

Characterisation of a New p97 Adaptor in Genome Stability

DPhil Thesis

Supervised by

Professor Kristijan Ramadan



**Katherine Elizabeth
Wiseman**



St Edmund Hall

CRUK/MRC Oxford Institute for Radiation Oncology

Department of Oncology

University of Oxford

Trinity Term 2017



CANCER
RESEARCH
UK

MRC

Medical
Research
Council

ACKNOWLEDGMENTS

First of all, I would like to thank Kristijan Ramadan for moving to the department at exactly the right time and allowing me to carry out my DPhil in his laboratory. His inputs, knowledge, enthusiasm, and guidance have been essential in getting me to this point. Thanks for never letting me give up!

One key member of the lab who I wouldn't be here without is Ignacio, the founding mass spec father of the entire project! Thank you for all of your hard work and discoveries, and your patience and teaching. Bruno, for all of your scientific inputs, but most of all, your friendliness and the good times, in and out of the lab. Swagat, a fellow DPhil and wonderful source of replication knowledge, who, along with Abhay, I shared many a tea with. Marta, my wonderful bench (invading) buddy. You always know what to say, helped me so much intellectually with my *in vitro* work, and managed to answer all of my self-doubting questions whilst making me feel like a capable scientist! I'll miss our coffee-sunshine and pub times! Judith, my Schmoooooo, thank you for everything! Ordering, coffee chameleon, co-ginoholic, shoulder to cry on, sanity maintainer, this definitely wouldn't have been possible without you. Also, thank you for proof-reading large portions of this thesis and keeping me going! Thank you to John, for all of your work on this project, hopefully it'll bring you many more results yet! Susan, my other bench wing-woman, for laughter in the lab! Also, to Divya, for tea, Indian weddings, and European adventures! As well as to previous lab members Domenic and Jenny who were a wonderful presence in our lab for a short time. Thank you to all of the Ramadan Lab!

A special thank you to all of my coffee-sunshine friends, for some respite between experiments. Especially Alexa, my other coffee chameleon, for keeping me company during long purifications, many a gin infused pub session, and introducing me to lindy hop! Also Sally, who's been there from the very beginning, through thick and thin, our sushi and Disney education evenings will have to continue! Claudia, for all of your purification knowledge, application magic, and extensive lab chats!

Thank you to the ORCRB team, Larry, Caroline, Maria, Rob and the stores team, for making our ordering challenges and lab maintenance perils easier!

Ally, Jessica, Jon, Cormac, Anja, Steph, and Nick, my college strange peeps. Thank you for wonderful memories in and out of college, which I'm sure will continue despite us spreading out around the world! Particularly thank you to Ally, for introducing me to the calming influence of Yogavenue. To Jenny and Becky, the greatest of friends, for our weekend getaways, always being there, and making sure I remember that there's a world outside of the lab and Oxford.

To Max, who's always been there for me, supporting me throughout everything, making sure I'm fed, and moving to Oxford with me. I couldn't have asked for a better partner in crime, thank you for everything, words don't seem enough.

Thank you to my parents for your support, encouragement, and advice throughout, and providing a haven to escape to.

DECLARATION

I declare that this thesis and the work therein is entirely my own work, aside from Figure 3-1, performed by Dr I Torrecilla, and Figure 6-4, Figure 6-5, and Figure 6-6, which were performed by J Fielden, who also created the CRISPR cell lines. All corresponding figure legends reflect these contributions.

ABSTRACT

Cells are subjected to constant genotoxic stress arising from both endogenous and exogenous sources which can drive the development of a wide range of diseases, especially cancers. Therefore, repair of these lesions is essential to maintain genomic stability and health. The DNA damage response is a key aspect of this process and requires the temporal and spatial regulation of many different factors to orchestrate the repair process. This requires the action of the AAA ATPase p97 which acts as a molecular segregase in order to facilitate the turnover of proteins on chromatin. This is also a crucial process to prevent the collision of DNA replication and transcription machinery with DNA bound proteins and DNA structures. These processes require the action of p97 cofactors which direct p97 to specific substrates and biological functions. The role of p97 and its cofactors in the DNA damage response is beginning to emerge, however there is still much to be investigated. Therefore, this thesis focuses on the characterisation of a new p97 adaptor protein discovered during a Mass Spectrometry analysis of the p97 interactome in response to ionising radiation. This adaptor protein, which has been named Ionising Radiation Enhanced VCP Interacting Protein 1, or IREVIP1, is responsible for maintaining genome stability even in unchallenged conditions. IREVIP1 was found to interact directly with p97 via a SHP domain located in the C-terminus of IREVIP1, as well as with Topoisomerases, and is responsible for bridging their interaction. IREVIP1 was observed as being part of the replication fork, and was required to prevent the accumulation of Topoisomerase cleavage complexes. Thus, IREVIP1 could be required to prevent genomic instability arising from collisions with Topoisomerase cleavage complexes and DNA structures arising from excessive DNA torsion. Therefore, this thesis begins to characterise the novel adaptor IREVIP1 and its functions within the cell.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	I
DECLARATION	II
ABSTRACT	III
LIST OF FIGURES	VIII
LIST OF TABLES	IX
LIST OF ABBREVIATIONS	X
1. INTRODUCTION	1
1.1 Genome Stability and Sources of Damage	1
1.2 Cell Cycle and Checkpoints.....	4
1.3 G1 and the G1-S Checkpoint	4
1.4 S phase and the Intra-S Phase Checkpoint.....	6
1.5 Mitosis and the G2-M Checkpoint	8
1.6 The DNA Damage Response	10
1.6.1 DNA-PK and NHEJ	10
1.6.2 ATM-CHK2, NHEJ and HR	11
1.6.3 NHEJ or HR Pathway Choice	12
1.6.4 ATR-CHK1	13
1.6.5 Other DNA Repair pathways	14
1.7 DNA Replication and Genome Instability	15
1.7.1 Replication-Transcription Collisions	20
1.8 Topoisomerases	23
1.8.1 Type I Topoisomerases.....	24
1.8.2 Topoisomerase I	24
1.8.3 Type II Topoisomerases.....	25
1.8.4 Topoisomerase II α	25
1.8.5 Topoisomerase Inhibitors and Topoisomerase Cleavage Complexes.....	27
1.9 The Ubiquitin Proteasome System.....	28
1.10 Ubiquitin and Ubiquitin Like Modifiers	29
1.11 p97.....	31
1.12 p97 and its Adaptors	33
1.12.1 p47.....	34
1.12.2 UFD1-NPL4	35
1.12.3 UBXD1	36
1.13 p97 and Protein Quality Control	36

1.14	p97 and the DNA Damage Response	38
1.15	p97 and DNA Replication	41
1.16	p97 and Topoisomerases	42
1.17	p97 and Cell Cycle Control	43
1.18	p97 and Disease	44
1.19	Aims of this Thesis.....	47
2.	MATERIALS AND METHODS	48
2.1	Buffers	48
2.1.1	10x Phosphate Buffer Saline (PBS), 1 L	48
2.1.2	10x Tris-Buffered Saline (TBS), 1 L.....	48
2.1.3	10x Tris/Glycine Buffer, 1 L	48
2.1.4	5x Laemmli Buffer (LB)	49
2.2	Plasmid Manipulation and Cloning	49
2.2.1	Sub-cloning.....	49
2.2.2	Mutagenesis	53
2.3	Bacterial Transformations	54
2.4	Native Purification.....	54
2.5	Denaturing Purification	55
2.6	ÄKTA system purification	56
2.7	In vitro interactions	57
2.8	Mammalian Cell Culture.....	58
2.8.1	Human Cell Lines	58
2.8.2	Generation of CRISPR/cas9 Knockout Cells.....	58
2.8.3	Maintenance of Human Cell Lines	58
2.8.4	Stable Cell Line Production	59
2.8.5	Transient siRNA Transfections	60
2.8.6	Plasmid Transfection	61
2.9	Preparation of Protein Lysates.....	61
2.9.1	Whole cell extracts (WCE)	61
2.9.2	Cellular Fractionation	61
2.10	SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)	62
2.11	Western Blotting	63
2.12	DNA Damage Induction.....	66
2.12.1	Ionising Radiation (IR)	66
2.13	Immunofluorescence (IF)	66
2.14	Flow Cytometry	67

2.15	Electromobility Shift Assay (EMSA).....	68
2.16	IREVIP1 Co-Immunoprecipitation (Co-IP) and Mass Spectrometry (MS).....	69
2.17	Isolation of Proteins on Nascent DNA (iPOND)	71
2.18	Rapid Approach to DNA Adduct Recovery (RADAR) DPC Isolation	71
2.19	DNA Fibre Assay	72
3.	RESULTS 1: IDENTIFICATION OF IREVIP1, A PREVIOUSLY UNCHARACTERISED PROTEIN ...	73
3.1	IREVIP1 Interacts With Key DNA Replication, Nuclear Pore, and Chromatin Topology Proteins	80
3.2	IREVIP1 Is Localised at the Nuclear Membrane by its Hydrophobic N-Terminal Domain 84	
3.3	Discussion.....	88
4.	RESULTS 2: PURIFICATION OF PROTEINS FOR <i>IN VITRO</i> STUDIES	91
4.1	Extensive Construct Optimisation was Required in Order to Purify IREVIP1.....	91
4.1.1	IREVIP1 N-Terminal GST	91
4.1.2	IREVIP1 C-terminal His	92
4.1.3	IREVIP1 C-terminal His and Hydrophobic N-terminal Removal	95
4.1.4	Mutants of IREVIP1 were Purified for Subsequent Biochemical Analysis.....	99
4.2	Discussion.....	106
5.	RESULTS 3: <i>IN VITRO</i> CHARACTERISATION OF IREVIP1	112
5.1	IREVIP1 Interacts with p97 <i>in vitro</i>	112
5.2	IREVIP1 Binds p97 via its SHP Domain	114
5.3	IREVIP1 Interacts with Topoisomerase II α and Topoisomerase I	114
5.4	p97 is Unable to bind Topoisomerase II α or Topoisomerase I Directly	115
5.5	p97 Requires IREVIP1 to Bridge the Interaction with Topoisomerase II α and Topoisomerase I.....	116
5.6	IREVIP1 is Capable of Binding Both Double Stranded and Single Stranded DNA	119
5.7	Discussion.....	120
6.	RESULTS 4: IREVIP1 AND GENOME STABILITY	123
6.1	IREVIP1 Depletion Causes Activation of the DNA Damage Response Signalling Pathway 123	
6.2	IREVIP1 Depletion Results in DNA Damage, Causing Histone H2AX Phosphorylation, but Reduces 53BP1 Recruitment, Both With and Without Exogenous DNA Damage	125
6.3	IREVIP1 Depletion Causes Cell Cycle Defects	128
6.4	IREVIP1 is a Part of the Replication Fork.....	130
6.5	IREVIP1 Depletion Results in Accumulation of Topoisomerase I and II α on DNA	131
6.6	IREVIP1 is required for efficient replication fork progression.....	132
6.7	Discussion.....	134

7. DISCUSSION	140
7.1 IREVIP1 is a Novel p97 Adaptor Able to Bind Topoisomerases.....	140
7.2 IREVIP1 is Located at the Nuclear Membrane, Where Torsional Stress is Often Exacerbated.....	144
7.3 IREVIP1 is Necessary to Prevent Genome Instability	145
7.4 Future Work	149
8. REFERENCES.....	150
9. APPENDIX	188
9.1 Publications	195

LIST OF FIGURES

Figure 1-1 The cell cycle and cell cycle checkpoints	5
Figure 1-2 Pre-Replicative Complex formation and replication initiation.....	17
Figure 1-3 Replication elongation	18
Figure 1-4 Replication termination	20
Figure 1-5 Topoisomerases are necessary to relieve the build-up of DNA torsion.....	23
Figure 1-6 The mechanisms of action of Topoisomerase I and Topoisomerase II α/β	26
Figure 1-7 p97 mechanism of action	32
Figure 1-8 Adaptor binding to p97	35
Figure 1-9 p97 is responsible for the removal of L3MBTL1 to allow 53BP1 recruitment to DSBs	39
Figure 3-1 p97EQ SILAC MS in response to IR.....	74
Figure 3-2 Domain schematic of IREVIP1	75
Figure 3-3 REVIP1 sequence alignment and protein characteristics	78
Figure 3-4 cBioPortal cross cancer alteration summary for IREVIP1	79
Figure 3-5 IREVIP1 MS interactions can be confirmed by Western blot	81
Figure 3-6 MS of IREVIP1 interacting proteins reveals its multi-pathway involvement.....	84
Figure 3-7 IREVIP1 is located at the nuclear rim via its hydrophobic N-terminal domain.....	87
Figure 4-1 IREVIP1 nGST is poorly soluble	92
Figure 4-2 Schematics of IREVIP1 constructs.....	94
Figure 4-3 IREVIP1 N Δ cHis has greater solubility	96
Figure 4-4 IREVIP1 N Δ cHis can be purified in batch under native and denaturing (8 M urea) conditions	98
Figure 4-5 Purification of IREVIP1 N Δ SHP Δ cHis	99
Figure 4-6 IREVIP1 N Δ SHP Δ cHis is similar to IREVIP1 N Δ cHis	101
Figure 4-7 Purification of IREVIP1 N Δ VBM* cHis.....	102
Figure 4-8 Purification of IREVIP1 N Δ SHP Δ VBM* cHis	103
Figure 4-9 Protein quantification and comparison of all four purified IREVIP1 cHis proteins .	104
Figure 4-10 Purification of p97cS●Tag	105
Figure 5-1 IREVIP1 interacts with p97 via its SHP domain in a salt dependent manner.....	113
Figure 5-2 IREVIP1, but not p97, interacts with Topoisomerase I and II α	115
Figure 5-3 IREVIP1 is required to bridge the interaction between p97 and Topoisomerase I or II α	118
Figure 5-4 IREVIP1 is capable of binding single stranded and double stranded DNA.....	120
Figure 6-1 IREVIP1 depletion causes activation of the DNA damage response	124
Figure 6-2 IREVIP1 depletion causes varied DNA damage response in a cell type dependent manner	127
Figure 6-3 IREVIP1 is required for normal cell cycle progression	129
Figure 6-4 IREVIP1 is part of the replication fork and prevents Topoisomerase I accumulation	131
Figure 6-5 Topoisomerase I and Topoisomerase II α DPCs accumulate in the absence of IREVIP1	132
Figure 6-6 Depletion of IREVIP1 results in DNA replication fork stalling, which can be rescued by Topoisomerase I, but not Topoisomerase II α , co-depletion.....	133
Figure 7-1 IREVIP1 is necessary to prevent genomic instability and interacts with p97 and Topoisomerases	148

Figure 9-1 IREVIP1 cHis enters denaturing but not native gels.....	188
Figure 9-2 IREVIP1 NTΔ cHis binds poorly, giving low elutions.....	189
Figure 9-3 IREVIP1 NTΔ cHis is found in the deal volume of Superdex 75 gel filtration column	190
Figure 9-4 IREVIP1 NTΔ cHis is found in the dead volume of Superdex 200 gel filtration column	191
Figure 9-5 IREVIP1 NTΔ cHis is lost during MonoQ Ion Exchange.....	192
Figure 9-6 IREVIP1 NTΔ cHis enters denaturing but not native gels.....	193
Figure 9-7 IREVIP1 NTΔ cHis purification under denaturing conditions.....	194

LIST OF TABLES

Table 1-1 Main DNA polymerases for different DNA repair pathways	15
Table 2-1 PCR reaction setup for Phusion DNA polymerase for sub-cloning	50
Table 2-2 Cycling conditions for Phusion DNA polymerase for sub-cloning	50
Table 2-3 List of plasmids	51
Table 2-4 List of Uniprot and NCBI Reference numbers.....	51
Table 2-5 List of primers	52
Table 2-6 PCR reaction setup for Phusion DNA polymerase for mutagenesis	53
Table 2-7 Cycling conditions for Phusion DNA polymerase for mutagenesis.....	53
Table 2-8 Purification constructs, growth temperatures and times, beads used, and elution conditions	54
Table 2-9 Buffers used during ÄKTA purification	56
Table 2-10 Interaction conditions for in vitro protein interactions.....	57
Table 2-11 siRNA sequences and working concentrations	60
Table 2-12 Homemade gel stacking and separating gel recipes	63
Table 2-13 List of Western blot antibodies and working concentrations	65
Table 2-14 List of Immunofluorescence antibodies and working concentrations	67
Table 2-15 Experiment outlines for flow cytometry for different cell lines	68
Table 2-16 Probe annealing for EMSA	69
Table 2-17 Reaction conditions for EMSA	69
Table 4-1 IREVIP1 purification optimisation.....	93

LIST OF ABBREVIATIONS

53BP1	p53 Binding Protein 1
AAA	ATPases Associated With Various Cellular Activities
ACS	ARS Consensus Sequences
AD	Alzheimer's Disease
ALS	Amyotrophic Lateral Sclerosis
APC/C	Anaphase-Promoting Complex Cyclosome
ATM	Ataxia-Telangiectasia Mutated
ATR	Ataxia Telangiectasia and Rad3-Related Protein
ATRIP	ATR-Interacting Protein
BER	Base Excision Repair
BLM	Bloom Syndrome Protein
BRCA	Breast Cancer Susceptibility Gene
BRCT	BRCA1 C-Terminus
BTR	BLM, Topoisomerase III α , and the RMI Proteins
CAD	Chromatin-Associated Degradation
CAK	CDK-Activating Kinase
Cdc	Cell Division Cycle
Cdh1	Cadherin 1
CDKs	Cyclin-Dependent Ser/Thr Kinases
CDT1	Chromatin Licensing and DNA Replication Factor 1
CFS	Chromosome Fragile Site
CHK	Checkpoint Kinase
CIN	Chromosomal Instability
CMG	CDC45, GINS, MCM
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRL	Cullin RING-Based E3 Ligase
CtIP	C-Terminal Binding Protein 1 Interacting Protein
DAPI	4',6-Diamidino-2-Phenylindole
DDB2	DNA Damage-Binding Protein 2
DDK	Dbf4-Dependent Kinase
DDR	DNA Damage Response
dHJ	Double Holliday Junction
DNA	Deoxyribonucleic Acid

DNA2	DNA Replication Helicase/Nuclease 2
DNA-PK	DNA-Dependent Protein Kinase
DNA-PKcs	DNA-Dependent Protein Kinase Catalytic Subunit
DOX	Doxycycline
DPC	DNA-Protein Crosslink
DSBs	Double Strand Breaks
dsDNA	Double Stranded DNA
DUB	Deubiquitinating Enzymes
E2F	E2 Factor
EME1	Essential Meiotic Structure-Specific Endonuclease 1
ER	Endoplasmic Reticulum
ERAD	Endoplasmic Reticulum-Associated Degradation
ERCC1	Excision Repair Cross-Complementation Group 1 (XPF)
ETAA1	Ewing's Tumour-Associated Antigen 1
EXO1	Exonuclease 1
FA	Formaldehyde
FAF1	Fas Associated Factor 1
FAM8A1	Family with Sequence Similarity 8, Member 1
FEN1	Flap Endonuclease 1
FTD	Frontotemporal Dementia
G0/1/2	Gap 0/1/2
GADD45a	Growth Arrest and DNA Damage Inducible Alpha
GCR	Gross Chromosomal Rearrangements
GFs	Growth Factors
GINS	Go, Ichi, Nii and San
GyrA	DNA Gyrase Subunit A
Gyrl	Gyrase Inhibitory Protein
HC	Heterochromatin
HD	Huntington's Disease
HECT	Homologous to the E6-AP Carboxyl Terminus
HERP	Homocysteine-Induced ER Protein
HR	Homologous Recombination
Hrd1	Hydroxymethyl Glutaryl-Coenzyme A Reductase Degradation Protein 1

IBMPFD	Inclusion Body Myopathy, Paget's Disease of the Bone and Frontotemporal Dementia
ICLs	Inter-Strand Crosslinks
iPOND	Isolation Of Proteins On Nascent DNA
IREVIP1	Ionising Radiation VCP Interacting Protein 1
JAMMs	JAB1/MPN/MOV34 Metalloenzymes
JMJD2A/B	Jumonji Domain-Containing Protein
KAP-1	KRAB-Associated Protein 1
L3MBTL1	Lethal (3) Malignant Brain Tumor-Like Protein 1
LB	Laemmli Buffer or Luria Broth Agar
M	Mitosis
MAD	Mitochondrial-Associated Degradation
MAT1	Ménage À Trois-1
MCM	Minichromosome Maintenance Protein
MDC1	Mediator of DNA Damage Checkpoint Protein 1
MDM2	Mouse Double Minute 2 Protein
MIN	Mini-Satellite Instability
MJD	Machado-Joseph Disease
MMEJ	Microhomology Mediated End Joining
MMR	Mismatch Repair
MPF	Mitosis Phase Promoting Factor
MRE11	Meiotic Recombination 11
MRN	MRE11-RAD50-NBS1
MS	Mass Spectrometry
MSH	Muts Homolog
MUS81	MMS and UV-Sensitive Protein 81
MVB	Multivesicular Body
NBS1	Nijmegen Breakage Syndrome Protein 1
NER	Nucleotide Excision Repair
NF- κ B	Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells
NHEJ	Non-Homologous End Joining
NPC	Nuclear Pore Complex
NPL4	Nuclear Protein Localisation Protein 4
NT	N-Terminal

NUPs	Nucleoporins
ORC	Origin Recognition Complex
Otu1/YOD1	Otubain 1
OTUs	Ovarian Tumour Proteases
p97	AAA ATPase p97 also known as VCP
PARP1	Poly [ADP-Ribose] Polymerase 1
PAXX	Paralog of XRCC4 and XLF
PBS	Phosphate Buffer Saline
PCNA	Proliferating Cell Nuclear Antigen
PDB	Paget's Disease Of The Bone
PI	Propidium Iodide
PI3K	Phosphoinositide 3-Kinase
PICHROS	Protein Induced Chromatin Stress
PIKK	Phosphoinositide 3-Kinase (PI3K)-Related Kinase
PIP	PCNA-Interacting Protein
PALB2	Partner and Localizer of BRCA2
Pol	Polymerase
Pre-RC	Pre-Replicative Complex
PRIM	Primase (DNA) Subunit 1
PTIP	PAX Transactivation Activation Domain-Interacting Protein
PTM	Posttranslational Modification
PUB	PNGase (Peptide N-Glycosidase)/Ubiquitin Associated
R Point	Restriction Point
RAD	Ribosomal-Associated Degradation
RAD51	Radiation Sensitive Protein
Rb	Retinoblastoma
RecQL4	RecQ Like Helicase 4
RFC	Replication Factor C
RIF1	Rap1-Interacting Factor 1
RING	Really Interesting New Gene
RIPA	Radioimmunoprecipitation Assay
R-Loops	RNA:DNA Hybrids
RNA	Ribonucleic Acid
RNF	RING Finger Protein

