Sidman, R.L., Rakic, P., 1973. Neuronal migration, with special reference to developing human brain: a review. *Brain Res.* 62, 1–35.

- Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.-K., Van Houten, B., Shuman, S., McKinnon, P.J., Jasin, M., 2011. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* 471, 245–248.
- Smet-Nocca, C., Wieruszeski, J.-M., Léger, H., Eilebrecht, S., Benecke, A., 2011. SUMO-1 regulates the conformational dynamics of thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. *BMC Biochem.* 12, 4.
- Song, J., Durrin, L.K., Wilkinson, T.A., Krontiris, T.G., Chen, Y., 2004. Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences* 101, 14373–14378.
- Song, W., Pascal, J.M., Ellenberger, T., Tomkinson, A.E., 2009. The DNA binding domain of human DNA ligase I interacts with both nicked DNA and the DNA sliding clamps, PCNA and hRad9-hRad1-hHus1. *DNA Repair* 8, 912–919.
- Sordet, O., Khan, Q.A., Pommier, Y., 2004. Apoptotic topoisomerase I-DNA complexes induced by oxygen radicals and mitochondrial dysfunction. *Cell Cycle* 3, 1095–1097.
- Sreerama, N., Woody, R.W., 2000. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Analytical Biochemistry* 287, 252–260.
- Stemmler, T.L., Lesuisse, E., Pain, D., Dancis, A., 2010. Frataxin and mitochondrial FeS cluster biogenesis. *Journal of Biological Chemistry* 285, 26737–26743.
- Ström, C.E., Mortusewicz, O., Finch, D., Parsons, J.L., Lagerqvist, A., Johansson, F., Schultz, N., Erixon, K., Dianov, G.L., Helleday, T., 2011. CK2 phosphorylation of XRCC1 facilitates dissociation from DNA and single-strand break formation during base excision repair. *DNA Repair* 10, 961–969.
- Sun, S., Schiller, J.H., Gazdar, A.F., 2007. Lung cancer in never smokers--a different disease. *Nature Publishing Group* 7, 778–790.
- Suraweera, A., Lim, Y., Woods, R., Birrell, G.W., Nasim, T., Becherel, O.J., Lavin, M.F., 2009. Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Hum. Mol. Genet.* 18, 3384–3396.
- Sykora, P., Croteau, D.L., Bohr, V.A., Wilson, D.M., 2011. Aprataxin localizes to mitochondria and preserves mitochondrial function. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7437–7442.
- Takashima, H., Boerkoel, C.F., John, J., Saifi, G.M., Salih, M.A.M., Armstrong, D., Mao, Y., Quijcho, F.A., Roa, B.B., Nakagawa, M., Stockton, D.W., Lupski, J.R., 2002. Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. *Nat. Genet.* 32, 267–272.

- Tatham, M.H., 2001. Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9. *Journal of Biological Chemistry* 276, 35368–35374.
- Tatham, M.H., Chen, Y., Hay, R.T., 2003. Role of two residues proximal to the active site of Ubc9 in substrate recognition by the Ubc9.SUMO-1 thiolester complex. *Biochemistry* 42, 3168–3179.
- Taylor, R.W., Turnbull, D.M., 2005. Mitochondrial DNA mutations in human disease. *Nat Rev Genet* 6, 389–402.
- The Protein Kinase CK2 Facilitates Repair of Chromosomal DNA Single-Strand Breaks, 2004. *The Protein Kinase CK2 Facilitates Repair of Chromosomal DNA Single-Strand Breaks* 1–12.
- Tudek, B., Boiteux, S., Laval, J., 1992. Biological properties of imidazole ring-opened N7-methylguanine in M13mp18 phage DNA. *Nucleic Acids Research* 20, 3079–3084.
- Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., Pommier, Y., Tazi, J., Coquelle, A., Pasero, P., 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature Cell Biology* 11, 1315–1324.
- Ulrich, H.D., 2008. The Fast-Growing Business of SUMO Chains. *Molecular Cell* 32, 301–305.
- Uziel, T., Lerenthal, Y., Moyal, L., Andegeko, Y., Mittelman, L., Shiloh, Y., 2003. Requirement of the MRN complex for ATM activation by DNA damage. *The EMBO Journal* 22, 5612–5621.
- van Waardenburg, R.C.A.M., 2004. Platinated DNA Adducts Enhance Poisoning of DNA Topoisomerase I by Camptothecin. *Journal of Biological Chemistry* 279, 54502–54509.
- Vilk, G., Weber, J.E., Turowec, J.P., Duncan, J.S., Wu, C., Derksen, D.R., Zien, P., Sarno, S., Donella-Deana, A., Lajoie, G., Pinna, L.A., Li, S.S.C., Litchfield, D.W., 2008. Protein kinase CK2 catalyzes tyrosine phosphorylation in mammalian cells. *Cell. Signal.* 20, 1942–1951.
- Virág, L., Robaszkiewicz, A., Rodriguez-Vargas, J.M., Oliver, F.J., 2013. Poly(ADP-ribose) signaling in cell death. *Mol. Aspects Med.* 34, 1153–1167.
- Walker, J.R., Corpina, R.A., Goldberg, J., 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412, 607–614.
- Wallace, B.A., 2003. Analyses of circular dichroism spectra of membrane proteins. *Protein Science* 12, 875–884.