RPA	Replication Protein A
S Phase	Synthesis Phase
SCE	Sister Chromatid Exchange
SCF	SKP, Cullin, F-Box Containing Complex
SDSA	Synthesis-Dependent Strand Annealing
SEP	Shp1, Eyc and P47
SHP/BS1	Binding Site 1
SIM	SUMO Interacting Motif
SLBP	Stem-Loop Binding Protein
Sld	Synthetic Lethality with Dpb11
SLX	Synthetic Lethal Of Unknown Function
SPRTN	Sprt-Like Domain-Containing Protein Spartan
SSBs	Single Strand Breaks
ssDNA	Single Stranded DNA
SSH	Strep-Strep-HA
STUb1	SUMO-Targeted E3-Ubiquitin Ligase
SUMO	Small Ubiquitin-Like Modifier
TBS	Tris-Buffered Saline
TDP	Tyrosyl-DNA Phosphodiesterase 1
TdT	Terminal Deoxynucleotide Transferase
TFII	Transcription Factor II
TLS	Trans-Lesion Synthesis
TNR	Trinucleotide Repeats
TopBP1	DNA Topoisomerase II Binding Protein 1
Topo	Topoisomerase
Topo-ccs	Topoisomerase Cleavage Complexes
UAE	Ubiquitin-Activating Enzyme
Ub	Ubiquitin
UBA	Ubiquitin Associated
UBA6	Ubiquitin-Like Modifier Activating Enzyme 6
UBD	Ubiquitin D
U-Box	Ubiquitin-Box
UBX	Ubiquitin Regulatory X
UBXL	Ubiquitin Regulatory X Like

UBZ	Ubiquitin-Binding Zinc Finger
UCH	Ubiquitin C-Terminal Hydrolases
UDR	Ubiquitination-Dependent Recruitment
UFD1	Ubiquitin Fusion Degradation Protein 1
UN	UFD1-NPL4
UPS	Ubiquitin Proteasome System
USPs	Ubiquitin-Specific Proteases
UV	Ultraviolet
VBM	VCP-Binding Motif
VCIP135	Valosin-Containing Protein [VCP][P97]/P47 Complex-Interacting Protein, P135
VCP	Valosin Containing Protein
VIM	VCP Interacting Motif
WB	Western blot
WCE	Whole Cell Extracts
XLF	XRCC4-Like Factor
XPB	Xeroderma Pigmentosum, Complementation Group B
XPC	Xeroderma Pigmentosum, Complementation Group C
XPF	Xeroderma Pigmentosum, Complementation Group F (ERCC1)
XRCC	X-Ray Repair Cross Complementing
βMe	β-Mercaptoethanol
Δ	Deletion

1. INTRODUCTION

1.1 Genome Stability and Sources of Damage

Genome stability is important to prevent mutation and chromosome rearrangements which could potentially lead to genetic diseases, predisposition to the development and progression of cancer, and aging. Although such mutation can often be pathological, they are also the source of genetic variation and immunoglobulin diversification, therefore play a role in human evolution (Aguilera and Gómez-González, 2008; Hanahan and Weinberg, 2011).

Each cell in the human body experiences tens of thousands of DNA lesions every day, such as DNA deamination, oxidation, formation of abasic sites, pyrimidine dimers, DNA inter-strand crosslinks (ICLs), DNA-protein crosslinks (DPCs), single-strand breaks (SSBs), and double-strand breaks (DSBs) (Lindahl and Barnes, 2000; Jackson and Bartek, 2009). These can arise from both endogenous and exogenous sources, as well as spontaneous decomposition (Lindahl, 1993). The foremost exogenous source that comes to mind is solar ultraviolet light (UVA, UVB) which can induce around 100,000 lesions per cell per hour in strong sunlight. Ionising radiation (IR) generates a multitude of damage types, however the most lethal are DSBs. IR can arise from natural sources, such as radioactive decay, producing radon gas in some homes. It also occurs during various medical imaging techniques based on X-ray such as radiography, computer tomography, fluoroscopy, or cancer radiotherapy treatment, as well as from nuclear-reactor disasters and weapons. Chemicals can also be a source of exogenous damage, notably tobacco products, which can increase the risk of several cancers, including lung, oral, and surrounding tissue, as well as food-derived chemical. Endogenous damage can arise from processes within the cell, such as DNA replication, Topoisomerase enzymatic activity, hydrolytic reactions and non-enzymatic methylation, reactive oxygen species arising from oxidative respiration, redox-cycling events, and macrophages and neutrophils at inflammation and infection sites. Chemical

attack can give rise to DNA replication/transcription blocking base adducts, base loss, or cause SSBs (Jackson and Bartek, 2009).

Genome instability can present in many different forms, including chromosomal instability (CIN) where there is aneuploidy due to failures in mitotic segregation (Soto *et al.*, 2017), micro- or mini-satellite instability (MIN) from the expansion and contraction of repetitive DNA during replication slippage, mismatch repair or homologous recombination (HR) impairment (Cortes-Ciriano *et al.*, 2017). Mutations, insertions, and deletions from replication errors, as well as base excision repair (BER) or mismatch repair (MMR) impairment, or error prone trans-lesion synthesis (TLS) can also be a source of genomic instability (Davis *et al.*, 2012; Mosbech *et al.*, 2012). Chromosome rearrangements arising from DNA breaks where genetic linkages are either altered via unequal sister chromatid exchange (SCE) or end-joining resulting in gross chromosomal rearrangements (GCRs) can likewise be problematic (Chaganti, Schonberg and German, 1974; Lambert *et al.*, 2017).

Chromosome fragile sites (CFSs) are DNA sequences that accumulate gaps and breaks following DNA synthesis. Rare fragile sites account for less than 5% of sites and often arise from DNA repeat expansion during folate deficiency or DNA minor groove binders. They are generally associated with diseases such as fragile X mental retardation syndrome, Friedreich's ataxia, Huntington's disease, and myotonic dystrophy. On the other hand, common fragile sites account for 95% of sites and are considered to be naturally occurring and present in all individuals. However, they are a hotbed of translocation, gene amplification, exogenous DNA integration, and rearrangements, often correlating with chromosomal breakpoints in tumours. These sites are devoid of replication origins, replicate late on, and contain lengthy genes, often showing perturbed DNA elongation, suggesting that some inherent feature of fragile sites cause replication delays (Sutherland *et al.*, 1983; Glover *et al.*, 1984; Zlotorynski *et al.*, 2003; Marnef, Cohen and Legube, 2017).

CFSs are often accompanied by trinucleotide repeats (TNR) such as CGG-CCG, CAG-CTG, GAA-TTC, and GCN-NGC, with particularly G-rich or long AT-rich repeats. CFSs are wide spread and conserved between mice and human, and are also present in yeast and bacteria, suggesting their biological importance despite their inherent DNA structure being the molecular basis of their instability. Their instability is linked with their potential to form distinct secondary structure, such as hairpins and DNA triplexes. These structures cause replication perturbation, resulting in replication fork stalling and DSBs, resulting in DNA fragility (Cleary *et al.*, 2002; Zlotorynski *et al.*, 2003; Mirkin, 2007).

Error-free replication is necessary for the accurate duplication of chromosomes prior to chromosome segregation. DNA replication during S phase is a very susceptible portion of the cell cycle and the replisome must overcome many obstacles in order to complete replication successfully. Adducts, secondary structures and tightly bound proteins can cause fork stalling, resulting in ssDNA gaps and DSBs that could lead to development of sporadic, but also hereditary cancers, such as in SPRTN, BRCA2 or Fanconi anaemia proteins mutations (Fanconi, 1927; de Winter and Joenje, 2009; Stingele *et al.*, 2016; Vaz *et al.*, 2016; Ramadan *et al.*, 2017).

Transcription of DNA can increase the frequency of recombination and mutation as ssDNA is less stable than dsDNA. It is also possible for the transcribed strand to form an RNA:DNA hybrid, referred to as an R-loop, which have been reported to be important for immunoglobulin class switch recombination, DNA methylation prevention, and transcription termination (Nguyen *et al.*, 2017). However, they have also been associated with replication fork stalling, DSB formation and genome instability (Zamiri *et al.*, 2015; Bayona-Feliu *et al.*, 2017; Marnef, Cohen and Legube, 2017).

This ensuing DNA damage can be sensed by checkpoints as a potential anti-cancer barrier. This prevents fork collapse and breakdown, and activates the DNA damage response (DDR), thus allowing replication to be eventually restarted, maintaining genomic integrity. These genome

maintenance surveillance mechanisms ensure that the mutation rate remains low. However, in becoming tumorigenic, cancer cells increase the mutation rate in order to progress by a breakdown in one of the genome surveillance mechanism (Hanahan and Weinberg, 2011).

1.2 Cell Cycle and Checkpoints

The cell cycle is a controlled series of events which orchestrates the duplication and division of a cell into two daughter cells. This is split into key phases, namely synthesis (S) phase where DNA replication duplicates the genome, mitosis (M) where the chromosomes are condensed and separated into two daughter cells, and 'gap' phases G1 and G2 separating S phase and mitosis and where cell cycle checks are carried out. Cells can also enter G0, a dormant state. The progression between these different phases is controlled by activation and inactivation of cyclin-dependent Ser/Thr kinases (CDKs). Different CDKs partner with different cyclins throughout the cell cycle: cyclin D/CDK4/6 in G1, cyclin E/CDK2 in G1-S, cyclin A/CDK2/1 in S-G2, and cyclin B/CDK1 in M. The cycle is completed by the destruction of cyclins A and B, facilitated by ubiquitination by the E3 ligase anaphase-promoting complex cyclosome (APC/C), eliciting mitotic exit (**Figure 1-1**) (Hustedt and Durocher, 2017).

1.3 G1 and the G1-S Checkpoint

The initial transition through G1 requires extracellular signals, such as growth factors (GFs), until a point when the cell no longer requires external signals and is committed to DNA replication. This point is called the restriction (R) point and occurs 2-3 hours prior to DNA replication initiation. The R point relies on short lived regulatory proteins synthesised in response to growth factors, which accumulate to a tipping point before the R point can be passed. Cyclin D1 is such a protein and is essential for G1/S progression. GFs regulate cyclin D1 by transcription, stabilisation, nuclear localisation, and assembly with CDK4/6. Active cyclin D1/CDK4/6 is then responsible for the phosphorylation of the retinoblastoma (Rb) protein. Unphosphorylated Rb

prevents G1/S progression by sequestering E2F and preventing it initiating transcription of S phase genes. Suppression of Rb or E2F accumulation is sufficient to drive G1-S progression. Cyclin D1/CDK4/6 is not sufficient for this transition though. Cyclin E/CDK2, which reaches its maximum potential in late G1 in response to initiation of the R point, is also required and adds a second phosphorylation to Rb to bring about Rb degradation and fully remove GF-dependency (Croy and Pardee, 1983; Weinberg, 1995; Lundberg and Weinberg, 1998). However, E2F is also responsible for inducing p21 transcription, which brings about Rb dephosphorylation and therefore attenuates the R point (Blagosklonny and Pardee, 2002).

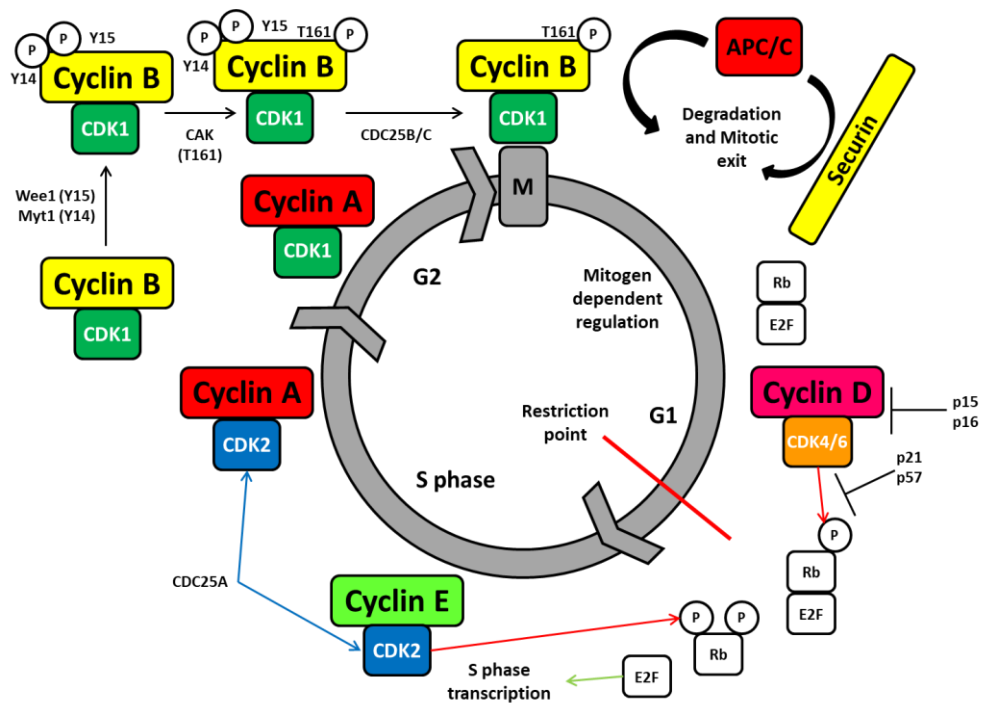


Figure 1-1 The cell cycle and cell cycle checkpoints

The cell cycle is split into four phases, which progression through is controlled by signalling cascades within the cell. Initially cells receive signals to begin mitogen dependent growth. This starts the G1 signalling cascade, causing expression of Cyclin D and its interaction with CDK4/6. Here, checkpoint signalling can kick in and prevent Cyclin D/CDK4/6 interaction or prevent their downstream signalling. Cyclin D/CDK4/6 is responsible for the initial phosphorylation of Rb, a protein which sequesters E2F. The upregulation and activation of Cyclin E/CDK2 from G1-S phase finishes off the phosphorylation of Rb, releasing E2F which initiates transcription of S phase genes. CDC25 phosphatases are responsible for removing the inhibitory phosphorylation from Cyclin/CDK complexes. Inhibition of CDC25A prevents activation of Cyclin D/CDK4/6, Cyclin E/CDK2, and Cyclin A/CDK2. Cyclin E/CDK2 and Cyclin A/CDK2/1 are necessary for DNA replication during S phase. Inactivation of these kinases during S phase allows time for DNA repair to be carried out before completion of DNA replication. Cyclin B/CDK1 is necessary for Mitosis and is controlled by inhibitory and activating phosphorylation. Cyclin B/CDK1 is phosphorylated by Wee1 and Myt1 in an inhibitory fashion, whereas CAK phosphorylates T161 in order for activation. CDC25B/C is required to remove the inhibitory phosphorylation, allowing Cyclin B/CDK1 to perform mitotic function. To exit Mitosis, Cyclin B/CDK1 degradation is brought about by the APC/C complex, along with securin, allowing separase to bring about sister chromatid separation during anaphase.

In response to DNA damage, the cell cycle can be arrested at the G1-S transition by accumulation of unphosphorylated Rb. This can be achieved by p53-independent and – dependent mechanisms. p53-independent mechanisms are fast and reversible. This can involve degradation of the phosphatase CDC25A following phosphorylation-dependent ubiquitination by the ubiquitin E3 ligase SCF^{BT_{RC}P} (Dozier *et al.*, 2017). CDC25A is one of three phosphatases in mammals, expressed from G1/S to G2 and is responsible for removal of the inhibitory phosphorylation from cyclin E/CDK2 following CDC25A stabilisation by cyclin D/CDK4/6 (Dozier *et al.*, 2017). However, CDC25A is also responsible for the degradation of cyclin D1 by the proteasome, in turn releasing p21 from CDK4 allowing CDK2 inhibition. p53-dependent pathways include the induction of p21 transcription and the inhibition of cyclin D/CDK4/6 and cyclin E/CDK2, thus preventing Rb phosphorylation. It is often this R point which is lost in cancer, via cyclin D overexpression, mutation of the CDK4/6 inhibitor p16, and loss of Rb function, often via p53 deletion or mutation or DNA tumour viruses such as the SV40 T antigen, E7 and E1A, allowing uncontrolled growth via Rb sequestering and p53 inhibition (Weinberg, 1995; Blagosklonny and Pardee, 2002; Karimian, Ahmadi and Yousefi, 2016; Vijayaraghavan *et al.*, 2017).

1.4 S phase and the Intra-S Phase Checkpoint

In addition to its control of Rb, cyclin E/CDK2 is also responsible for the phosphorylation of components of the pre-replicative complex (Pre-RC) (Jackson *et al.*, 1995; Woo and Poon, 2003). Cyclin A is also responsible for DNA replication and the commencement of mitosis, associating with CDK2 at the start of S phase, and CDK1 towards the end. Cyclin A has a C-terminal domain which is responsible for the binding of the replication licensing factor MCM7, which is essential for cyclin A promotion of S phase entry alongside cyclin A/CDK2 phosphorylation of other DNA replication initiation protein for Pre-RC formation (Chibazakura *et al.*, 2011).

The function of the S phase checkpoint is also to monitor for damaged DNA. As CDK2 is vital for S phase initiation, it is predictable that this is targeted by the DSB DNA damage checkpoint. In contrast to the G1 checkpoint which results in a slow, yet robust halt of DNA replication, the S phase checkpoint is rapid, yet short-lived, at the point of replication initiation and is largely p53-p21-independent. One such mechanism involves the protein kinase ATM, which can lead to the phosphorylation of another protein kinase CHK2 at Thr68 (Matsuoka *et al.*, 2000). This leads to CHK2, ATM, and CHK1 phosphorylation of CDC5A and its subsequent degradation, thus leaving CDK2 inactive, inhibiting S phase progression (Bornhorst and Falke, 2000; Woo and Poon, 2003). Another mechanism by which CDC25A dephosphorylation of CDK2 is prevented is via another protein kinase, ATR, which phosphorylates and activates the protein kinase CHK1 in response to ssDNA, which is discussed in greater detail later on. CHK1 also promotes the proteasomal degradation of CDC25A, thus preventing CDK2 activation, halting or preventing cell cycle progression, allowing time for DNA repair (Falck *et al.*, 2001). ATR is also responsible for the basal turnover of CDC25A to prevent its excessive accumulation thereby mediating its replication-promoting activity. In order for this signalling to increase this ATR-dependent degradation in the presence of DNA damage, CHK2 is activated, which further amplifies the signalling. Therefore, both ATM-CHK2 and ATR-CHK1 pathways work together to maintain accurate S phase progression (Falck *et al.*, 2001; Bartek, Lukas and Lukas, 2004; Blackford and Jackson, 2017).

Little is known about the progression from S phase to G2, when cyclin A/CDK1 is most active. S phase is when histone mRNA is transcribed to be able to assemble the newly replicated DNA into chromatin. Histone mRNA does not need splicing, instead it ends in a conserved stem-loop structure which needs to be bound by the stem-loop binding protein (SLBP) for mRNA 3' processing and translation. SLBP is also cell cycle-regulated and is quickly degraded at the end of S phase, initiated by Thr60/61 phosphorylation by cyclin A/CDK1, thus controlling histone biogenesis (Koseoglu, Graves and Marzluff, 2008). Cyclin A/CDK1 is also important for the

phosphorylation of the Orc1 subunit of the Origin recognition complex (ORC), responsible for deciding whether DNA replication starts at licensed origins, in order to prevent re-replication prior to mitosis (Li, Vassilev and DePamphilis, 2004). This is also the case for FEN1, a 5' exonuclease necessary for the maturation of Okazaki fragments during lagging strand DNA synthesis. Its phosphorylation prevents its interaction and stimulation by the replication protein PCNA, thus promoting S phase exit (Henneke, Koundrioukoff and Hubscher, 2003).

1.5 Mitosis and the G2-M Checkpoint

The transition from G2-M requires cyclin A and cyclin B, who both associate with CDK1. Cyclin B/CDK1 is necessary to coordinate mitotic cellular reorganisation, including microtubules, caspases, Golgi apparatus, the nucleolus, nuclear lamins, nuclear pore complexes, and the APC/C (Gavet and Pines, 2010). The cyclin B/CDK1 complex, or M phase promoting factor (MPF), is kept inactive during G2 by phosphorylation of Tyr15 and Tyr14 of CDK1 by Wee1 (Parker *et al.*, 1991) and Myt1 (Mueller *et al.*, 1995; Liu *et al.*, 1997) respectively. Cyclin B/CDK1 is also sequestered in the cytoplasm until their nuclear import just prior to nuclear envelope breakdown (Hagting *et al.*, 1998; Gavet and Pines, 2010). Nuclear CDK-activating kinase (CAK), composed of CDK7 (Fesquet *et al.*, 1993; Poon *et al.*, 1993; Solomon, Harper and Shuttleworth, 1993), Cyclin H (Fisher and Morgan, 1994; Makela *et al.*, 1994), and MAT1 (Tassan *et al.*, 1995), is then able to bind and phosphorylate CDK1 at its activating loop Thr161, further stabilising the interaction between CDK1 and cyclin B. All three of the mammalian CDC25 proteins (A (Galaktionov and Beach, 1991), B (Nagata *et al.*, 1991), and C (Sadhu *et al.*, 1990)) are involved in the G2-M transition. CDC25B is responsible for removing the inhibitory phosphorylations from cyclin B/CDK1, which in turn phosphorylates and activates CDC25C. This creates an autoactivation loop which drives the entry into mitosis (Booher, Holman and Fattaey, 1997; Hoffmann, 2000; Taylor and Stark, 2001; Coulonval, Kookken and Roger, 2011; Gardino and Yaffe, 2011). The E3 ubiquitin ligase APC/C and its activator Cdh1 are needed in G1 to prevent the premature accumulation of

cell cycle positive regulators. However, on G1-S transition, APC/C is inactivated by cyclin A/CDK2/1-mediated phosphorylation, which dissociates Cdh1 from APC/C. During G2, Cdh1 accumulates but APC/C^{Cdh1} remains inactive due to inhibitory CDK-dependent phosphorylations on Cdh1. Upon anaphase, Cdh1 is dephosphorylated by the yeast phosphatase Cdc14, however, the mammalian phosphatase responsible for this remains unclear (Bassermann *et al.*, 2008). The progression from metaphase to anaphase of mitosis requires MPF inactivation by degradation of cyclin B by APC/C^{Cdh1}. APC/C^{Cdh1} is also responsible for the degradation of securin, an inhibitor of separase. This, along with the alleviation of cyclin B/CDK1-mediated phosphorylation of separase, allows separase to remove cohesion from between sister chromatids at the onset of anaphase, allowing them to be pulled to different spindle poles (Chang, Xu and Luo, 2003).

The G2 checkpoint is also activated by DNA damage, either initiated in G2 or accumulating from S or G1. The main target of the G2-M checkpoint is the mitotic cyclin/CDK complex cyclin B/CDK1. This is achieved via the ATM/CHK2, ATR/CHK1 pathways regulating the CDC25 phosphatases required for G2-M transition. ATM/CHK2 and ATR/CHK1 can phosphorylate Ser 15/37 and Ser20 of p53 respectively, mediating upregulation of p21, which binds to and inhibits CDK1, production of 14-3-3 ϵ and γ , which binds to phosphorylated CDC25B/C (S323 and S216 respectively), and 14-3-3 σ , which binds and exports cyclin B/CDK1 to the cytoplasm, and GADD45a, which prevents CDK1 binding to cyclin B (Taylor and Stark, 2001; Dalal, Yaffe and DeCaprio, 2004). p53 and p21 can also bind to and inhibit the CAK kinase and CDK1 respectively, therefore prevent the activating phosphorylation of CDK1. p53 is also capable of downregulating Topoisomerase II, which is necessary for the G2-M transition in order to establish proper DNA topology. Often cells deficient in p53 tend to accumulate in G2-M, suggesting that this checkpoint can be mediated sufficiently by p53-independent mechanisms, however this arrest is not permanent (Taylor and Stark, 2001; Kastan and Bartek, 2004). One further layer of control exists, involving APC/C^{Cdh1} which is required for the degradation of the DDR mediator protein Claspin in order to relieve the G2-M checkpoint (Bassermann *et al.*, 2008).

1.6 The DNA Damage Response

In response to DNA damage, a signalling cascade is activated which initiates the DNA damage response (DDR) and cell cycle checkpoints to aid DNA repair and prevent further damage. The posttranslational modifications phosphorylation and ubiquitination are two key signalling events that regulate the DDR. Phosphorylation is carried out by a combination of three phosphoinositide 3-kinase (PI3K)-related kinases (PIKKs); DNA-PK (Carter *et al.*, 1990; Jackson *et al.*, 1990; Lees-Miller, Chen and Anderson, 1990), ATM (Savitsky *et al.*, 1995), and ATR (Bentley *et al.*, 1996). All three PIKKs have a preference for targeting serine/threonine followed by a glutamine (S/T-Q) motif for phosphorylation. Therefore, the substrate targeting of these three often have overlapping functions, as well as being able to autophosphorylate at S/T-Q motifs as well. In order to regulate their activity and prevent aberrant activation of the DDR, each kinase requires a co-factor to recruit it to damage sites; Ku80 for DNA-PK (Gell and Jackson, 1999; Singleton *et al.*, 1999), NBS1 for ATM (Falck, Coates and Jackson, 2005), and ATRIP for ATR (Zou and Elledge, 2003; Blackford and Jackson, 2017).

1.6.1 DNA-PK and NHEJ

DNA-PKcs, the catalytic subunit of the DNA-PK complex, is recruited to DNA and is activated by Ku to form DNA-PK (Mari *et al.*, 2006). Ku is able to bind dsDNA ends, thereby acting as a sensor for DSBs. DNA-PK plays a major role in promoting non-homologous end joining (NHEJ), which is responsible for repairing 80% of breaks not associated with replication forks when the homologous recombination (HR) pathway is not in use, throughout the cell cycle (Kakarougkas, Ismail, Klement, *et al.*, 2013). NHEJ repairs breaks by joining DNA ends without a homology template and therefore can be error-prone and mutagenic. This is especially the case in the absence of classical-NHEJ pathways (cNHEJ) when an Ku and DNA-PK independent alternative-NHEJ pathway (altNHEJ) must be used, which relies on PARP1, DNA ligase III and DNA polymerase θ (van Schendel *et al.*, 2015). cNHEJ is initiated by Ku and DNA-PKcs, where DNA-

PKcs anchors the DNA ends (Graham, Walter and Loparo, 2017). Additional factors are then recruited in order to align and ligate the ends, such as XRCC4, PAXX (Ochi *et al.*, 2015), XLF, and DNA ligase IV, as well as, in some cases, DNA polymerases and nucleases, such as Artemis (Blackford and Jackson, 2017; Hustedt and Durocher, 2017).

1.6.2 ATM-CHK2, NHEJ and HR

ATM is recruited to DSBs and is stimulated via its interaction with NBS1, a component of the MRE11-RAD50-NBS1 (MRN) complex. Around 15% of DSBs are located in heterochromatin (HC) which, in G0/G1, requires ATM activation of the NHEJ pathway by MRN-ATM activation of DNA-PK, promoting Artemis recruitment and phosphorylation of the HC factor KAP-1 (Kakarougkas, Ismail, Klement, *et al.*, 2013). Prior to this, ATM is responsible for phosphorylation of histone variant H2AX at Ser139 (γ H2AX) which marks DSBs and serves as a recruitment platform for downstream proteins and subsequent signalling. MDC1 is the main protein recruited to γ H2AX via its BRCT domain. MDC1 is then phosphorylated and stabilised by ATM, which in turn recruits more MRN-ATM, acting as a feedforward loop. Phosphorylated MDC1 is recognised by the ubiquitin E3 ligase RNF8, which, along with the ubiquitin-conjugating enzyme UBC13, ubiquitinates the linker histone H1. Ubiquitinated H1 recruits another E3 ligase, RNF168, which marks histone H2AK15 with K63 ubiquitin chains (Schwertman, Bekker-Jensen and Mailand, 2016). H2AK15Ub, along with H4K20me2, recruits 53BP1 via its UDR and tudor domains respectively. 53BP1 acts as a scaffold protein, following ATM-dependent N-terminal phosphorylation, for the recruitment of subsequent proteins, such as RIF1 and PTIP, important for promoting NHEJ (Panier and Boulton, 2014; Blackford and Jackson, 2017; Hustedt and Durocher, 2017). 53BP1 works with RIF1 in order to block end resection and promote NHEJ (Chapman *et al.*, 2013). 53BP1 and RIF1 could potentially be forming higher order structures which are capable of stabilising the chromatin and preventing nuclease access, thus preventing resection and favouring NHEJ (Panier and Boulton, 2014).

1.6.3 NHEJ or HR Pathway Choice

ATM is also involved in the HR pathway, which involves DNA resection in order to find a homology template for repair during S and G2. ATM phosphorylates CtIP, which in turn represses the NHEJ pathway by RNF138-dependent Ku80 ubiquitination and removal (Schwertman, Bekker-Jensen and Mailand, 2016). ATM also phosphorylates BRCA1 which further promotes HR over 53BP1-mediated NHEJ, relocating 53BP1 to the periphery of repair foci, possibly to prevent excessive resection (Nimonkar *et al.*, 2011; Chapman *et al.*, 2012; Kakarougkas, Ismail, Katsuki, *et al.*, 2013; Blackford and Jackson, 2017; Shibata, 2017). CDKs are also involved in activation of CtIP, RAD51 and MRE11, thus promoting HR and linking it with the cell cycle and sister chromatid availability. Furthermore, the lack of H4K20me2 on newly replicated chromatin prevents recruitment of 53BP1, and consequently NHEJ. Therefore, it can be said that 53BP1-RIF1 and BRCA1-CtIP compete for pathway choice and break site localisation (Hustedt and Durocher, 2017).

HR is carried out when the sister chromatid is available for use as a template. ATM-activated CtIP stimulates the 3'-5' endonuclease activity of MRE11, which works alongside the 5'-3' nucleases EXO1 and BLM/DNA2 to create a 3' overhanging ssDNA filament (Eid *et al.*, 2010; Nimonkar *et al.*, 2011). This ssDNA is coated with replication protein A (RPA), but is shortly replaced with RAD51 via the action of BRCA1, PALB2 and BRCA2 (Zhang *et al.* 2009), however its accumulation on ssDNA is antagonised by RADX in order to attenuate RAD51 mediated fork reversal, which can result in DNA damage and fork collapse (Dungrawala *et al.*, 2017). RAD51 is then capable of mediating a homology search, strand invasion, and initiating DNA synthesis to fill the break site. This process forms structures which can form D-loops or double Holliday junctions (dHJs) which need to be disassembled. D-loops can be unwound by helicases post-ligation via synthesis-dependent strand annealing (SDSA), whereas dHJs result in hemicatenane which must be resolved by Topoisomerase III α and the BTR complex (BLM helicase bound to RMI1 and RMI2). Otherwise, dHJs can be cleaved by dHJ resolvases via resolution by the SLX4

complex (SLX4-SLX1/MUS81-EME1 and XPF-ERCC1) (Guervilly *et al.*, 2015). BTR-mediated resolution results in non-crossover products, preventing loss of heterozygosity, whereas SLX4 complex resolution in the absence of BTR increases the frequency of sister chromatid exchanges (SCEs) (Matos and West, 2014; Hustedt and Durocher, 2017).

ATM can also phosphorylate the N-terminal of p53, especially S15, which prevents the interaction of p53 with its inhibitory protein MDM2, an E3 ubiquitin ligase, thus resulting in p53 stabilisation and activation (Momand *et al.*, 1992; Kastan and Bartek, 2004; Blackford and Jackson, 2017). ATM is also responsible for phosphorylation of the kinase CHK2 at T68, which promotes homodimerisation and transactivation at Thr383/387, subsequent CDC25 inhibition, and p53 activation (Lee and Chung, 2001; Dai and Grant, 2010; Blackford and Jackson, 2017).

1.6.4 ATR-CHK1

ATR is important for responding to replication stress induced DNA damage, and is essential in proliferating cells. ATRIP recruits ATR to extended tracts of ssDNA coated with RPA, which can be generated by nucleolytic processing or helicase-polymerase uncoupling during replication (Zou and Elledge, 2003). Once ATR is recruited it requires additional activator proteins, such as TopBP1 and ETAA1 which are recruited indirectly or directly via RPA. Activated ATR phosphorylates and activates the kinase CHK1, which promotes the proteasomal degradation of CDC25A, preventing intra-S or G2-M transition and allowing time for DNA repair. Timeless and Tipin associate with RPA and are required for ATR-dependent phosphorylation of CHK1 at Ser317/345, and subsequent recruitment of Claspin. Claspin promotes BRCA1 recruitment and phosphorylation, further promoting CHK1 recruitment in a feedforward loop (Dai and Grant, 2010). CHK1 is also responsible for p53 phosphorylation and activation, further amplifying ATM-CHK2 signalling (Dai and Grant, 2010). ATR also stabilises replication forks by inhibiting cell cycle CDKs which prevents replication origin firing and exhaustion of repair factors, such as RPA (Blackford and Jackson, 2017; Toledo, Neelsen and Lukas, 2017).

Notably, γ H2AX, which can be phosphorylated by all three PIKKs, maintains chromatin as transcriptionally inactive, preventing replication-transcription collisions (Blackford and Jackson, 2017).

1.6.5 Other DNA Repair pathways

In order to cope with the many varied types of DNA damage occurring in the cell, a whole host of DNA repair pathways exist, each employing different DNA polymerases (**see Table 1-1 for a list of polymerases**) (Lindahl, 1993; Helleday, Eshtad and Nik-Zainal, 2014; Villani and Tanguy Le Gac, 2014). Base excision repair (BER) is characterised by the action of a DNA glycosylase which removes altered bases from the DNA, resulting in mutagenic apurinic or apyrimidinic sites. These sites are then excised by the lyase APEX1 and the gap filled by polymerase β , and the gap ligated by DNA ligase III-XRCC1 (Kubota *et al.*, 1996). Mismatch repair (MRR) occurs after replication and corrects miss-incorporated nucleotides. The miss-incorporated base is recognised by MSH2 and MSH6, followed by nicking and excision by PMS2 and EXO1 respectively, followed by polymerase δ gap filling (Jiricny, 2006). Nucleotide excision repair (NER) is responsible for the removal of bulky, misshapen, DNA, such as bulky adducts including UV-induced lesions. The bulky region is recognised by XPC-RAD23B or during transcription, and is excised, followed by gap filling by polymerase δ , ϵ , or κ (Aboussekhra *et al.*, 1995). Bulky lesions are usually repaired slowly by NER, therefore they often remain into DNA replication and block the high fidelity DNA polymerase. Translesion synthesis (TLS) polymerases are used to navigate across the lesion as they can accommodate bulky bases, the best characterised of these in mammals is polymerase η , which is recruited to UV-induced damaged by monoubiquitinated PCNA (Kannouche, Wing and Lehmann, 2004).

Repair Pathway	Main Eukaryotic DNA Polymerases
BER	Pol β
NER	Pol δ , ϵ , κ
MMR	Pol δ
TLS	Pol η
cNHEJ	Pol μ , λ
altNHEJ	Pol δ , η , ζ , λ
HR	Pol η

Table 1-1 Main DNA polymerases for different DNA repair pathways

1.7 DNA Replication and Genome Instability

DNA replication is carried out in S phase of the cell cycle and is particularly vulnerable to genome instability. The replisome can encounter challenges during assembly, firing, elongation and termination, resulting in replication stress, which has been proposed as one of the leading causes of sporadic cancers, tissue degeneration, and ageing. This originates from repetitive sequences, replication-transcription collisions, and CFSs, amongst others, and can manifest as replication fork slowing, fork stalling, dormant origin firing, extensive ssDNA, and ATR-CHK1/ATM-CHK2 activation (Mazouzi, Velimezi and Loizou, 2014; Saugar, Ortiz-Bazán and Tercero, 2014; Ramadan *et al.*, 2017).

Replication licensing is the process of replisome assembly prior to replication firing, which in *S. cerevisiae* occurs at specific, conserved sequences called ARS consensus sequences (ACSs), however a consensus sequence has not been identified in metazoans (Heller *et al.*, 2011). The pre-replicative complex (Pre-RC) is assembled during licensing at the end of M phase and during early G1 (Mazouzi, Velimezi and Loizou, 2014). This involves the loading of the hexameric minichromosome maintenance 2-7 (MCM2-7) complex with the help of the origin recognition complex (ORC), which contains ORC1-6 (Yekezare, Gómez-González and Diffley, 2013; Zhai *et al.*,

2017). The ORC binds to all replication origins and licenses them, along with cell division cycle 6 (CDC6), via loading of MCM2-7-chromatin licensing and DNA replication factor 1 (Cdt1) as two oppositely orientated double hexamers, followed by the release of CDC6 and CDT1. ORC, CDC6, and MCM2-7 are all AAA ATPases whose ATP binding is essential for DNA binding and MCM2-7 loading. CDT1 binding to MCM2-7 is necessary to open the MCM2-7 complex, allowing threading onto DNA. On the other hand, upon ORC and CDC6 binding, the MCM2-7 complex is in a more closed conformation (**Figure 1-2**) (Yekezare, Gómez-González and Diffley, 2013; Frigola *et al.*, 2017).

Origin firing occurs at both early and late time points throughout DNA replication in a bi-directional manner, involving the recruitment of CDC45 and the GINS complex (for Go, Ichi, Nii and San; five, one, two, three in Japanese) to form the CMG helicase complex, where MCM acts as an ATP-dependent motor, which is necessary for firing and elongation (Frigola *et al.*, 2017). Additional firing factors are needed; RecQL4, Treslin, Sld7, TopBP1 (which interacts with phosphorylated Treslin and RecQL4), and two kinases – CDK (which phosphorylates RecQL4 and Treslin) and Dbf4-dependent kinase (DDK) (which phosphorylates MCM2-7), as well as the leading strand polymerase, DNA polymerase ϵ (Yekezare, Gómez-González and Diffley, 2013; Ramadan *et al.*, 2017). The Treslin-RecQL4-TopBP1 complex enables the stable binding of CDC45 and GINS to MCM2-7 to form the CMG helicase (Mazouzi, Velimezi and Loizou, 2014). However, only a small proportion of origins that are licenced are actually fired, allowing some origins to remain dormant to safeguard against under-replication (Shima and Pederson, 2017).

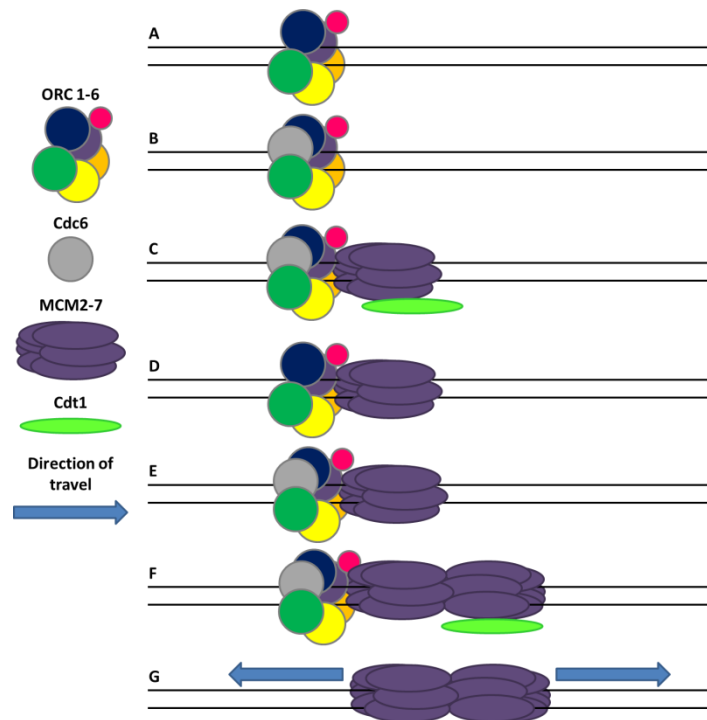


Figure 1-2 Pre-Replicative Complex formation and replication initiation

A) The Origin Recognition Complex (ORC 1-6) binds to the DNA, **B)** followed by CDC6. **C)** This ORC-CDC6 complex is responsible for recruitment of MCM2-7 and CDT1. Two MCM2-7 complexes are loaded in opposite directions. First, one MCM2-7 CDT1 complex is loaded, **D)** followed by the release of CDT1 and CDC6. **E)** Another CDC6 is then recruited, **F)** allowing the binding of the second MCM2-7 CDT1 complex. **G)** The ORC-CDC6 complex and CDT1 are then released prior to origin firing. Adapted from (Bell, 2017).

Replication elongation is carried out by the core replication machinery; proliferating cell nuclear antigen (PCNA), DNA polymerases α , δ , and ϵ , RPA, and flap endonuclease 1 (FEN1) (Balakrishnan, Gloor and Bambara, 2010; Stillman, 2015). DNA replication is always carried out in the 5'-3' direction, therefore one strand is continuously replicated (leading) and one strand as 150-200 bp fragments (lagging) referred to as Okazaki fragments, utilising different DNA polymerases. It is generally believed that DNA polymerase ϵ is required for leading strand synthesis, and a heterotetramer of RNA polymerase and DNA polymerase α (Primase) initiates lagging strand synthesis by creating an 8-12 nt RNA primer for Okazaki fragment synthesis by DNA polymerase δ (Deshmukh *et al.*, 2016). Once synthesis of one Okazaki fragment reaches the previous fragment, the synthesising strand displaces the 5' end of the fragment, producing a 5' flap. This short flap is then cleaved by FEN1, leaving a nick between the two fragments (Henneke, Koundrioukoff and Hubscher, 2003; Tsutakawa *et al.*, 2017). Following removal of the

RNA priming region, the gap is ligated by DNA ligase I (**Figure 1-3**) (Balakrishnan, Gloor and Bambara, 2010; Stillman, 2015). Unwinding of the DNA in order for replication to proceed results in the over-winding of the DNA ahead of the replication fork, resulting in high topological stress. This requires the action of Topoisomerases to be released (Pommier *et al.*, 2016).

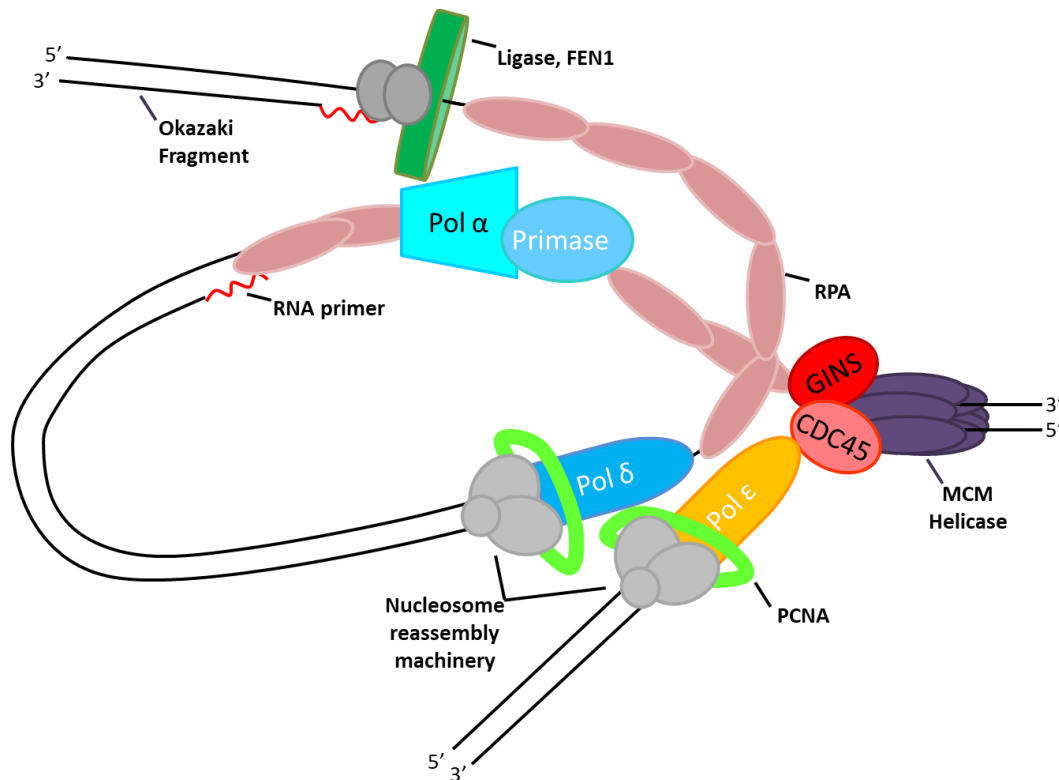


Figure 1-3 Replication elongation

Origins are fired by the action of the MCM2-7 complex, along with CDC45 and GINS to form the CMG helicase. This complex acts to open up the DNA helix, allowing passage of the core replication machinery. The leading strand is replicated by the action of polymerase ε and PCNA in a continuous 5'-3' fashion. The lagging strand is also replicated 5'-3', therefore must be carried out in segments, called Okasaki fragments. As the DNA opens up, polymerase α and an RNA polymerase are responsible for laying down RNA primers. This is then the starting point for DNA replication by polymerase δ away from the replication fork. Polymerase δ continues until it reaches the previous Okasaki fragment, where its 5' end is displaced by the polymerase. The ends are then processed by FEN1 and ligated together by DNA ligase I. Adapted from (Stillman, 2015).