- Wallis, J.W., Aerts, J., Groenen, M.A.M., Crooijmans, R.P.M.A., Layman, D., Graves, T.A., Scheer, D.E., Kremitzki, C., Fedele, M.J., Mudd, N.K., Cardenas, M., Higginbotham, J., Carter, J., McGrane, R., Gaige, T., Mead, K., Walker, J., Albracht, D., Davito, J., Yang, S.-P., Leong, S., Chinwalla, A., Sekhon, M., Wylie, K., Dodgson, J., Romanov, M.N., Cheng, H., de Jong, P.J., Osoegawa, K., Nefedov, M., Zhang, H., McPherson, J.D., Krzywinski, M., Schein, J., Hillier, L., Mardis, E.R., Wilson, R.K., Warren, W.C., 2004. A physical map of the chicken genome. *Nature* 432, 761–764.
- Wang, Q.-E., Zhu, Q., Wani, G., El-Mahdy, M.A., Li, J., Wani, A.A., 2005. DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation. *Nucleic Acids Research* 33, 4023–4034.
- Ward, J.F., 1988. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog. Nucleic Acid Res. Mol. Biol.* 35, 95–125.
- Webster, A.D., Barnes, D.E., Arlett, C.F., Lehmann, A.R., Lindahl, T., 1992. Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. *Lancet* 339, 1508–1509.
- Weger, S., Hammer, E., Heilbronn, R., 2005. Topors acts as a SUMO-1 E3 ligase for p53 in vitro and in vivo. *FEBS Letters* 579, 5007–5012.
- Weinfeld, M., Mani, R.S., Abdou, I., Aceytuno, R.D., Glover, J.N.M., 2011. Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends in Biochemical Sciences* 36, 262–271.
- Weissman, L., de Souza-Pinto, N.C., Stevnsner, T., Bohr, V.A., 2007. DNA repair, mitochondria, and neurodegeneration. *Neuroscience* 145, 1318–1329.
- Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89, 392–400.
- World Health Organization, 2009. WHO Handbook on Indoor Radon. World Health Organization, Geneva.
- Yakes, F.M., Van Houten, B., 1997. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proceedings of the National Academy of Sciences* 94, 514–519.
- Yamaguchi-Iwai, Y., Sonoda, E., Sasaki, M.S., Morrison, C., Haraguchi, T., Hiraoka, Y., Yamashita, Y.M., Yagi, T., Takata, M., Price, C., Kakazu, N., Takeda, S., 1999. Mre11 is essential for the maintenance of chromosomal DNA in vertebrate cells. *The EMBO Journal* 18, 6619–6629.
- Yamazoe, M., Sonoda, E., Hocheegger, H., Takeda, S., 2004. Reverse genetic studies of the DNA damage response in the chicken B lymphocyte line DT40. *DNA Repair* 3, 1175–1185.

- Yang, S.-H., Sharrocks, A.D., 2004. SUMO promotes HDAC-mediated transcriptional repression. *Molecular Cell* 13, 611–617.
- Yang, S.W., Burgin, A.B., Huizenga, B.N., Robertson, C.A., Yao, K.C., Nash, H.A., 1996. A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proceedings of the National Academy of Sciences* 93, 11534–11539.
- Yu, S.-W., Wang, H., Poitras, M.F., Coombs, C., Bowers, W.J., Federoff, H.J., Poirier, G.G., Dawson, T.M., Dawson, V.L., 2002. Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* 297, 259–263.
- Yüce, O., West, S.C., 2013. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Molecular and Cellular Biology* 33, 406–417.
- Zandomeni, R., Mittleman, B., Bunick, D., Ackerman, S., Weinmann, R., 1982. Mechanism of action of dichloro-beta-D-ribofuranosylbenzimidazole: effect on in vitro transcription. *Proceedings of the National Academy of Sciences* 79, 3167–3170.
- Zeng, Z., Cortes Ledesma, F., El-Khamisy, S.F., Caldecott, K.W., 2011. TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *Journal of Biological Chemistry* 286, 403–409.
- Zeng, Z., Sharma, A., Ju, L., Murai, J., Umans, L., Vermeire, L., Pommier, Y., Takeda, S., Huylebroeck, D., Caldecott, K.W., El-Khamisy, S.F., 2012. TDP2 promotes repair of topoisomerase I-mediated DNA damage in the absence of TDP1. *Nucleic Acids Research* 40, 8371–8380.
- Zhao, Y., Sun, Y., 2012. CUL4B ubiquitin ligase in mouse development: a model for human X-linked mental retardation syndrome? *Cell Res* 22, 1224–1226.
- Zou, Y., Liu, Q., Chen, B., Zhang, X., Guo, C., Zhou, H., Li, J., Gao, G., Guo, Y., Yan, C., Wei, J., Shao, C., Gong, Y., 2007. Mutation in CUL4B, which encodes a member of cullin-RING ubiquitin ligase complex, causes X-linked mental retardation. *Am J Hum Genet* 80, 561–566.