In addition to carrying out unperturbed DNA replication, there are a subset of DNA polymerases, the DNA polymerase X family, which are responsible for the replication of unusual DNA structures and initiation of DNA damage repair. The DNA polymerase X family includes Pol β, Pol λ, Pol μ, and terminal deoxynucleotide transferase (TdT). Pol β is responsible for the removal of 5'-terminal abasic sugars on the 5' end of broken DNA, important for base excision repair (BER)

and DSB repair. Pol λ is required for gap filling during NHEJ, as well as accurately incorporating cytosine opposite to 8-oxoguanisine, which mimics thymidine, thus preventing transversion mutations, as well as BER, and TLS (Ramadan, Shevelev and Hübscher, 2004; Mentegari *et al.*, 2016; van Loon, Hübscher and Maga, 2017). Indeed, lack of Pol λ can result in replication stress and S phase checkpoint arrest (Zucca *et al.*, 2013). Furthermore, in addition to preventing genomic instability, Pol β can cause the expansion of repetitive sequence trinucleotide repeats in non-replicative cells in response to oxidative DNA damage via promotion of microhomology mediated end joining (MMEJ) in a FEN1 dependent manner (Crespan, Hübscher and Maga, 2015). Replication fork stalling can result in the uncoupling of DNA polymerases and the CMG helicase, resulting in long stretches of RPA coated ssDNA, which initiates monoubiquitination of PCNA at Lys164. This activates TLS in order to bypass the DNA lesion. TLS requires specialised DNA polymerases which are generally low fidelity, and includes four members of the Y-family of polymerases, Pol η , REV1, Pol κ , and Pol ι , one B-family polymerase, Pol ζ , and two members of the A-family, Pol ν , and Pol θ (Davis *et al.*, 2012; Mosbech *et al.*, 2012; Saugar, Ortiz-Bazán and Tercero, 2014). Polyubiquitination of PCNA at Lys 164 initiates the error-free DNA damage avoidance sub-pathway template switching, where the DNA lesion is bypassed by using the sister chromatid as a template to replication across the lesion (Saugar, Ortiz-Bazán and Tercero, 2014).

Finally, when two replication forks meet, DNA replication needs to be terminated. When the last region of DNA is unwound, the two converging CMG helicases meet the lagging strand from the opposite fork, and the remaining gaps are filled and ligated. Any intertwined DNA is resolved by Topoisomerase II and the CMG helicases are removed from the dsDNA (**Figure 1-4**) (Priego Moreno *et al.*, 2014; Bailey, Priego Moreno and Gambus, 2015; Dewar, Budzowska and Walter, 2015; Pommier *et al.*, 2016).

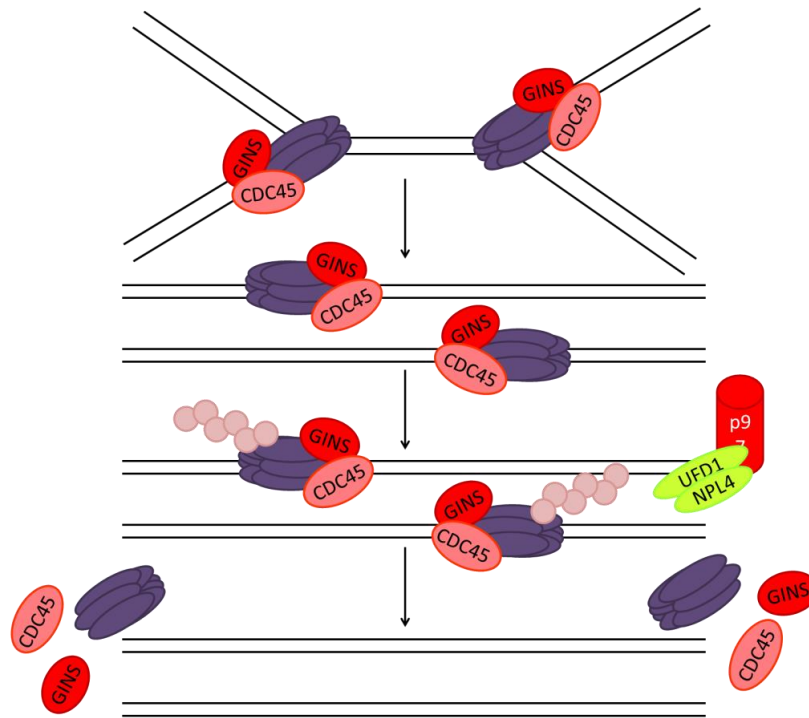


Figure 1-4 Replication termination

Replication termination occurs when two replication forks meet. The leading strand polymerase meets the first lagging strand Okazaki fragment and any gaps are filled and ligated. Any topological issues are resolved by Topoisomerases and the CMG helicases pass each other. MCM7 is then ubiquitinated and the CMG helicase is removed from the DNA by the p97-UFD1-NPL4 complex. Adapted from (Maric *et al.*, 2017).

1.7.1 Replication-Transcription Collisions

One other such process which uses DNA as a template is transcription, and conflicts between replication and transcription have been found to be a major natural source of genome instability, therefore both must be highly coordinated (Bermejo, Lai and Foiani, 2012; García-Muse and Aguilera, 2016). Replication and transcription often occur together, as both are necessary for viability and proliferation. In contrast to replication, the elongating RNA polymerase binds to dsDNA and catalyses the synthesis of the RNA chain, forming a RNA-DNA hybrid, within its active pocket. Multiple RNA polymerases can transcribe the stretch of DNA sequentially. DNA polymerase cannot pass an elongating RNA polymerase therefore conflicts can occur, whether they are in the same orientation, or more dramatically, head on (Helmrich *et al.*, 2013; García-Muse and Aguilera, 2016). Co-directional collision generally does not lead to replication pausing as it can be resolved upon transcription termination, however, head on collisions generally result in replication pausing. Mammalian cells often spatially separate

replication and transcription, particularly within highly transcribed ribosomal DNA, which takes place in the nucleoli, as well as temporally with genes which are transcribed in early S phase being replicated in late S phase. In addition to potentially direct collisions, indirect collisions can arise from transcription or replication generated features. DNA supercoiling ahead of the unwinding DNA in both transcription and replication requires Topoisomerases for release, therefore attenuation of their activity can attenuate both processes (**Figure 1-5**). Secondary structures, such as hairpins, triplex DNA, or G-quadruplexes, can form, especially at repetitive sequences (Bochman, Paeschke and Zakian, 2012; Zamiri *et al.*, 2015). These can form at ssDNA or in negatively supercoiled DNA behind the RNA polymerase and have the potential to stall replication forks. RNA-DNA hybrids (R-loops) can also arise from transcription which can lead to fork stalling. In order to counteract these collisions, RNA polymerase can backtrack to remove the last incorporated nucleotide, which allows elongation restart after pausing, thus preventing DSB formation. RNA polymerase can also be evicted from the DNA following a collision to allow progression of the DNA replication machinery (García-Muse and Aguilera, 2016). The DNA helicase RecQL5 interacts with RNA polymerase II and PCNA and has a negative role in transcription and replication elongation respectively, and can play a role in preventing RNA:DNA hybrids by promoting Topoisomerase I activity to resolve supercoiling (Yadav *et al.*, 2014; Brambati *et al.*, 2015; García-Muse and Aguilera, 2016).

Messenger RNA (mRNA) synthesis is coupled with mRNA processing, messenger ribonucleoparticle (mRNP) assembly, and nuclear export to the cytoplasm (Bermejo *et al.*, 2011). This is arbitrated by tethering of transcribed genes to the nuclear envelope via THO/TREX and TREX-2 complexes and nucleoporins (NUPs) in a process called gene gating (Gómez-González *et al.*, 2011). Nuclear pore complexes (NPCs) are made up of around 30 different NUPs arranged in an octagonal pore and a filamentous basket. NPCs span the nuclear envelope and regulate macromolecular transport, controlling most cellular processes (Ibarra and Hetzer, 2015). Gene gating can lead to topological stress building up between the elongating RNA polymerase and

the attachment site as the tethering impedes rotation, which has to be relieved by Topoisomerase I to prevent the formation of R-loops in the supercoiled region (**Figure 1-5**) (Bermejo, Lai and Foiani, 2012; Zeng *et al.*, 2012; Yadav *et al.*, 2014). However, tethering also provides topological barriers, a means by which to contain topological distortions, mediating their removal, and providing separation from replication. This separation can also be facilitated by DNA looping, which can be arbitrated by gene gating (Bermejo, Lai and Foiani, 2012). Additionally, Topoisomerase II has been implicated in relieving torsional stress from gene gating, where it is seen localising to intergenic regions during S phase (Bermejo *et al.*, 2009). RNase H is responsible for specifically degrading RNA in RNA:DNA hybrids, and is localised to R-loops by RPA binding to ssDNA, therefore is likely to be critical for preventing transcription-replication interference. Gene gating during S phase can also be controlled by the ATR checkpoint in response to ssDNA accumulation in order to promote fork stability (Bermejo *et al.*, 2011). Gene gating could allow for the transcribed mRNA to be exported through the NPCs during transcription, thus preventing it becoming entangled with the DNA and forming R-loops, however during S phase, it is presumably more favourable for replication to proceed uninhibited to avoid checkpoint activation. There is limited knowledge about how genes are re-gated following replication, however it would seem logical that tethering of the same genomic regions to NPCs could aid sister chromatid cohesion by maintaining proximity, therefore aiding in mitosis (Bermejo *et al.*, 2011; Helmrich, Ballarino and Tora, 2011; Bermejo, Lai and Foiani, 2012; Brambati *et al.*, 2015; Nguyen *et al.*, 2017).

Therefore Topoisomerases are essential for preventing torsional stress arising from both replication and transcription, and preventing replication-transcription collisions, highlighted by the fact that loss of wild type p53 sensitises cells to Topoisomerase II poisons which can be alleviated by transcription inhibition (**Figure 1-5**) (Tuduri *et al.*, 2009; Yeo *et al.*, 2016).

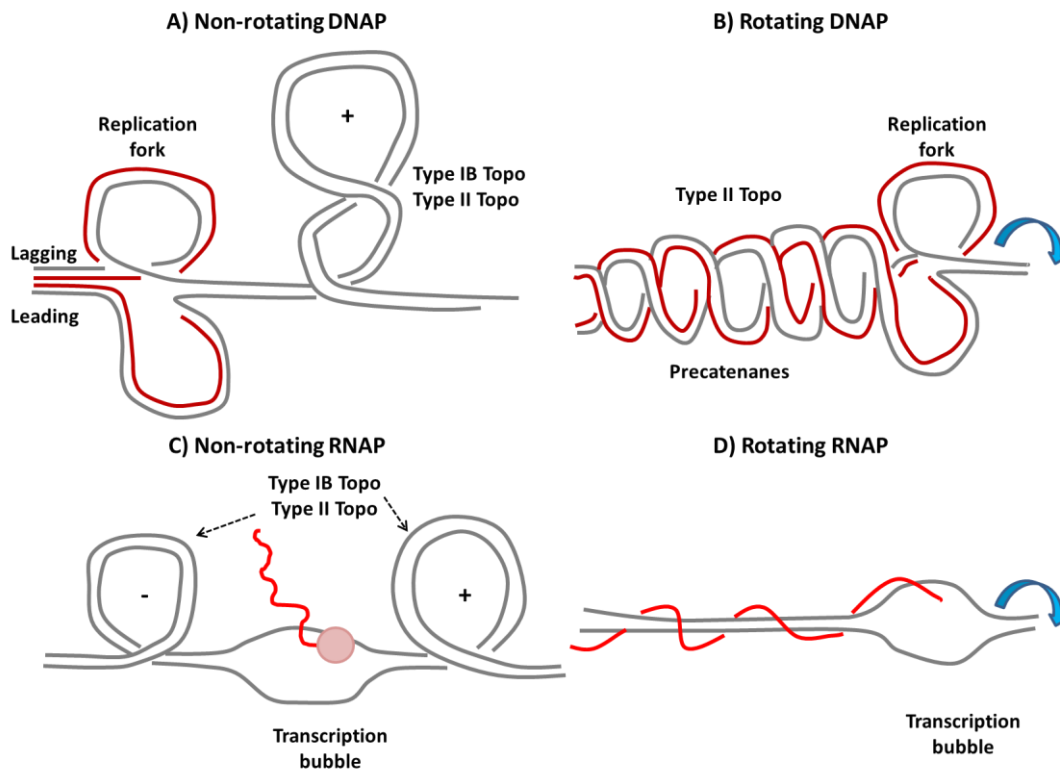


Figure 1-5 Topoisomerases are necessary to relieve the build-up of DNA torsion

Different types of DNA torsion can build up, originating from various DNA processes. A) DNA replication generates positive supercoiling ahead of the replication fork due to the opening of the DNA helix, which can be released by both Type I and Type II Topoisomerases. Generally, Topoisomerase I acts ahead of the fork. B) Unless this torsion is released, if possible, the DNA polymerase can rotate and spin the DNA helix. However, this generates precatenanes behind the fork which require Topoisomerase II for their resolution. C) Transcription caused positive supercoils in front and negative supercoils behind the transcribed region, requiring Topoisomerase mediated resolution. D) If the RNA polymerase rotates, this can release the torsion, however the newly formed RNA can become entangled with the DNA. Gene gating can be used to alleviate this entanglement. Adapted from (Bermejo, Lai and Foiani, 2012).

1.8 Topoisomerases

Topoisomerases are able to solve DNA topological problems by allowing DNA strands or double helices to pass through each other. This is achieved by sequential transesterification reactions between a nucleophilic tyrosyl oxygen of the Topoisomerase and the phosphotyrosine link of the DNA, forming a DNA gate for strand passage or rotation. Two types of Topoisomerases exist, type I and type II, and further A, B, and C subclasses (Wang 2002; Vos et al. 2011; Bermejo, Lai, and Foiani 2012).

1.8.1 Type I Topoisomerases

Type I enzymes break one strand of DNA and include all odd numbered Topoisomerases. Type IA enzymes are ssDNA strand passing enzymes, where one strand is opened and another is passed through. These include bacterial Topo IA, which relaxes negatively coiled DNA, and eukaryotic Topo III, which resolves ssDNA entanglements. With type IA enzymes, the 5' end of the cleaved DNA strand is covalently attached to the active site tyrosyl group in the 'lid' of the enzyme, whereas the 3' end is non-covalently bound to the 'base' of the enzyme. The non-cleaved strand is then passed through the gap between the lid and the base, followed by DNA religation (Wang 2002; Pommier et al. 1998). Type IB enzymes are DNA swivelases which preferentially bind negatively or positively supercoiled DNA, and work by nicking a single DNA strand and allowing one end to rotate with respect to the other. Type IB are common to all eukaryotes and are represented by a single group; Topo IB. Type IB enzymes bind the 3' end of the cleaved DNA covalently to the active site tyrosyl group, which holds the end tightly. The 5' DNA end is mostly bound ionically, allowing rotation of the DNA and multiple strand passaged per cleavage cycle (Wang 2002; Pommier et al. 1998). The only Type IC enzyme, Topo V, has been found in the archaeal genus *Methanopyrus*, and, like type IB, relaxes negative and positive supercoils via nicking and rotation, however type IC are structurally different to type IB (Wang 2002; Vos et al. 2011).

1.8.2 Topoisomerase I

Topoisomerase I (Topo IB) is a key eukaryotic enzyme heavily involved in maintaining genome stability by regulating biological processes, such as DNA replication and transcription, amongst others (Tuduri *et al.*, 2009; Yadav *et al.*, 2014; Pommier *et al.*, 2016). Topoisomerase I binds to bent or supercoiled DNA, which it then cleaves and religates by transesterification reactions. The active tyrosine (Tyr723) hydroxyl links to the 3' phosphate of the phosphodiester bond, thereby releasing the 5' hydroxyl, generating a strand break. Topoisomerase I bridges the gap by covalent and non-covalent bonds, allowing processive single strand rotation passage; the

cleaved strand rotates multiple times around the intact stationary strand (**Figure 1-6**) (Pommier *et al.*, 1998; Vos *et al.*, 2011). This provides a mechanism for removing supercoils as fast as they are generated during replication or transcription (Pommier *et al.*, 1998). Ligation is carried out by a second transesterification reaction; the exposed 5' hydroxyl attacks the phosphotyrosine linkage and re-joins the DNA (Pommier *et al.*, 1998; Vos *et al.*, 2011). Eukaryotes generally depend upon Topoisomerase I to relieve the positive supercoils generated ahead of the replication fork and transcription machinery, as well as the negative supercoils that are generated behind RNA polymerase to prevent R-loop formation (Vos *et al.*, 2011).

1.8.3 Type II Topoisomerases

Type II enzymes break both strands of the DNA and include all even numbered Topoisomerases. Type IIA enzymes cleave both strands of DNA and pass another duplex through the gap in an ATP dependent manner. This allows resolution of positive and negative supercoils, as well as detangling of chromosomes and catenanes. Type IIA can be split into three subfamilies; vertebrate Topo II including two paralogues, Topo II α and Topo II β , bacteria Topo IV, which relaxes positive supercoils, and bacterial and archaeal gyrase, which adds negative supercoils to DNA. The only type IIB enzyme, Topo VI, is found in archaea, plants and some bacteria, and works by strand passage to unlink and unwind positive and negative supercoils in an ATP-dependent manner (Wang 2002; Vos *et al.* 2011).

1.8.4 Topoisomerase II α

Topoisomerase II α is essential in proliferating cells and aids in chromosome condensation, segregation and replication (Vos *et al.*, 2011). Like Topoisomerase I, it binds DNA in a structure dependent manner, whereby binding to bent DNA is enhanced. Topoisomerase II α carries out pre-passage cleavage by making two corresponding nicks with a 4-base 5' overhang in the approximately 40 bp G-segment, or gate. A covalent bond between the active site tyrosine and 5' phosphate of the DNA is formed. Strand passage happens when Topoisomerase II α binds ATP,

causing a conformational change, resulting in double strand DNA T-segment to pass through the G-segment via a 'two-gate' mechanism. Topoisomerase II α then carries out post strand passage ligation, followed by ATP hydrolysis which opens the clamp and releases the product. The enzyme can then be reset for another round of activity, or recycled for use on a new substrate (processive vs distributive) (**Figure 1-6**) (Osheroff, Shelton and Brutlag, 1983; Burden and Osheroff, 1998; Vos *et al.*, 2011; Gao *et al.*, 2014). Topoisomerase II can substitute for Topoisomerase I ahead of the replication fork, but generally, Topoisomerase II is involved in detangling chromosomes prior to cell division and resolving catenanes arising behind and between two converging replication forks. The type IA Topoisomerase Topo III also participates in catenane disentanglement. Like Topoisomerase I, Topoisomerase II helps to resolve negative supercoiling behind transcription to prevent R-loop formation (Vos *et al.*, 2011).

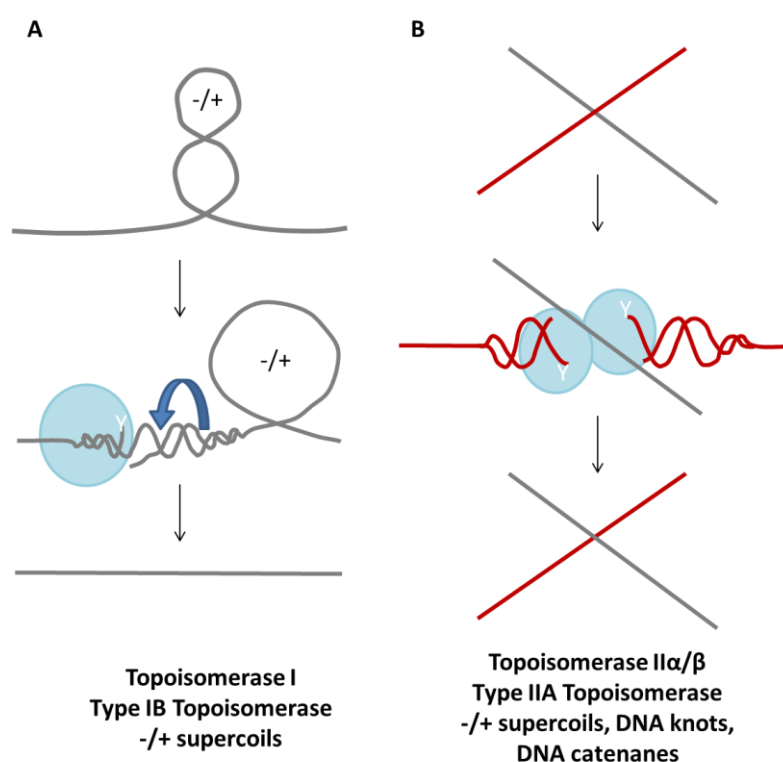


Figure 1-6 The mechanisms of action of Topoisomerase I and Topoisomerase II α/β

A) Topoisomerase I acts to remove positive and negative supercoils from DNA. To do this, it creates a single strand break in the DNA, allowing the other strand to rotate in a controlled fashion around the other strand, thereby alleviating torsion. **B)** Topoisomerase II acts to remove positive and negative supercoils, DNA knots, and catenanes. To do this, Topoisomerase II forms a homodimer, which cleaves a double stranded break in the DNA, which remains attached to the enzyme. The other strand is then passed through the gap, followed by religation of the cleaved strand. Adapted from (Pommier *et al.*, 2016).

1.8.5 Topoisomerase Inhibitors and Topoisomerase Cleavage Complexes

Topoisomerases can be trapped on the DNA in several different ways either from endogenous or exogenous sources (Yang *et al.*, 1996; Bandele and Osheroff, 2007; Koster *et al.*, 2007; Ledesma *et al.*, 2009; Zeng *et al.*, 2012; Gao *et al.*, 2014; Maede *et al.*, 2014; Vaz *et al.*, 2016; Ramadan *et al.*, 2017). Endogenous sources include dietary flavonoids (Austin *et al.*, 1992), which have been implicated in infant leukaemias (Strick *et al.*, 2000), Topoisomerase II oxidation (Hunter *et al.*, 2004), and the presence of ribonucleotides in the DNA as Topoisomerase I creates nonligatable SSBs when ribonucleotides are encountered (Wang *et al.*, 1999; Deweese and Osheroff, 2009; Yadav *et al.*, 2014). Inhibitors attenuate activity, whereas poisons stabilise Topoisomerase cleavage complexes (Topo-ccs). Camptothecin and derivatives intercalate into active site of Topoisomerase I, whereas etoposide and derivatives binds between the 5' and 3' DNA ends, both preventing religation. Other compounds act via blocking ATPase activity, or preventing DNA binding or cleavage (Koster *et al.* 2007; Maede *et al.* 2014; Vos *et al.* 2011; Pommier *et al.* 1998; Burden and Osheroff 1998; Lin *et al.* 2009). These Topo-ccs cause replication fork arrest and consequently fork collapse, DSB formation, and RNA polymerase arrest. Altogether, these lead to the exposure of Topoisomerase I SSBs, S-phase toxicity, G2 arrest, elevated p53 levels, and chromosome rearrangements (Lin *et al.* 2009; Wu and Liu 1997). As well as these drugs being most well-known for their clinical importance, especially as cancer-therapies, they can also have normal tissue toxicity, such as secondary malignancies and therapy related acute myeloblastic leukaemias (Burden and Osheroff, 1998; Pommier *et al.*, 1998; Gao *et al.*, 2014). Therefore, understanding their processing mechanisms is of great interest.

A protein inhibitor of bacterial DNA gyrase was identified in 1998 by Nakanishi *et al.* They discovered an 18 kDa protein in *E. coli*, gyrase inhibitory protein (GyrI), which is capable of directly inhibiting the supercoiling activity of DNA gyrase, and was found to be identical to the *yeeB* and *sbmC* genes. In contrast to other proteinaceous gyrase inhibitors, GyrI was the first chromosomally encoded regulatory factor which can directly inhibit DNA gyrase (Nakanishi *et*

al., 1998). In contrast, CcdB and microcin B17 are both protein inhibitors of DNA gyrase, however, they are encoded on plasmids specific to gram-negative eubacteria. Unlike other inhibitors which prevent ATP hydrolysis or alter the cleavage-religation equilibrium, GyrI acts as an antidote against other DSB inducing inhibitors potentially by either sequestering DNA gyrase, or preventing it binding to DNA (Chatterji and Nagaraja, 2002). However, GyrI has no protective effect against ATPase inhibitors which bind to the ATPase GyrB subunit, such as coumarins and cyclothialidines, but does protect against inhibitors which stabilise the GyrA subunit active centre gyrase-DNA complex, such as microcin B17, CcdB and quinolones (Chatterji, Sengupta and Nagaraja, 2003). This suggests that any role that GyrI plays in inhibition of DNA gyrase is likely carried out via the GyrA subunit.

The role of Topoisomerases in the regulation of many biological processes highlights the importance of appropriate protein maintenance on DNA.

1.9 The Ubiquitin Proteasome System

Protein homeostasis is essential for cellular metabolism and maintains the balance between protein synthesis, folding, quality control and degradation, with around a third of newly synthesised proteins being degraded within 10 minutes (D'Arcy, Wang and Linder, 2015). The ubiquitin proteasome system (UPS) is a key part of this degradation process, which in short consists of an ubiquitin tagging system and degradation by the 26S proteasome (Kravtsova-Ivantsiv and Ciechanover, 2012; Anderson *et al.*, 2015; D'Arcy, Wang and Linder, 2015). This system is responsible for the degradation of 80% of cellular proteins, and regulates many systems including protein quality control, membrane fusion, the cell cycle, transcription, DNA damage repair, DNA replication, apoptosis, and autophagy (Meyer, Bug and Bremer, 2012; D'Arcy, Wang and Linder, 2015). Disruption of the UPS has been implicated in many diseases, such as cancer, neurodegenerative disorders, cardio vascular diseases, and viral diseases (Bedford *et al.*, 2011; D'Arcy, Wang and Linder, 2015).

1.10 Ubiquitin and Ubiquitin Like Modifiers

For a protein to be targeted for degradation it is usually tagged with a chain of ubiquitin molecules. Ubiquitin is a small, highly conserved, 76 amino acid protein of approximately 8.5 kDa, which is attached to the ϵ -amino group of a substrate protein lysine via an isopeptide linkage to the carboxy glycine residue (G76) of ubiquitin. There are three fundamental classes of ubiquitination; monoubiquitination involving a single ubiquitin, multi- or poly-monoubiquitination where several ubiquitins are attached to a protein, and poly-ubiquitination where proteins are tagged with ubiquitin chains. Furthermore, the type of ubiquitin linkage can vary between seven different lysine residues (K6, K11, K27, K29, K33, K48, and K63) as well as the first methionine of Ubiquitin (M1), which determine the fate of the protein (Nalepa, Rolfe and Harper, 2006; D'Arcy, Wang and Linder, 2015; Akutsu, Dikic and Bremm, 2016; Yau and Rape, 2016). The foremost ubiquitin signal for proteasomal degradation is K48, however other non-canonical signals have been identified, including chains anchored to non-lysine residues such as threonines, serines, cysteines, or the α -amino group of the substrate protein N-terminal, and mixed chains of both ubiquitin and ubiquitin-like modifiers such as small ubiquitin-like modifier (SUMO) (Nalepa, Rolfe and Harper, 2006; Komander and Rape, 2012; Kravtsova-Ivantsiv and Ciechanover, 2012; D'Arcy, Wang and Linder, 2015). Furthermore, chains other than K48 have been implicated in proteasomal degradation, such as K11 which plays an important role in endoplasmic reticulum-associated degradation (ERAD) (Xu *et al.*, 2009). K63 has also been implicated in degradation, however its main role seems to be in regulating processes such as DNA repair, DNA replication and signalling (Saeki *et al.*, 2009; D'Arcy, Wang and Linder, 2015). On the contrary, M1-linked ubiquitin chains are synthesised in response to inflammatory signals and are recognised by linkage specific UBD domains, such as UBA1, in order to mediate immune signalling, NF- κ B activation, and interferon production (Yau and Rape, 2016).

Ubiquitination requires the activity of three enzymes, the Ub-activating E1, the Ub-conjugating E2, the Ub-ligating E3, and the Ub-chain extending E4, however the substrate set of E4 enzymes is limited (Kravtsova-Ivantsiv and Ciechanover, 2012; D'Arcy, Wang and Linder, 2015). Firstly, the activation of the ubiquitin occurs in the presence of ATP by the E1 enzyme, which is usually an ubiquitin-activating enzyme (UAE), however ubiquitin-like modifier activating enzyme 6 (UBA6) can also play this role. The E1 binds ATP and ubiquitin and catalyses their binding to form an Ub-acyl adenylate intermediate. The E1 active cysteine then attacks this bond, forming a high-energy thioester bond between the carboxy-terminal glycine of the ubiquitin and the cysteine. This is then repeated so that a second ATP and Ub form an Ub-acyl adenylate, which is also bound to the E1, creating a ternary complex. Following activation, the ubiquitin is transferred from the E1 to the catalytic cysteine of one of the 30-40 E2 Ub carriers via a transthioleation reaction. The E3 enzyme is responsible for substrate specificity, with there being over 600 in the cell, as it binds the target protein and transfers the ubiquitin to the substrate via a covalent bond. There are four families of E3 ligases, the HECT domain family which directly received the ubiquitin from the E2, forming an ubiquitin-E3 intermediate, the RING finger family which mediates a bridge between the E2 and the substrate, the U-box family, which can often function as E4 enzymes as well, however they lack the canonical cysteine yet function similarly to the RING finger family, and multi-subunit E3s containing a RING finger protein (Hatakeyama *et al.*, 2001; Komander, 2009; Bedford *et al.*, 2011; de Bie and Ciechanover, 2011; Komander and Rape, 2012; D'Arcy, Wang and Linder, 2015).

Deubiquitinating enzymes (DUBs) are responsible for the removal of ubiquitin signals in order to prevent degradation or alter signalling, with nearly 100 known DUBs being present in humans. They are also responsible for the cleavage of the C-terminal of Ub pro-proteins, thereby generating their mature forms. DUBs are split into five different families, ubiquitin C-terminal hydrolases (UCHs), ubiquitin-specific proteases (USPs), ovarian tumour proteases (OTUs), and Josephins, which are Cys proteases, and JAB1/MPN/MOV34 metalloenzymes (JAMMs) which are

zinc metalloproteases (Komander, Clague and Urbe, 2009; Bedford *et al.*, 2011; D'Arcy, Wang and Linder, 2015).

p97 is a key component of the UPS and is important for ubiquitin-dependent cellular processes, where it acts as a molecular chaperone which unfolds and segregates ubiquitinated proteins from various cellular locations including the endoplasmic reticulum, chromatin, or even protein macro-complexes, and targets them to the proteasome for degradation, or recycles them (Kloppsteck *et al.*, 2012; Meyer, Bug and Bremer, 2012; Vaz, Halder and Ramadan, 2013; Meyer and Weihl, 2014; Bodnar and Rapoport, 2017).

1.11 p97

p97/VCP is a type II AAA+ ATPase (ATPases associated with various cellular activities) found in all eukaryotes and is essential for viability. p97 forms a homo-hexamer with two stacked rings surrounding a central pore. Each p97 subunit has an N-terminal domain, two tandem AAA domains, D1 and D2, each with Walker A and Walker B motifs for ATP binding and hydrolysis respectively, a short flexible linker, and a flexible C-terminal domain. Until recently it was unclear how p97 pulled on substrates and how they were released and passed on. Bodner and Rapoport recently discerned this mechanism in yeast using the p97 homologue Cdc48. They found that when the D1 domain is bound to ADP, the N-terminal domains are in a 'down' conformation. Upon ATP binding, the D1 N-terminal domains move to an 'up' conformation, which inhibits ATPase activity. The UN complex and substrate is then able to bind. Substrate binding stimulates the ATPase activity of the D2 domain, whereby ATP hydrolysis causes a conformational change in the D2 domain, allowing the substrate to be unfolded and pulled through the central pore. The substrate is then released upon D1 ATP hydrolysis, causing the N-terminal domains to return to a 'down' conformation, allowing access of the Ub-chain to the DUB Otu1 (mammalian YOD1) (**Figure 1-7**) (Chia *et al.*, 2012; Bodnar and Rapoport, 2017).

p97 is present in both the cytosol and the nucleus and shuttling occurs between the two (Chia *et al.*, 2012; Song *et al.*, 2015). Song *et al.* confirmed that the N-terminal domain is responsible for nuclear entry, whereby mutations throughout the N-terminal domain prevent nuclear localisation, suggesting that the conformation of p97 is necessary for nuclear entry and potentially docking with nuclear import factors. Furthermore, the C-terminal sequence G₇₈₀ASPSQ is responsible for nuclear export, although it is unknown by what mechanism (Song *et al.*, 2015).

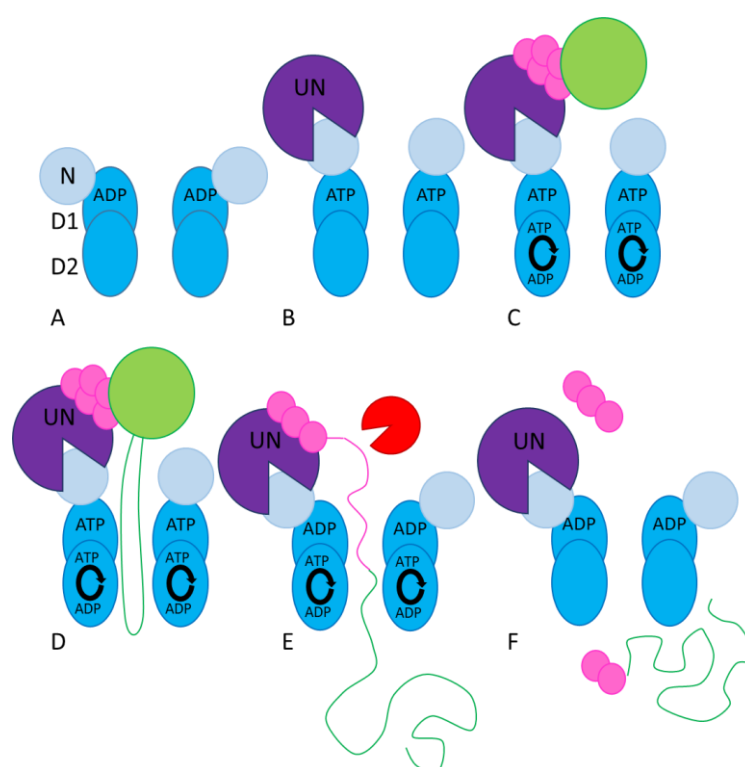


Figure 1-7 p97 mechanism of action

p97 goes through an ATP hydrolysis cycle and conformational changes during substrate processing. A) When bound to ADP, the N-terminal domains of p97 are 'down'. B) Upon binding to ATP, the N-terminal domains move 'up'. C) UN is able to bind to p97 along with C) the ubiquitinated (pink) substrate (green) which stimulates ATP hydrolysis by the D2 domain. D) This drives substrate unfolding by p97 and the substrate is passed through the central pore of p97. E) The substrate is then cleaved from the ubiquitin tag by YOD1 (red), F) allowing substrate release for degradation. Adapted from (Bodnar and Rapoport, 2017).

p97 is a highly abundant protein, with 2.7×10^6 copies per HeLa cell, and making up 1% of liver cell cytosol, highlighting its biological relevance and predicting its involvement in many different processes, much like the UPS (Yeung *et al.*, 2008; Chia *et al.*, 2012; Meyer and Wehl, 2014). The main functions of p97 are mitochondrial-associated degradation (MAD), ribosomal-associated

degradation (RAD), chromatin-associated degradation (CAD), and endoplasmic reticulum (ER)-associated degradation (ERAD). ERAD is one of the best understood roles, where p97 extracts and processes polyubiquitinated misfolded proteins and sends them to the proteasome for degradation (Anderson *et al.*, 2015; Bodnar and Rapoport, 2017).

1.12 p97 and its Adaptors

In order to participate in this multitude of pathways, p97 employs a plethora of co-factors, or adaptor proteins, which confer pathway specificity to p97 (Hänzelmann, Buchberger and Schindelin, 2011). There are at least 30 different adaptors responsible for directing p97 to specific pathways and substrates, however, the full extent of how many adaptors there are is as yet unknown (Vaz, Halder and Ramadan, 2013; Meyer and Wehl, 2014)

In order to interact with p97, adaptors have p97 interacting motifs. These include the UBX (ubiquitin regulatory X) domain which makes up the largest family of p97 adaptor proteins, of which there are 13 members in humans (Kloppsteck *et al.*, 2012). Some UBX proteins also contain a ubiquitin associated (UBA) domain for ubiquitin binding, which includes p47, FAF1, SASK1, UBXD7 and UBXD8, whereas others only have a UBX, such as p37, ASPL, UBXD1 and UBXD6. However, the UBA domain in UBXD8 is unable to bind ubiquitin, therefore is actually a UBA-like domain (Christianson *et al.*, 2012). The ubiquitin D (UBD) domain is present in NPL4 and is responsible for mutually exclusive competition with the UBX of p47 for p97 binding due to their structural homology. The PNGase (peptide N-glycosidase)/ubiquitin associated (PUB) domain is present in UBXD1, which also contains a UBX domain, and mediates binding with the C-terminal of p97, whereas most other cofactors interact with the N-terminal of p97. The SHP motif/BS1 (binding site 1) is present in p47 and UFD1, as well as several other UBX-containing proteins. Its consensus sequence, FXGXGQRU (where X is any aa and U is a non-polar residue, however serine at position two, and leucine at position eight are often also highly conserved), is always found to the N-terminal of the UBX domain suggesting spatial constraints with p97

binding. The SHP and UBX binding sites have been shown to be mutually compatible on p97, suggesting distinct binding sites. Indeed, an in detail characterisation of the p97-SHP interactions was recently performed by Hänzelmann and Schindelin (Hänzelmann and Schindelin, 2016). The VCP interacting motif (VIM) was first identified in gp78 (Ballar *et al.*, 2006), and is also present in UBXD1. Finally, the VBM (VCP-binding motif), first identified in the DUB Ataxin-3, and also present in Hrd1, has four conserved basic amino acids which are essential for p97 binding, with a consensus sequence of (L/I/V/Y)R(K/R/W)(R/K/L)RXX(Y/F)(F/K/R/Y) (Yeung *et al.*, 2008; Hänzelmann and Schindelin, 2011; Kloppsteck *et al.*, 2012; Vaz, Halder and Ramadan, 2013).

Currently it is thought that there are three main adaptor complexes which bind p97 mutually exclusively: p97^{UFD1-NPL4}, p97^{p47}, and p97^{UBXD1}. This is the starting point for a hierarchical system of adaptor binding, whereby different ‘accessory’ adaptor proteins further direct p97 specificity towards substrates and/or provide enzymatic activities, such as with the E3 ubiquitin ligase Hrd1 or with deubiquitinating enzymes such as VCIP135, Ataxin-3, and YOD1 (Hänzelmann and Schindelin, 2011; Meyer, Bug and Bremer, 2012; Vaz, Halder and Ramadan, 2013). This hierarchical binding could also change the conformation of p97 in order to facilitate binding of specific adaptors, or regulate the subcellular localisation of p97. Indeed it has been reported that both FAF1 and UBXD8 are unable to bind p97 in the absence of the UFD1-NPL4 complex (Hänzelmann, Buchberger and Schindelin, 2011).

1.12.1 p47

p47 binds to p97 as a homotrimer and is required to recruit p97 for membrane fusion and autophagy, and is one of the best studied p97 adaptors. p47 contains three domains, UBA for Ub binding, SEP for trimerisation, and UBX for interaction with p97, all connected by flexible linkers, with the SEP-UBX linker containing a SHP motif as well (Beuron *et al.*, 2006; Meyer and Wehl, 2014). The binding of p47 to p97 is bipartite in nature, due to the presence of both UBX and SHP domains, which bind to the N-terminal of p97. These two domains could either bind different N-

terminal domains, therefore all six N-terminal domains (**Figure 1-8 panel A**), or the same N-terminal domain, therefore only three out of the six available p97 monomers (**Figure 1-8 panel B**). Both of these models are so far consistent with electron microscopy studies, and may both exist depending on the ATP status of p97 in order to rearrange target proteins (Beuron *et al.*, 2006; Hänzelmann and Schindelin, 2011, 2016).

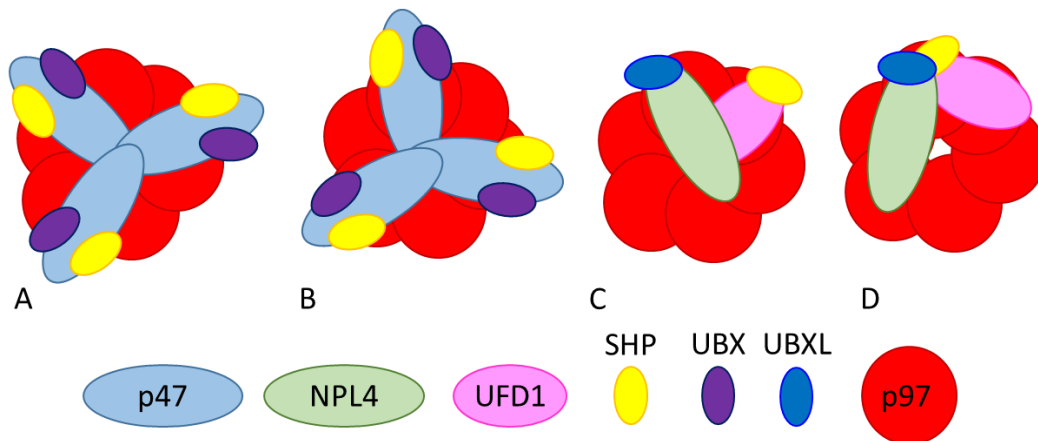


Figure 1-8 Adaptor binding to p97

p97 adaptors can bind to p97 via multiple or individual p97 binding domains. p47 has both a UBXL and a SHP domain and binds to p97 in a bipartite manner as a homotrimer and could either A) bind two different p97 monomers with each domain, or B) bind the same p97 monomer with each domain. UFD1 and NPL4 form a heterodimer and bind via a SHP domain and a UBXL domain respectively to either A) two different p97 monomers with each domain or B) the same p97 monomer with each domain. Adapted from (Hänzelmann and Schindelin, 2016).

1.12.2 UFD1-NPL4

UFD1-NPL4 (UN) is required for p97 participation in ERAD and CAD. UN binds to the p97 hexamer as a heterodimer in a ratio of 6:1:1, also using a bipartite mechanism via the UBXL of NPL4 and the SHP motif in UFD1, which provides the correct binding orientation for interaction with ubiquitinated substrates (Bebeacua *et al.*, 2016). UFD1 can also bind mono- and poly-Ub via its N-terminal domain, whereas its C-terminal domain is required for NPL4 and p97 binding. NPL4 also has a C-terminal zinc finger domain which binds Ub. It has been proposed that UFD1 and NPL4 bind different N-terminal domains of p97 as a homodimer (**Figure 1-8 panel C**), which would leave four other N-terminal domains available for subsequent adaptor binding, however it cannot be ruled out that they bind the same p97 N-terminal domain (**Figure 1-8 panel D**). Whichever is the case, both facilitate the hierarchical binding model. The bipartite binding

mechanism has been suggested to strengthen adaptor binding, thereby allowing for individually weaker interactions for greater binding specificity. As p47 and UN both bind p97 via the same domains, this creates a mutually exclusive environment. Additionally, the binding of these adaptors changes the stoichiometry of p97, thus further prevent binding of the other adaptor set (Hänzelmann and Schindelin, 2011, 2016; Bebeacua *et al.*, 2016). The choice between whether p47 or UN bind to p97 has been shown to be linked to the concentration of ATP and its binding to the D1 domain, but not its hydrolysis, with ATP binding strengthening the binding of UN. Several different binding states for UN were also observed which supports work by Bebeacua *et al.* suggesting that UN can bind p97 via either a single or double binding pattern (Chia *et al.*, 2012; Bebeacua *et al.*, 2016).

1.12.3 UBXD1

UBXD1, despite having a UBX domain, does not use this to bind p97. UBXD1 is not known to bind ubiquitinated substrates directly, but rather interacts with substrate processing components of the UPS, such as Hrd1 and HERP. UBXD1 has an uncharacterised N-terminal domain, a central PUB domain, and a C-terminal UBX domain. UBXD1 uses a bipartite mechanism using its PUB to bind the C-terminal of p97 and its VIM to bind the N-terminal at a ratio of three UBXD1 to the p97 hexamer (6:3), similarly to p47, explaining their competition due to spatial constraints. They are not mutually exclusive as such, but rather impede each other's binding. The binding of UBXD1 to both the N- and C-terminal of p97, which in itself is unique, could potentially restrict the range of motion of p97 during conformational changes, as well as preventing binding of other adaptors to its C-terminal (Hänzelmann and Schindelin, 2011; Kloppsteck *et al.*, 2012).

1.13 p97 and Protein Quality Control

Misfolded proteins can cause proteotoxic stress within the cells, therefore protein quality control mechanisms are needed to avoid and resolve this. Misfolded protein accumulate in the endoplasmic reticulum (ER), where they can form aggregates, and are cleared in a process

referred to as ER associated degradation (ERAD) (Lippincott-Schwartz *et al.*, 1988; Bonifacino, Cosson and Klausner, 1990; McCracken and Brodsky, 1996), whereby misfolded protein are degraded by the UPS (Hiller *et al.*, 1996; Christianson and Ye, 2014). Hrd1 and Doa10 are RING E3 ligases in budding yeast involved in ERAD which both sit in the ER membrane, and are evolutionarily conserved (Bays *et al.*, 2001; Swanson, Locher and Hochstrasser, 2001). Hrd1 ubiquitinates the target misfolded protein luminal or transmembrane domains, whereas Doa10 ubiquitinates misfolded cytosolic domains (Christianson and Ye, 2014). Many more E3 ligases exist in mammals, including Hrd1 (Mehnert, Sommer and Jarosch, 2010). Generally, ERAD consists of four distinct steps. Firstly, misfolded proteins are recognised by chaperones, and secondly, are relocated to the ER membrane where they are translocated across the membrane in a p97 dependent manner (Ye, Meyer and Rapoport, 2001). Thirdly, misfolded proteins are ubiquitinated by ER-membrane resident E3 ligases in the cytoplasm. Fourthly, this results in protein degradation by the proteasome (Olzmann, Kopito and Christianson, 2013). For example, misfolded proteins are recognised and relocated to Hrd1 for subsequent ubiquitination by its cofactors SEL1L, Herp, and FAM8A1, along with many other accessory factors (Schulz *et al.*, 2017). Derlin-1 is another such component which is responsible for the retrotranslocation of proteins across the ER membrane as well as coordinating their ubiquitination by bringing p97, with the help of VIMP, and ubiquitin ligases into close proximity (Ye *et al.*, 2005). p97-UN is responsible for the high energy process of retrotranslocation of most ERAD substrates through the ER membrane. It has been reported that p97 itself interacts with a portion of the misfolded protein, followed by UFD1-NPL4 interaction with the ubiquitin signal on the target protein and its subsequent removal from the membrane for degradation by the proteasome (Ye, Meyer and Rapoport, 2001, 2003).

1.14 p97 and the DNA Damage Response

Similarly to its role in ERAD, p97 is responsible for the timely removal of ubiquitinated proteins on chromatin (CAD) to prevent protein-induced chromatin stress (PICHROS). This protein turnover directly affects the DNA damage response and chromatin remodelling, and thus genome stability, therefore highlighting p97 as a genome caretaker (Vaz, Halder and Ramadan, 2013; Torrecilla, Oehler and Ramadan, 2017). In particular, p97 is responsible for the physical removal of the polycomb protein L3MBTL1 from methylated histones to allow access of 53BP1 to sites of DNA damage (Acs *et al.*, 2011), removal of the Ku70/80 complex from sites of DSB to allow proper DSB repair (van den Boom *et al.*, 2016), or removal of DNA damage sensors XPC and DDB2 to allow proper NER and genome stability (Puumalainen *et al.*, 2014). In general, it is predicted that p97 removes many as yet unidentified proteins from the sites of DNA damage as was demonstrated by hyper-accumulation of K48-ubiquitin signal at sites of DNA damage after p97 inactivation (Meerang *et al.*, 2011).

DSBs are a major platform for post-translational modifications such as phosphorylation, ubiquitination, and SUMOylation, which orchestrate the DDR. Ubiquitin modifications at the sites of DSBs are carried out by two main E3 ligases RNF8 and RNF168. K63-linked and K27-linked ubiquitin signals are mainly involved in signalling and protein recruitment, whereas K48-linked ubiquitin signals are responsible for p97 recruitment and protein removal. The main complex for p97 recruitment to DSBs is UFD1-NPL4, which, in addition to RNF8-dependent ubiquitination, can also be recruited by the action of the SUMO-targeted E3-ubiquitin ligase (STUbL) RNF4 (Torrecilla, Oehler and Ramadan, 2017).

53BP1 accumulation at DSBs requires recognition of RNF168-dependent H2AK15 K63-Ub via its non-canonical UBD, as well as H4K20me2 recognition by its tudor domain. However, H4K20me2 is masked by L3MBTL1 and JMJD2A/B, preventing the recruitment of 53BP1 due to their higher affinity. Thus, L3MBTL1 removal by p97 following RNF8- and RNF16-mediated ubiquitination is

required to enable 53BP1 recruitment to DSBs. JMJD2A/B is also ubiquitinated by RNF8, however a role for p97 in this has not yet been demonstrated (**Figure 1-9**) (Acs *et al.*, 2011; Dantuma, Acs and Luijsterburg, 2014).

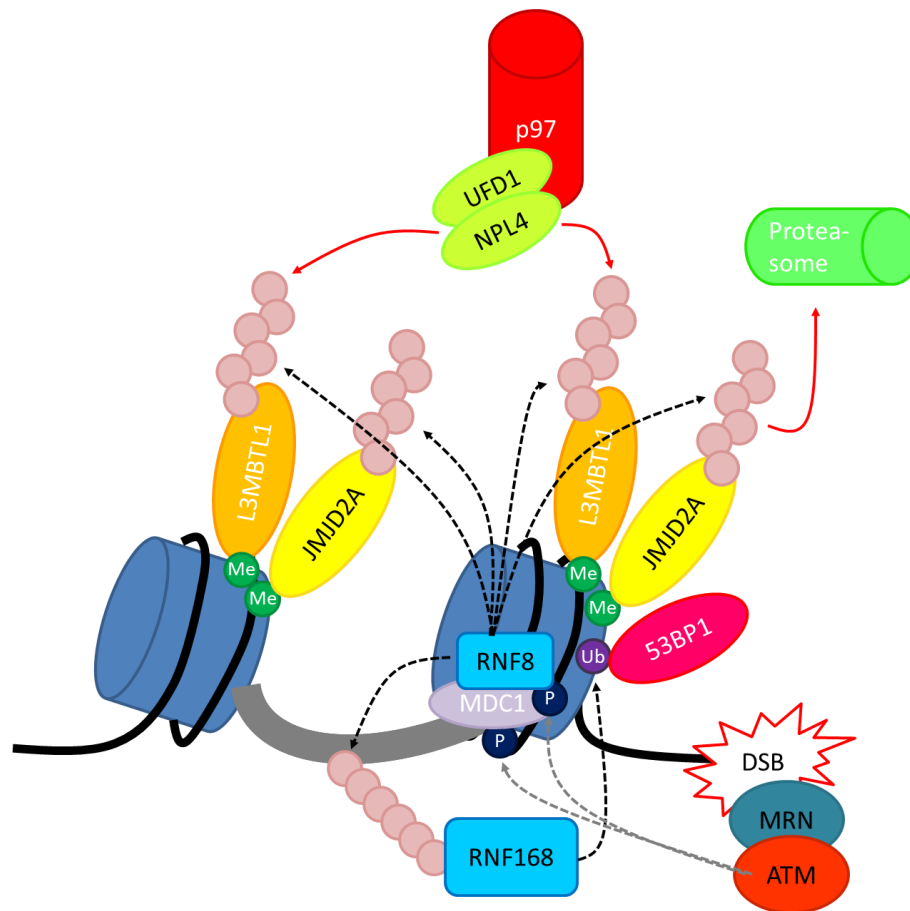


Figure 1-9 p97 is responsible for the removal of L3MBTL1 to allow 53BP1 recruitment to DSBs
For recruitment of 53BP1 to chromatin, it must bind RNF168-dependent H2AK15Ub and H4K20me2. However, H4K20me2 is masked by L3MBTL1 and JMJD2A. In order for this to be removed from, L3MBTL1 and JMJD2A must be ubiquitinated with K48-linked chains by RNF8, which is recruited to ATM-mediated phosphorylated MDC1 in response to DSBs. Following L3MBTL1 ubiquitination, it is removed from chromatin by the p97-UFD1-NPL4 molecular segregase, and directed to the proteasome for degradation. A role for p97 in the removal of JMJD2A from chromatin has not yet been established; however, it is degraded by the proteasome. This represents a well-recognised role of p97 in CAD. Adapted from (Acs *et al.*, 2011; Dantuma, Acs and Luijsterburg, 2014; Torrecilla, Oehler and Ramadan, 2017).

Ku is a chief target of p97, in conjunction with UFD1, following NHEJ and enabling HR, thus facilitating pathway choice. The Ku70/80 complex is tightly bound to DNA in a sterically trapped ring, especially following DNA end ligation during NHEJ, but also on free DNA ends. Therefore, p97 is needed to free Ku from DNA. Ku80 is ubiquitinated with K48 chains by three potential E3 ligases: SCF^{FBXL12}, RNF8, and RNF138, which carry out ubiquitination under different

circumstances. For example, RNF138 functions primarily during S/G2 of the cell cycle, therefore RNF138-dependent Ub could be responsible for p97 removal of Ku in order to facilitate HR in the presence of the sister chromatid, thus facilitating pathway choice (van den Boom *et al.*, 2016; Torrecilla, Oehler and Ramadan, 2017).

RAD51 is required during HR to coat the ssDNA and perform a homology search for the sister chromatid. Rad51 assembly is controlled in *Saccharomyces cerevisiae* by SUMOylation of Rad52, the binding partner of Rad51, following DNA damage. Rad52-SUMO binds to the SUMO interacting motif (SIM) in Rad51, promoting interaction, as well as the SIM in Ufd1, thus recruiting Cdc48/p97-Ufd1-Npl4 which disrupts the Rad51-Rad52 interaction, and hence curtailing Rad51-dependent HR. However, this Rad52-dependent Rad51 loading has yet to be established in mammals, where RAD51 loading relies on BRCA2 (Bergink *et al.*, 2013; Dantuma, Acs and Luijsterburg, 2014).

MDC1 is a crucial early player in the DDR, responsible for the recruitment of BRCA1 via RNF8-mediated histone ubiquitination, however, persistence of MDC1 at DSB prevents accumulation of BRCA1. MDC1 removal is mediated by RNF4-dependent ubiquitination (Shi *et al.*, 2008), which is counteracted by the DUB Ataxin-3 to prevent premature MDC1 removal (Pfeiffer *et al.*, 2017). MDC1 accumulates at DSBs in response to proteasome inhibitors, however p97 has not yet been implicated in its removal (Shi *et al.*, 2008; Galanty *et al.*, 2012; Vaz, Halder and Ramadan, 2013). RPA turnover is also similar to this mechanism, whereby non-SUMOylatable RPA is defective in RAD51 loading (Galanty *et al.*, 2012).

1.15 p97 and DNA Replication

During DNA replication, the replisome also has to undergo spatio-temporal regulation, which is carried out by p97. p97 has been shown to be involved in both initiation (CDT1 degradation) and termination (MCM7 degradation) of DNA replication (Franz *et al.*, 2011; Raman *et al.*, 2011; Priego Moreno *et al.*, 2014; Bailey, Priego Moreno and Gambus, 2015; Moreno and Gambus, 2015; Maric *et al.*, 2017; Ramadan *et al.*, 2017).

CDT1 is critical for the initiation of DNA replication where it 'licenses' origins. However, CDT1 is degraded upon replication initiation to prevent re-licensing of origins to ensure that the genome is only replicated once. CDT1, when bound to PCNA, is ubiquitinated by the Cullin RING-based E3 ligase (CRL) CRL4^{CDT2} upon replication initiation during the cell cycle and in response to DNA damage. p97, in union with its hierarchical adaptors UFD1-NPL4 and UBXN-3/FAF1, was identified in *Xenopus laevis* as being critical for the removal of CDT1 from chromatin, and its unfolding, for subsequent degradation by the proteasome (Franz *et al.*, 2011; Raman *et al.*, 2011; Franz, Ackermann and Hoppe, 2016).

Furthermore, the CDC-45/GINS complex persists on chromatin when CDT1 is stabilised in the absence of CDC-48/p97, suggesting that the degradation of CDT1 is important for CDC-45/GINS disassembly at the end of replication in both *Caenorhabditis elegans* and *X. laevis*. This persistence on chromatin results in replication stress and eventually genomic instability (Franz *et al.*, 2011; Franz, Ackermann and Hoppe, 2016).

CDC45 and GINS are part of the CMG complex, along with the MCM2-7 complex. Loading of this complex is required for replication initiation. The complex remains assembled throughout DNA replication, until termination when the replisome is disassembled. In *S. cerevisiae*, this disassembly is driven by dissociation of the CMG helicase which is carried out by Cdc48/p97-

UFD1-NPL4 in response to Mcm7 K29 ubiquitination by the SCF^{Dia2} complex (Maric *et al.*, 2014, 2017; Priego Moreno *et al.*, 2014).

In addition to these roles, p97, via its adaptor protein SPRTN, is responsible for the timely removal of the error-prone translesion synthesis polymerase (Pol η) once the DNA lesion has been bypassed in order to limit mutagenesis (Davis *et al.*, 2012; Mosbech *et al.*, 2012; Vaz *et al.*, 2016). In response to UV induced DNA damage, SPRTN is recruited by its PIP-box to PCNA, and by its UBZ domain to ubiquitin, thereby recruiting the p97-UN complex, via SPRTN's SHP domain, to stalled replication forks (Davis *et al.*, 2012; Mosbech *et al.*, 2012).

1.16 p97 and Topoisomerases

SPRTN exists as both mono-Ub or unmodified. In response to UV, SPRTN is not deubiquitinated, therefore remains in the mono-Ub state (Stingle *et al.*, 2016). However, SPRTN is deubiquitinated in response to the DNA-protein crosslink (DPC) inducing agent formaldehyde (FA), which allows SPRTN access to chromatin, thereby stimulating SPRTN's proteolytic activity for DPC removal (Stingle *et al.*, 2016; Vaz *et al.*, 2016). SPRTN is able to proteolyse DNA bound proteins such as histone H3, H4, H2A, H2A, H1, and of particular interest, Topoisomerase I and Topoisomerase II α cleavage complexes (Topo-ccs) (Stingle *et al.*, 2016; Vaz *et al.*, 2016).

It has been known for a while that both TDP1 and TDP2 are involved in processing Topoisomerase I and Topoisomerase II α Topo-ccs respectively (Yang *et al.*, 1996; Ledesma *et al.*, 2009; Lin *et al.*, 2009; Zeng *et al.*, 2011, 2012; Gao *et al.*, 2014). However, TDP1 and TDP2 are able to process the peptide remnants on the DNA but are unable to process the bulky Topo-cc. Furthermore, the proteasome has been implicated in this process, however the precise details are still unknown (Lin *et al.*, 2009; Gao *et al.*, 2014; Vaz *et al.*, 2016). Recently, SPRTN was identified as the missing link for the processing of Topo-ccs and is responsible for their proteolytic digestion, therefore allowing access of TDP1/TDP2 to the tyrosyl-DNA bond (Lin *et*

al., 2009; Gao *et al.*, 2014; Vaz *et al.*, 2016). As SPRTN is known for its processing of many different substrates, it begs the question how is SPRTN targeted to specific substrates?

1.17 p97 and Cell Cycle Control

As with the previously mentioned pathways, cell cycle progression also requires the turnover of cell cycle proteins for progression as well as checkpoint arrest, in which p97 plays a role (Vaz, Halder and Ramadan, 2013).

CDC25A is responsible for dephosphorylation of CDK1, necessary for the G2/M transition. In response to DNA damage, activation of ATM/ATR leads to CHK2/CHK1 activation, which subsequently phosphorylates CDC25A. This phosphorylation is recognised by the SCF^{βTrCP} ubiquitin ligase, which in turn ubiquitinates CDC25A, recruiting p97-UN to bring about CDC25A degradation, thus preventing G2/M transition. This therefore highlights another role for p97 in the DDR, this time via cell cycle control (Bartek, Lukas and Lukas, 2004; Riemer *et al.*, 2014)

Proper alignment and formation of the mitotic spindle is important to ensure correct segregation of daughter cells in asymmetrically dividing cells and tissue. This must be spatiotemporally coordinated to achieve proper orientation of the mitotic spindle. The CDC-48/p97 adaptor UBXN-2 (p37/p47) in *C. elegans* was identified as being required for the degradation of AIR-1 (Aurora A) to prevent its excessive early recruitment to centrosomes and premature maturation and mitotic spindle formation defects. Aurora A can act as a microtubule organisation centre, and therefore its excessive accumulation speeds up spindle formation, with excessive astral-spindles required for orientation, resulting in orientation defects. In human cells, p37/p47 are phosphorylated and inhibited by CDK1, therefore a mitotic increase in CDK1 activity would gradually allow the accumulation of Aurora A at centrosomes at the right time. Thus p97 is responsible for coordinating centrosome maturation with cell cycle progression (Kress *et al.*, 2013).

Aurora B is a mitotic kinase involved in mitotic processes, such as nucleus formation and chromosome segregation. During progression from prophase to metaphase, Aurora B is re-localised from the entire chromosome length to the centromere, controlling spindle attachment to the kinetochores. This helps to prevent anaphase onset and chromosome separation until all of the spindles are properly attached. Following this, Aurora B is important for spindle dynamics and cytokinesis. These roles need to be tightly controlled to ensure proper chromosome segregation. Indeed, Dobrynin *et al.* showed that, in HeLa cells, p97-UN is required for removal of non-centromeric Aurora B as cells progress towards metaphase, thereby antagonising Aurora B activity, therefore controlling proper chromosome alignment and segregation (Dobrynin *et al.*, 2011).

Following mitosis and chromosome segregation, the nuclear envelope has to be re-established, and this process requires p97. Removal of Aurora B from chromatin at the end of mitosis by CDC-48/p97-UN was shown in *X. laevis* to be necessary for chromosome decondensation and nuclear envelop formation (Ramadan *et al.*, 2007)

1.18 p97 and Disease

Reflective of the role of p97 in so many different biological functions, it is unsurprising that perturbation of p97 function underlies several diseases. Dominant mutations in *VCP*, the gene that encodes p97, causes the rare degenerative disorder inclusion body myopathy, Paget's disease of the bone and frontotemporal dementia (IBMPFD), which is the only disease directly linked to p97 so far. p97 has been linked to around 2% of amyotrophic lateral sclerosis (ALS) cases. More than 30 missense mutations have been identified which result in these diseases. N-terminal mutations all reside in the interface between the N-terminal and D1 domains, implying that impaired communication between the two is causative of the diseases. Generally, oligomerisation is not affected, however ATP hydrolysis by the D2 domain is often impaired, but not in all cases. The source of disease pathogenesis appears to be via specific cofactor

association, explaining why, unlike complete loss of p97, these mutations are not embryonic lethal. Mutations in p97 autophagy adaptors responsible for degradation of aggregated proteins in inclusion bodies and autosomal degradation, such as p62/SQSTM1 and optineurin, are also associated with frontotemporal dementia (FTD), ALS and Paget's disease of the bone (PDB). Mutations affecting endolysosomal degradation, affecting multivesicular body (MVB) biogenesis have also been implicated in ALD and FTD. The role of p97 in mitochondria degradation has also been linked to Parkinsonism, where p97 extracts proteins from the outer mitochondrial membrane which have been ubiquitinated by the E3 ligase Parkin (Meyer and Weihl, 2014). p97 has also been implicated in dealing with protein aggregates which can be linked to aging, such as neurodegenerative diseases arising from the aggregation of proteins with expanded poly-glutamine stretches (PolyQ). There are 10 different polyQ diseases linked to specific protein aggregation, including Huntington's disease (HD) and Machado-Joseph disease (MJD). p97 co-localises with aggregated proteins involved in both of these diseases. Furthermore, p97 associates with ubiquitinated inclusions in Alzheimer's disease (AD) and prion diseases (Franz, Ackermann and Hoppe, 2014).

Many of the pathways employed by cancerous cells include those regulated in some way by p97. In fact, p97 has been implicated in breast, colorectal, lung, prostate and pancreatic cancers (Vaz, Halder and Ramadan, 2013). Cancerous cells often rely heavily on protein quality control due to their increased level of protein synthesis and aberrant proteins, often showing increased levels of p97 which correlates with progression, poor prognosis, and metastasis. Alterations in p97 adaptor functions, such as FAF1, have been observed in 30% of cervical carcinomas and 12.5% of mantle-cell lymphomas, as well as being downregulated in gastric carcinomas and many mesotheliomas (Vaz, Halder and Ramadan, 2013). Furthermore, biallelic germline mutations in the p97 adaptor SPRTN were identified to be the underlying cause of a new segmental progeroid syndrome represented by genomic instability and susceptibility towards hepatocellular carcinoma (Lessel *et al.*, 2014). Previously to this discovery, SPRTN was implicated in TLS (Davis

et al., 2012; Mosbech *et al.*, 2012) and recently has also been implicated in DNA-protein crosslink repair (Stingle *et al.*, 2016; Vaz *et al.*, 2016), revealing diverse mechanisms by which SPRTN mutation may be contributing towards aging and cancer. Exact mechanisms of p97 involvement in oncogenesis are gradually being elucidated. Thus considering the plethora of pathways discussed earlier it is not difficult to see how p97s role in maintaining genome stability could be involved in cancer development and progression, and how it could be targeted for effective therapies (Vaz, Halder and Ramadan, 2013; Franz, Ackermann and Hoppe, 2014).

Proteasome inhibitors have been developed to tackle some of these diseases, with the first success being with Bortezomib for multiple myeloma. The success of this inhibitor plays on the fact that cancer cells generate more aberrant proteins and higher levels of oncoproteins, therefore are more sensitive to proteasome inhibition (Adams *et al.*, 1999; Adams, 2004; Nalepa, Rolfe and Harper, 2006).

The first p97 inhibitor was reported by showing the prevention of turnover of ERAD substrates (Fiebigler *et al.* 2004; Wang, Li, and Ye 2008) by being able to bind to the D1 domain of p97 and the ER cell membrane (Wang *et al.* 2010), and had similar effects to that of Bortezomib with which it can synergise as an anticancer agent (Wang *et al.* 2009). Several different p97 inhibitors have been developed since, such as DBE (Chou *et al.*, 2011), NMS-873 (Magnaghi *et al.*, 2013), and CB-5083 (Anderson *et al.*, 2015; Zhou *et al.*, 2015). However, CB-5083 is the only inhibitor so far which displays the pharmacological properties necessary for clinical trials. The principal cause of cell death upon treatment with CB-5083 seems to be via unavoidable proteotoxic stress, however, at later time points cells seem to die due to signalling not linked to the action of CB-5083, highlighting the numerous roles of p97 within the cell (Anderson *et al.*, 2015). Therefore, it could be prudent to target specific functions of p97 in order to have fewer off target effects, as well as to provide a two pronged targeting strategy to try and limit resistance.

1.19 Aims of this Thesis

A better understanding of the role of p97 in response to DNA damage, in particular to IR induced DSBs, could provide druggable targets with greater pathway specificity in order to develop therapies with fewer off target toxicities. Therefore, in order to gain further insights into the p97 proteome, a Mass Spectrometry (MS) analysis of p97 interacting factors was performed in response to IR. Both early and late responses were analysed in order to gain an understanding of protein dynamics in the cell. From this, a novel protein with intriguing dynamics was identified, which here will be referred to as Ionising Radiation Enhanced VCP Interacting Protein 1, or IREVIP1. Due to the nature of the predicted domains in IREVIP1, and its dynamic interaction with p97 following IR, it was believed to be an interesting target for further investigation. Thus, this thesis aims to characterise IREVIP1 in regards to p97 and its involvement in maintaining genome stability.

2. MATERIALS AND METHODS

All reagents were purchased from Sigma, unless stated otherwise.

2.1 Buffers

2.1.1 10x Phosphate Buffer Saline (PBS), 1 L

The following were dissolved in 800 mL ddH₂O:

- 18.1 g Na₂HPO₄•2H₂O
- 80 g NaCl
- 2 g KCl
- 2.4 g KH₂PO₄

The pH was adjusted to 7.4 with HCl and the volume then adjusted to 1 L with ddH₂O.

The 10x stock of PBS was diluted 1:10 for use as a 1x working solution (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4).

2.1.2 10x Tris-Buffered Saline (TBS), 1 L

The following were dissolved in 900 mL ddH₂O:

- 88 g NaCl
- 24 g Tris Base

The pH was adjusted to 7.6 with HCl and the volume then adjusted to 1 L with ddH₂O.

The 10x stock of TBS was diluted 1:10 for use as a 1x working solution (20 mM Tris, 150 mM NaCl).

For use as 1x TBS-Tween, 1 mL of Tween-20 was added to 1 L of 1x TBS for a final concentration of 0.1%.

2.1.3 10x Tris/Glycine Buffer, 1 L

The following were dissolved in ddH₂O:

- 30 g Tris

- 144 g Glycine

The 10x stock of Tris/Glycine was diluted 1:10 for use as a 1x working solution for running and transfer buffer for Western blots.

For running buffer, SDS was added to 1x Tris/Glycine to a final concentration of 0.1%.

For transfer buffer, 10x Tris/Glycine was diluted with 20% methanol and 0.025% SDS in ddH₂O.

2.1.4 5x Laemmli Buffer (LB)

- Bromophenol Blue (SLS, 0.25%)
- Glycerol (Fisher Scientific, 50%)
- SDS (10%)
- Tris-HCl (0.25 M, pH 6.8)

5x Laemmli was used at a 1x working concentration (50 mM Tris-HCl pH 6.8, 2% SDS, 10% Glycerol, 0.05% bromophenol blue), supplemented with 15% β -mercaptoethanol (β Me).

2.2 Plasmid Manipulation and Cloning

2.2.1 Sub-cloning

The insert of interest was amplified from a source plasmid (See **Table 2-3**, for uniprot numbers see **Table 2-4**) with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) using the desired primers and cycling conditions (See **Table 2-1**, **Table 2-2** and **Table 2-5**). The PCR product was separated from the template plasmid on a TBE agarose gel, the band of the correct size excised, and purified using a QIAquick Gel Extraction Kit (Qiagen). The purified product was digested at 37°C using the desired restriction enzymes (New England Biolabs), and cleaned up using a QIAquick PCR Purification Kit (Qiagen). The desired backbone (See **Table 2-3**) was digested using the desired restriction enzymes at 37°C and purified using the QIAquick Gel Extraction Kit. The purified, digested, insert was ligated with the digested backbone T4 DNA Ligase (ThermoFisher Scientific) according to the manufacturer's instructions. The ligated products were transformed into Sub-cloning Efficiency™ DH5 α ™ Competent Cells (ThermoFisher

Scientific) (See Bacterial Transformations). Colonies were analysed by colony PCR using Taq DNA Polymerase (ThermoFisher Scientific) and gel electrophoresis. Colonies positive by colony PCR were grown in culture and the plasmids extracted using NucleoSpin® Plasmid (Macherey-Nagel) according to the manufacturer's instructions. The plasmids extracted from each clone were sent for sequencing (Source Bioscience). One correctly sequenced clone was then chosen for storage and future use. If required for protein purification, the plasmid was transformed into competent Rosetta cells (Grigory Dianov Laboratory) and the screening process repeated.

Component	50 μ L	Final Concentration
H ₂ O	To 50 μ L	
5X Phusion HF Buffer	10 μ L	1X
10 mM dNTPs	1 μ L	200 μ M each
Fw Primer	Varies	0.5 μ M
Rv Primer	Varies	0.5 μ M
Template DNA	Varies	1 pg – 10 ng
Phusion DNA Polymerase	0.5 μ L	0.02 U/ μ L

Table 2-1 PCR reaction setup for Phusion DNA polymerase for sub-cloning

Temperature	Step	Time
95°C	Initial Denaturing	30 s
95°C	Denaturing	10 s
See Table 2-5 List of primers	Annealing	30 s
72°C	Extension	30 s/kb
72°C	Final Extension	10 minutes
4°C	Hold	∞

Table 2-2 Cycling conditions for Phusion DNA polymerase for sub-cloning

Plasmid	Source
pOTB7 IREVIP1 cDNA Clone IRAUp969H0427D	Source Bioscience
pGEX-4t-2	Lab Stock
pGEX-4t-2 IREVIP1 nGST	Sub-Cloning
pET21a(+) cHis	Ivan Ahel, Sir William Dunn School of Pathology
pCDNA5 FRT/TO FlpIn cSSH	ThermoFisher Scientific
pET21a(+) IREVIP1 NTΔ cHis	Sub-Cloning
pET21a(+) IREVIP1 NTΔ SHPΔ cHis	Mutagenesis
pET21a(+) IREVIP1 NTΔ VBM* cHis	Mutagenesis
pET21a(+) IREVIP1 NTΔ SHPΔ VBM* cHis	Mutagenesis
pCDNA3.1 p97 cS•Tag	John Christianson, Oxford Ludwig Institute
pET21a(+) p97 cS•Tag	Sub-Cloning
pCDNA5 FRT/TO FlpIn IREVIP1 cSSH	Sub-Cloning
pCDNA5 FRT/TO FlpIn IREVIP1 NTΔ cSSH	Sub-Cloning

Table 2-3 List of plasmids

Protein	Uniprot number	NCBI Reference
IREVIP1 Human	Q9Y6I9	
IREVIP1 Chicken (Gallus gallus)	E1C8H2	
IREVIP1 Mouse (Mus musculus)	Q99LS5	
IREVIP1 Rat (Rattus norvegicus)	Q5XIF0	
IREVIP1 Chimp (Pan troglodytes)	K7CBB3	
IREVIP Pig (Sus scrofa)	I3L9L8	
IREVIP1 Dog (Canis lupus familiaris)	E2RAC8	
IREVIP1 Xenopus tropicalis		XP_002935663.2
IREVIP1 Zebrafish (Danio rerio)		XP_001335278.2
Gyrl	P33012	
p97/VCP	P55072	

Table 2-4 List of Uniprot and NCBI Reference numbers

Primer	Backbone	Sequence	Annealing Temperature
IREVIP1nGST Fw EcoRI	for pGEX-4t-2	TTTACAGCGGCCGCTTACTCCTTGCCCTTCTCAG GGGC	55°C
IREVIP1nGST Rv NotI	for pGEX-4t-2	AAAAGAATTCCGATGCGACCTTGGGGCAGC	55°C
IREVIP1cHis Fw BamHI	for pCDNA5	ATATGGATCCATGTCGGACCTGCTACTACTGG	55°C
IREVIP1cHis Rv NotI		AATTGCGGCCGCCTCCTTGCCCTTCTCAGGGG	
IREVIP1 NTΔ Fw NdeI	for pET21a(+)	ATATCATATGGGGTACTCAGGGCTACTGG	55°C
IREVIP1 Rv NotI		AATTGCGGCCGCCTCCTTGCCCTTCTCAGGGG	
VCPcS•Tag Nde1 Fw	for pET21a(+)	ATAT CATATGGCCTCTGGAGCCGATTC	67°C
VCPcS•Tag Not1 Rv WITH STOP		ATATGCGGCCGCTTAGCTGTCCATGTGCTGGC	
IREVIP1 SHP deletion 272-285 Fw	for IREVIP1 NTΔ pET21a(+)	CAGCGGCTCCTCTGAGTCACGGCTGG	55°C
IREVIP1 SHP deletion 272-285 Rv		CCAGCCGTGACTCAGAGGAGCCGCTG	
IREVIP1 VBM mutagenesis 1 and 3 Fw round 1	for IREVIP1 NTΔ pET21a(+) or IREVIP1 NTΔ SHPΔ pET21a(+)	GCCTTGGACACCTACATCGCGGAG GCGAAGCTGTGTGCCTATCC	65°C
IREVIP1 VBM mutagenesis 1 and 3 Rv round 1		GGATAGGCACACAGCTCGCCTCCGC GATGTAGGTGTCCAAGGC	
IREVIP1 VBM mutagenesis 2 and 4 Fw round 2	for IREVIP1 NTΔ VBM Mt1 pET21a(+) or IREVIP1 NTΔ SHPΔ VBM Mt1 pET21a(+)	CACCTACATCGCGGCGGCGGCG CTGTGTGCCTATC	64°C
IREVIP1 VBM mutagenesis 2 and 4 Rv round 2		GATAGGCACACAGCGCCGCCGCCGCGATGTAG GTG	

Table 2-5 List of primers

2.2.2 Mutagenesis

Plasmid mutagenesis was carried out for the desired backbone (See Table 2-3) and primers (See Table 2-5) using AccuPrime™ Pfx DNA Polymerase (ThermoFisher Scientific) according to the reaction parameters below (See **Table 2-6** and **Table 2-7**).

Component	50 μ L	Final Concentration
H ₂ O	To 50 μ L	
10X AccuPrime™ Pfx Reaction mix with dNTPs	5 μ L	1X
Fw Primer	Varies	10 pmol
Rv Primer	Varies	10 pmol
Template DNA	Varies	300 ng
Phusion DNA Polymerase	0.5 μ L	0.05 U/ μ L

Table 2-6 PCR reaction setup for Phusion DNA polymerase for mutagenesis

Temperature	Step	Time
95°C	Initial Denaturing	2 minutes
95°C	Denaturing	15 s
See Table 2-5 List of primers	Annealing	30 s
72°C	Extension	1 m/kb
4°C	Hold	∞

Table 2-7 Cycling conditions for Phusion DNA polymerase for mutagenesis

The mutagenesis PCR reaction was digested with DpnI (ThermoFisher Scientific) for one hour at 37°C. Following which, the undigested product was transformed into DH5 α (ThermoFisher Scientific) and plated onto LB agar plates with the appropriate antibiotics overnight at 37°C. Single colonies were then picked and grown in 5 mL of LB media overnight shaking at 250 rpm at 37°C. The plasmids were purified using NucleoSpin® Plasmid (Macherey-Nagel) and the mutation analysed by sequencing.

2.3 Bacterial Transformations

Bacteria were defrosted on ice before addition of plasmid to no more than 10% of the volume of bacteria. The plasmid was incubated with the bacteria on ice for 30 minutes. The bacteria were heat shocked by incubating at 42°C in a water bath for 1 minute 30 seconds, followed by two minutes on ice. Bacteria were left to recover, shaking at 250 rpm at 37°C, in 1 mL of LB media. Bacteria were then spun at 2,800 xg for 3 minutes, the supernatant was mostly removed leaving around 30 µL in which the bacteria were resuspended and spread on a LB agar plate with the required antibiotics. Bacteria were left to grow over night in an incubator at 37°C.

2.4 Native Purification

Rosetta *E. coli* containing the desired construct were grown at 37°C until they reached an OD₆₀₀ of 0.6, at which point protein production was induced by the addition of IPTG to a final concentration of 1 mM and were grown at the indicated temperature for the indicated amount of time (See **Table 2-8**). Cells were collected by centrifugation at 5000 xg for 30 minutes before storing at -80°C overnight prior to lysis.

Construct	Growth Temperature	Growth Time	Beads	Elution
pET21a IREVIP1 cHis NTA	21°C	6 h	Ni-NTA Agarose (Qiagen)	200 mM – 1 M Imidazole
pET21a IREVIP1 cHis NTA SHPΔ	37°C	6 h	Ni-NTA Agarose (Qiagen)	200 mM – 1 M Imidazole
pET21a IREVIP1 cHis NTA VBM*	21°C	6 h	Ni-NTA Agarose (Qiagen)	200 mM – 1 M Imidazole
pET21a IREVIP1 cHis NTA SHPΔ VBM*	21°C	6 h	Ni-NTA Agarose (Qiagen)	200 mM – 1 M Imidazole
pET21a p97 cS•Tag	37°C	6 h	S-Protein Agarose (Merck Millipore)	200 mM Citric Acid pH 2.0

Table 2-8 Purification constructs, growth temperatures and times, beads used, and elution conditions

Cell pellets were resuspended in 1x native resuspension buffer (800 mM NaCl, 20 mM Tris HCl pH 7.0) containing 1% Triton X-100 and 1 mM PMSF, prior to the addition of lysozyme and incubation on ice for 30 minutes. Lysates were sonicating for two minutes at 20% power, 50% pulsing, using an Ultrasonic Homogeniser Model 300V/T (Biologics Inc.), followed by centrifugation at 20,000 xg for 30 minutes. Cell lysates were incubated with beads pre-washed in native resuspension buffer for two hours at 4°C rolling at 15 rpm. The lysate was then spun at 500 xg for 5 minutes and the supernatant collected as 'flow through'. The beads were washed, rotating at 15 rpm at 4°C for 15 minutes, twice with native resuspension buffer, followed by one wash with native resuspension buffer including 50 mM imidazole, and one wash with native resuspension buffer including 100 mM imidazole. Proteins were eluted from the beads by incubating with the desired elution buffer (See **Table 2-8**), rotating at 15 rpm at 4°C for 15 minutes. Proteins were analysed by SDS-PAGE and coomassie staining (Instant blue, Expedeon). The elution fractions containing the protein of interest were combined and the buffer exchanged three times with storage buffer (150 mM NaCl, 10 mM Tris HCl pH 7.5, 0.1 mM EDTA, 15% Glycerol), using EMD Millipore Amicon™ Ultra-15 Centrifugal Filter Units (Fisher Scientific), and concentrated to the desired final concentration.

2.5 Denaturing Purification

Rosetta *E. coli* containing pET21a IREVIP1 cHis NTA were grown, induced, and collected as described for native purification (See chapter 2.4). Cell pellets were resuspended in 1x denaturing resuspension buffer (800 mM NaCl, 20 mM Tris HCl pH 7.0, 8 M Urea) containing 1% Triton X-100 and PMSF prior, to the addition of lysozyme and incubation at RT for 30 minutes. Lysates were processed as described for native purifications. Cell lysates were incubated with Ni-NTA agarose beads (Qiagen) pre-washed in denaturing resuspension buffer for two hours at RT rolling at 15 rpm. The lysate was spun at 500 xg for 5 minutes and the supernatant collected as 'flow through'. The beads were washed, rotating at 15 rpm at RT, twice with 1x denaturing

resuspension buffer, once with 1x denaturing resuspension buffer with 50 mM imidazole, and once with 1x denaturing resuspension buffer with 100 mM imidazole.

Whilst still coupled to the beads, the proteins were refolded by three washes, rolling at 15 rpm for 15 minutes each at RT, in 1x denaturing resuspension buffer without Triton X-100, containing a decreasing concentration of Urea (6 M, 4 M, 2 M, 0 M). The washes in 2 M Urea and below were carried out at 4°C.

Proteins were eluted from the beads by rotating at 15 rpm at 4°C in increasing amount of imidazole (200 mM – 1 M) in 1x native resuspension buffer without Triton X-100. Proteins were analysed by SDS-PAGE and coomassie staining. The fractions of interest were pooled and the buffer exchanged three times with storage buffer (150 mM NaCl, 10 mM Tris HCl pH 7.5, 0.1 mM EDTA, 15% Glycerol), using EMD Millipore Amicon™ Ultra-15 Centrifugal Filter Units (Fisher Scientific), and concentrated to the desired final concentration.

2.6 ÄKTA system purification

The ÄKTA FPLC system was operated according to the manufacturer's instructions. HisTRAP FF Crude 5 mL affinity purification, Superdex 75 and Superdex 200 size exclusion, and MonoQ 1 mL ion exchange columns were also used according to the manufacturer's instructions. The buffers used are as listed below (See **Table 2-9**, see Native Purification for affinity purification buffers).

Column	Buffer A	Buffer B
Gel filtration	20 mM Tris HCl pH 7.0, KCl 150 mM, EDTA, 41 mM, Glycerol 5%	N/A
Ion Exchange	20 mM Tris HCl pH 7.0, KCl 50 mM, EDTA 1 mM, Glycerol 5%	20 mM Tris HCl pH 7.0, KCl 1 M, EDTA 1 mM, Glycerol 5%

Table 2-9 Buffers used during ÄKTA purification

2.7 In vitro interactions

Beads were washed three times using 1x binding buffer 1 before coupling with the desired first protein for two hours at 4°C (See **Table 2-10**). The beads were washed three times using 1x binding buffer 2 before adding the second protein of interest for two hours at 4°C (See **Table 2-10**). The beads were washed three times in 1x binding buffer 2 before eluting using 2x LB with 15% β Me and boiling at 95°C for five minutes. Samples were analysed directly by Western blotting.

Beads	Binding Buffer 1	First Protein	Binding Buffer 2	Second Protein
HisPur™ Ni-NTA Magnetic Beads (Thermo Scientific)	137 mM NaCl 20 mM Tris HCl pH 7.0	IREVIP1 cHis NTΔ Homemade 500 ng	137 mM NaCl 20 mM Tris HCl pH 7.0	Topoisomerase I (Inspiralis) 50 ng
HisPur™ Ni-NTA Magnetic Beads (Thermo Scientific)	137 mM NaCl 20 mM Tris HCl pH 7.0	IREVIP1 cHis NTΔ Homemade 500 ng	137 mM NaCl 20 mM Tris HCl pH 7.0	Topoisomerase IIα (Inspiralis) 267.75 ng
S-Protein Agarose (Merck Millipore) Blocking, 5% BSA overnight	137 mM NaCl 20 mM Tris HCl pH 7.0	p97 S•Tag Homemade 5 µg	137 mM NaCl 20 mM Tris HCl pH 7.0 0.5% Triton X-100	IREVIP1 cHis NTΔ/ SHPΔ/VBM*/ SHPΔ VBM* 500 ng
S-Protein Agarose (Merck Millipore) Blocking, 5% BSA overnight	137 mM NaCl 20 mM Tris HCl pH 7.0	p97 S•Tag Homemade 5 µg	137 mM NaCl 20 mM Tris HCl pH 7.0 0.5% Triton X-100	Topoisomerase I (Inspiralis) 50 ng
S-Protein Agarose (Merck Millipore) Blocking, 5% BSA overnight	137 mM NaCl 20 mM Tris HCl pH 7.0	p97 S•Tag Homemade 5 µg	137 mM NaCl 20 mM Tris HCl pH 7.0 0.5% Triton X-100	Topoisomerase IIα (Inspiralis) 267.75 ng

Table 2-10 Interaction conditions for in vitro protein interactions

2.8 Mammalian Cell Culture

2.8.1 Human Cell Lines

The human cervix adenocarcinoma (HeLa), human embryonic kidney (HEK293) and human osteosarcoma (U2OS) cell lines were purchased from ATCC (American Type Culture Collection).

The retinal pigmented epithelial cells immortalised with hTERT (RPE1) were obtained from the lab of Prof. Esther Hammond, CRUK/MRC Oxford Institute of Radiation Oncology.

The normal human foetal lung fibroblasts (TIG-1) were obtained from the lab of Prof. Grigory Dianov, CRUK/MRC Oxford Institute of Radiation Oncology.

2.8.2 Generation of CRISPR/cas9 Knockout Cells

CRISPR knockout cell lines were produced in our laboratory by J. Fielden. Two plasmids, a CRISPR/Cas9 KO plasmid containing 3 guide RNAs targeting *IREVIP1*, and a homology directed repair plasmid containing a puromycin resistance cassette, were purchased from Santa Cruz. 2.5 µg of each plasmid was transfected into early-passage HEK292 or HeLa cells using Eugene HD (Promega). After 72 hours, media supplemented with puromycin was added to the cells. The puromycin dose required to kill wild-type cells was determined by puromycin kill curve (1.25 µg/ml for HEK293 cells, 0.6 µg/ml for HeLa cells). After 72 hours, the puromycin-containing media was removed and cells were sorted using a cell sorter into single-cell populations on a 96-well plate. IREVIP1 expression was analysed by Western blot. Three HEK293 and three HeLa clones showing loss of all detectable IREVIP1 were selected for subsequent analysis.

2.8.3 Maintenance of Human Cell Lines

U2OS, HeLa, HEK293, TIG-1, RPE1 cells, and CRISPR knockout cell lines of the corresponding cell type, were cultured in Dulbecco's Modified Eagle's (DMEM, Sigma) High Glucose Medium, supplemented with 15% (TIG-1) or 10% (HeLa, U2OS, HEK293, RPE1) Foetal Bovine Serum (FBS, Sigma), and 1x Penicillin/Streptomycin (Pen/Strep, Gibco) (DMEM complete media). Cells were kept at 37°C, 5% CO₂, in a humid incubator. Cells were passaged upon reaching a confluency of

80-90%. Cells were washed with 1x PBS, trypsinised with 1x trypsin (0.05%), and re-seeded at a density of 1:10 (HeLa, U2OS, HEK293, RPE1, and TIG-1). Cells were used until passage 25 (HeLa, U2OS, HEK293, and RPE1) or 20 (TIG-1).

To thaw cells from liquid nitrogen the vials were thawed in a 37°C water bath. Cells were then diluted with 9 mL complete media. The cells were left to attach to the plate for 24 h, after which the media was changed. Cells were then grown for two passages before use.

To freeze cells, cells were trypsinised, centrifuged at 500 xg, washed once with 1x PBS, centrifuged at 500 xg, and resuspended in freezing media (45% DMEM, 45% FBS, 10% DMSO (Sigma)). Aliquots of 1 mL were frozen down at -80°C for 24 h in a freezing box (CoolCell) before transferring to liquid nitrogen.

2.8.4 Stable Cell Line Production

Flp-In™ T-REx™-293 cells (ThermoFisher Scientific) were grown to a confluency of 30-40% in DMEM complete media containing Zeocin (100 µg/mL, ThermoFisher Scientific) and Blasticidin (15 µg/mL Cambridge Biosciences). Cells were then co-transfected with pCDNA5 IREVIP1 cSSH and the Flp recombinase expression plasmid, pOG44 (ThermoFisher Scientific). In short, for a 10 cm dish, 2.3 µg of pCDNA5 IREVIP1 cSSH and 20 µg of pOG44 were mixed (1:9 ratio), or 20 µg of pOG44 only, in 1.2 mL of Opti-MEM (ThermoFisher Scientific), followed by the addition of 60 µL of FuGENE HD (Promega), mixed gently, and incubated at RT for 15 minutes. Meanwhile, the cells were washed twice with 1x PBS and the media replaced with 10 mL of DMEM complete with Blasticidin (15 µg/mL) only. The Opti-MEM mixture was added to the cells. Two days later, the cells transfected with pcDNA5/pOG44 or pOG44 only were split in the following way in DMEM complete media containing Blasticidin (15 µg/mL) and Hygromycin B (50 µg/mL, Cayman Chemical). The cells were trypsinised and collected in a 50 mL Falcon tube (Greiner) in 25 mL of DMEM complete media containing Blasticidin and Hygromycin. 20 mL (80%) was plated in a 15 cm plate (Greiner) and media was topped up to 20 mL. The remaining 20% were further diluted 1/5 and 80% was plated again on a fresh 15 cm plate. This dilution step was repeated for a 1:25

dilution. The cells were left to grow until single colonies had formed in the pOG44/pCDNA5 plate and nearly all of the cells in the pOG44 plate were dead. Single colonies were then transferred to 12 well plates (Greiner) in DMEM complete media containing Blasticidin and Hygromycin B and grown up as individual clones. Clones were then tested for IREVIP1cSSH expression by the addition of 1.2 µg/mL Doxycycline and analysed by Western blot.

2.8.5 Transient siRNA Transfections

Cells were seeded at an appropriate density for the cell type and plate size. The following day, cells were transfected using Lipofectamine RNAiMAX (Invitrogen). The transfection was carried out in half the usual volume for the plate size, for example, a 6 cm plate would usually contain 4 mL of DMEM complete media, which would be reduced to 2 mL of DMEM complete media. For a 6 cm plate, the required amount of siRNA (See **Table 2-11**) was then resuspended in 137.5 mL of OptiMEM (Gibco), and 4 µL of Lipofectamine RNAiMAX was resuspended in 137.5 mL of OptiMEM. These were incubated separately for five minutes at room temperature. The OptiMEM containing the siRNA was then added to the OptiMEM containing Lipofectamine RNAiMAX, which were then incubated together for 15 minutes at room temperature. The mixture was then added in a dropwise manner to the cells. The depletion was carried out over 72 hours prior to the required end point. The depletion efficiency was evaluated by Western blotting. If the cells required splitting during the depletion, this was carried out the day after transfection.

Target Gene	Sequence	Working Concentration	Supplier
siNS (Non-silence, Luciferase)	CGUACGCGGAUACUUCGA	Varies	Microsynth AG
siIREVIP1-7	CUCAUCGACCUCUACCAGAA	20 nM	Microsynth AG
siIREVIP1-8	CGGCUGGAGAUCUACCAGGAA	20 nM	Microsynth AG
siTopoisomerase I	GGUCCUGUUGAGAAACGA	5 nM	Silencer Ambion
siTopoisomerase IIα	CAACUUUGAUGUGCGUGAA	5 nM	Silencer Ambion

Table 2-11 siRNA sequences and working concentrations

2.8.6 Plasmid Transfection

Cells were seeded in the required plate size and allowed to reach a confluency of 50%. Transfections were carried out using FuGENE HD (Promega) in a volume half of that usually used for that dish size. For example, a 6 cm plate would usually be cultured in 4 mL of DMEM media, therefore transfections were carried out in 2 mL DMEM complete media. For 2 mL of DMEM media, 2 µg of the desired plasmid was resuspended in 100 µL of OptiMEM (Gibco) along with 6 µL of Fugene HD. This was then vortexed gently five times and incubated at room temperature for five minutes. The mixture was then added in a dropwise manner to the DMEM complete media and left over night, after which the desired end point was carried out.

2.9 Preparation of Protein Lysates

2.9.1 Whole cell extracts (WCE)

Cells were collected by scraping in 1x PBS, centrifuged at 500 xg, followed by two washes with 1x PBS. Cell pellets were then lysed in 1x radioimmunoprecipitation assay (RIPA) buffer (25 mM TrisHCl, pH 7.4, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100), supplemented with protease (1 µg/mL Aprotinin, Chymostatin, Leupeptin, and Pepstatin, in addition to 1 mM PMSF), and phosphatase inhibitors (10 mM Sodium Fluoride, 1 mM Sodium Orthovanadate, 2 mM Sodium Pyrophosphate). Cells were resuspended in a volume of 1x RIPA double the volume of the cell pellet and left on ice for 30 minutes. Lysates were sonicated using the Bioruptor (Diagenode) at high intensity for 30 sec on, 30 sec off for 5 rounds to shear the DNA. Lysates were then centrifuged at 21,000 xg for 15 minutes at 4°C and the supernatant was collected. Protein concentration was determined using the RC DC Protein Assay Kit from BioRad.

2.9.2 Cellular Fractionation

To analyse the different subcellular fractions, cells were fractionated according to the method described by Mendez and Stillman (Méndez and Stillman, 2000). Briefly, cell were resuspended in buffer A (10 mM Hepes pH 7.45, 10 mM KCl, 340 mM Sucrose, 10% Glycerol, Protease and

Phosphatase Inhibitors (See WCEs above), and 2 mM EDTA). Triton X-100 was added to a final concentration of 0.1% and cells were left on ice for five minutes. Cells were then centrifuged at 350 xg for three minutes, the supernatant was collected as the cytosolic fraction, and the remaining pellet (nuclei) was then washed twice in buffer A without 0.1% Triton X-100. The nuclei were then resuspended in buffer B (3 mM EDTA, 0.2 mM EGTA, 5 mM Hepes pH 7.9, Protease and Phosphatase Inhibitors (See WCEs above)) and left on ice for 10 minutes. NP-40 was then added to a final concentration of 1% to remove membranes and was left to incubate on ice for a further five minutes. The nuclear fraction was then centrifuged at 1700 xg for five minutes, the supernatant was collected as the nuclear soluble fraction. The remaining pellet was washed twice more in buffer B with 1% NP-40, and then twice in Benzonase buffer (50 mM Tris HCl pH 7.9, 50 mM NaCl, 5 mM KCl, 3 mM MgCl₂, and protease and phosphatase inhibitors (See WCEs above)). The chromatin pellet was then centrifuged at 5000 xg for five minutes and resuspended in Benzonase buffer, 125 U of Benzonase (Merk Millipore) was added and incubated overnight rolling at 15 rpm at 4°C. The next morning, the Benzonase digestion was centrifuged at 20,000 xg for five minutes, the supernatant was collected as the chromatin soluble fraction. The pellet was resuspended in 2x Laemmli buffer with 15% βMe and boiled at 95°C for five minutes. This was then centrifuged at 5000 xg for 10 minutes and the supernatant was collected as the chromatin insoluble fraction. Fractions were then loaded as a percentage of the total volume, relative to the volume of the cytosolic fraction, in order to maintain the natural ratio in the cell.

2.10 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Acrylamide gels were homemade at a specific single percentage (e.g. 10%) or a gradient (e.g. 4-18%) using the mini-PROTEAN tetra hand casting system from BioRad. The stacking and separating gels were mixed as described below (See **Table 2-12**). Samples were made at the desired protein concentration by dilution with ddH₂O to an equal volume. 5x LB was added to a

final working concentration of 1x. Samples were then boiled at 95°C for five minutes and then centrifuged briefly using a benchtop centrifuge. Samples were loaded alongside 3 µL of precision plus dual colour standards (BioRad).

	Stacking gel (4%)	Separating gel (10%)	
Volume per gel	2.00	6.00	mL
ddH ₂ O	1.22	2.47	mL
4x buffer*	0.50	1.50	mL
30% acrylamide	0.27	2.00	mL
10% APS (final = 0.5%)	10.00	30.00	µL
100% TEMED (final = 0.06%)	2.00	4.00	µL
*4x stacking buffer: 0.5 M Tris, pH 6.8, 0.4% SDS			
*4x Separating buffer: 1.5 M Tris, pH 8.8, 0.4% SDS			

Table 2-12 Homemade gel stacking and separating gel recipes

2.11 Western Blotting

SDS-PAGE gels were transferred onto PVDF (BioRad) activated with 100% methanol for 15 seconds. Activated PVDF, sponges (BioRad), 3MM paper (Fisher) and gels were all equilibrated in transfer buffer prior to cassette assembly. Transfer tanks (BioRad) were filled with pre-cooled 1x transfer buffer and an ice pack. Proteins were transferred at 100V for 75 minutes.

After transfer, membranes were washed once with 1x TBST (0.1%) prior to blocking for one hour in 5% BSA in 1x TBST (0.1%) at RT. Antibody dilutions were prepared in 2% BSA-TBST (0.1%), in which membranes were incubated overnight rolling at 4°C. The following day, membranes were washed three times for five minutes in 1x TBST (0.1%) prior to incubation for one hour in HRP-conjugated secondary antibodies, prepared in 1% BSA-TBST (0.1%) at RT. Membranes were then washed three times for five minutes in 1x TBST (0.1%) at RT before activation, according to the manufacturer's instructions, with either SuperSignal™ West Pico Chemiluminescent Substrate

(ThermoFisher Scientific), Clarity™ Western ECL Substrate (BioRad), or SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific). The chemiluminescent signal was detected using the ChemiDoc MP system (BioRad). The antibodies used for Western blotting are listed in **Table 2-13**. All non-commercial antibodies were produced in the laboratory of Raimundo Freire, Unidad de Investigacion, Tenerife, Spain.

Antibody	Dilution	Source
Vinculin	1:1000	Abcam
β -actin	1:5000	Ambion
IREVIP1	1:1000	Non-commercial
HA	1:2000	Roche
6xHis	1:1000	ThermoFisher Scientific
Strep HRP (for Edu)	1:10,000	Abcam
dsDNA	1:5000	Abcam
Histone H3	1:3000	Abcam
pATM (Ser 1981)	1:10,000	Abcam
pATR (Ser 428)	1:1000	Millipore
Chk1	1:1000	Cell Signalling
pChk1 (Ser 345)	1:1000	Cell Signalling
Chk2	1:1000	Cell Signalling
pChk2 (Thr 68)	1:1000	Cell Signalling
γ H2AX (Ser 139)	1:1000	Novus Biologicals
SPRTN N-terminal	1:1000	Non-commercial
Topoisomerase I	1:5000	Bethyl
Topoisomerase II α	1:10,000	Bethyl
p97 (Mouse)	1:1000	Abcam
p97 (Rabbit)	1:1000	Non-commercial
UBXD8	1:1000	Non-commercial
UFD1	1:1000	Non-commercial
NPL4	1:1000	Non-commercial
PCNA	1:1000	Abcam
MCM7	1:1000	Santa Cruz
MCM3	1:1000	Santa Cruz
Polymerase δ	1:1000	Abcam
NPC (MAb414)	1:1000	Covance
Mouse-HRP	1:50,000	Sigma
Rabbit-HRP	1:50,000	Sigma
Rat-HRP	1:10,000	BioRad

Table 2-13 List of Western blot antibodies and working concentrations

2.12 DNA Damage Induction

2.12.1 Ionising Radiation (IR)

Cells were treated with the desired Grays (Gy) of IR using a ¹³⁷Caesium source (Gamma-Service Medical GmbH) at a dose rate of 1.3 Gy/min. Cells were left to recover as described for each experiment.

2.13 Immunofluorescence (IF)

Cells were seeded onto 10 mm No. 1 cover glasses (VWR) in the desired dish size. After the indicated time, cover glasses were moved one cover glass per well into a 12-well plate (Greiner) and were washed once with 1x PBS. If pre-extraction was desired, cells were incubated in pre-extraction buffer (25 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 3 mM MgCl₂, 300 mM Sucrose, 0.5% Triton X-100) for 2-5 minutes on ice, followed by two washes in 1x PBS. Either following pre-extraction or immediately after washing, cells were fixed in 4% formaldehyde in 1x PBS for 15 minutes, washed once in 1x PBS and then permeabilised using 0.5% Triton X-100 in 1x PBS for 5 minutes. Alternatively cells were fixed at RT in ice cold 100% MeOH for five minutes, followed by one wash in 1x PBS. Cells were then either kept in 1x PBS or blocked in 5% BSA in PBS at 37°C for one hour, followed by antibody staining. Subsequently, cells were washed in 1x PBS and transferred to a humid chamber, cells facing down, on 20 µL of diluted primary antibody on parafilm. Antibodies were diluted in 2% BSA in 1x PBS at the desired dilutions (See Table 2-14) and incubated on the cover glass for one hour at RT. The cover glasses were then transferred back to the 12-well plate and washed twice with 1x PBS, before repeating the process with the fluorescent dye conjugated secondary antibody (See **Table 2-14**). Cover glasses were then transferred back to the 12-well plate, washed twice with 1x PBS, once with 1x PBS containing DAPI (1 µg/mL) for 20 minutes at RT, followed by a final wash in 1x PBS. Cover glasses were then briefly dipped into ddH₂O and left to dry on a tissue before mounting onto a microscopy slide (Thermo Fisher Scientific) using FluoroMount-G hard mounting medium

(Interchim). Slides were imaged using the 60x objective on either the Nikon 90i or Nikon Ni-E epifluorescent microscope equipped with a Nikon DS-Qi1Mc digital microscope camera. If required, foci were counted automatically using ImageJ. Each nuclei was selected using an automated detection tool and the number of foci counted using the find maxima function set at a uniform threshold between siNS and siREVIP1 conditions. GraphPad Prism was then used to plot the mean of each data set with error bars for the SEM.

Antibody	Dilution	Source
IREVIP1 (Mouse)	1:100	Novus Biologicals
γ H2AX (Mouse)	1:1000	Millipore
53BP1 (Rabbit)	1:1000	Santa Cruz
HA (Rat)	1:1000	Roche
Alexa Fluor 488 donkey anti-mouse	1:1000	ThermoFisher Scientific
Alexa Fluor 488 donkey anti-rabbit	1:1000	ThermoFisher Scientific
Alexa Fluor 594 donkey anti-mouse	1:1000	ThermoFisher Scientific
Alexa Fluor 488 goat anti-rat	1:1000	ThermoFisher Scientific

Table 2-14 *List of Immunofluorescence antibodies and working concentrations*

2.14 Flow Cytometry

Cells were seeded at the desired density (See **Table 2-15**) and were treated with siRNA the following day. The day after siRNA depletion, cells were split in order to be no more than 50% confluent 72 hours after siRNA transfection, at which time EdU (ThermoFisher Scientific) was added to the media to a final concentration of 10 μ M and left for 30 minutes. Cells were then collected by trypsinisation and washed twice in 1x PBS. Cells were then resuspended in 4% PFA in 1x PBS and incubated for 15 minutes before washing again in 1x PBS. Cells were then resuspended in FACS buffer (0.25% Saponin, 1% FBS in 1x PBS) by vortexing and incubated for 15 minutes. The Click reaction was then prepared according to the manufacturer's instructions using the Click-iT® EdU Alexa Fluor® 647 imaging kit (ThermoFisher Scientific), and added directly

to the FACS buffer and left for a maximum of 30 minutes at RT. The cells were then washed once in 1x FACS buffer. Cells were then stained using phospho-Histone H3 (S10) Alexa Fluor® 488 conjugate (Cell Signalling), a marker of Mitosis, at a dilution of 1:100 in 1x FACS buffer for one hour. Cells were washed once in 1x FACS buffer. Cells were then resuspended in 1% BSA in 1x PBS containing 10 µg/µL RNase and 20 µg/mL Propidium Iodide and incubated for 30 minutes. Samples were then run on the FACSCalibre (BD Biosciences) and analysed using FlowJo®.

Day 1	Day 2	Day 3	Day 4	Day 5
HEK293 6cm 200 x10 ³ (siNS) 500 x10 ³ (siIREVIP1)	Transfect	Split ½ into 2 6cm the day after siRNA transfection	x	Add EdU 30 min, Collect for FACS
U2OS 6cm 90 x10 ³ (siNS) 120 x10 ³ (siIREVIP1)	Transfect	Split ½ into 2 6cm the day after siRNA transfection	x	Add EdU 30 min, Collect for FACS
RPE1 6cm 1:10 (400 µL of 10 mL from 10 cm into 4 mL media in 6 cm)	Transfect	Split ¼ and ¾ into 2 6cm the day after siRNA transfection	x	Add EdU 30 min, Collect for FACS
TIG1 6 cm 120 x10 ³ (siNS) 250 x10 ³ (siIREVIP1)	Transfect	Split 1/3 and 2/3 into 2 6cm the day after siRNA transfection	x	Add EdU 30 min, Collect for FACS

Table 2-15 Experiment outlines for flow cytometry for different cell lines

2.15 Electromobility Shift Assay (EMSA)

DNA probes labelled at their 5' end with IRDye800 (Integrated DNA Technologies) were either used as a single stranded template or were annealed to a complementary unlabelled strand (Eurogentec) by placing in a heat block at 95°C and allowing to cool to 37°C (See **Table 2-16**). The

DNA probes were then incubated in a reaction with either no protein, BSA, purified SP1, or purified IREVIP1 at 37°C for 15 minutes (See **Table 2-17**). Samples were then run on a 0.5% agarose gel in 1x TBE at 80 V for one hour at 4°C before scanning using the Licor Odyssey Scanner.

Reagent	Volume
Annealing Buffer (125 mM NaCl, 100 mM Tris pH 8.0)	40 µL
Labelled Probe (100 µM) 5'-ATTCGATCGGGGCGGGGCGAGC-3'	1 µL
Complementary Probe (100 µM) 5'-GCTCGCCCCGCCCCGATCGAAT-3'	1.5 µL
ddH ₂ O	57.5 µL

Table 2-16 Probe annealing for EMSA

Binding Reaction	Volume
2x Binding buffer (200 mM KCl, 40 mM Tris HCl, pH 7.5, 10 µM ZnSO ₄ , 0.4% NP-40, 40% Glycerol)	10 µL
0.1 M DTT	1 µL
1 µg Salmon DNA	1 µL
1 µg protein	x
ddH ₂ O	5.0 µL
Probe (1 pmol/µL)	1 µL

Table 2-17 Reaction conditions for EMSA

2.16 IREVIP1 Co-Immunoprecipitation (Co-IP) and Mass Spectrometry (MS)

HEK293 Flp/In T-Rex doxycycline (DOX) inducible IREVIP1 cSSH were plated at a density of 8.5 x10⁶ cells in 10x 15 cm plates. The following day, DOX was added to a final concentration of 1.2 µg/mL, to half of the plates, and left overnight. The following morning, cells were collected by trypsinisation, pooled, and washed twice in 1x PBS. Cells were lysed in 750 µL of 1x lysis buffer (20 mM Hepes pH 7.4, 10 mM KCl, 137 mM NaCl, 0.25% digitonin, 20 mM NEM, 5 mM EDTA, and protease and phosphatase inhibitors) and left to rotate at 15 rpm at 4°C for 20 minutes. The lysate was then centrifuged at 15,000 xg at 4°C for 15 minutes, the supernatant was then collect, referred to as the soluble fraction. The soluble fraction was then incubated with Strep-Tactin® Sepharose® (IBA) rolling at 15 rpm at 4°C overnight. The remaining pellet was then washed once with 1x lysis buffer supplemented with 0.1% Triton X-100, followed by three washes with 1x benzonase buffer (25 mM NaCl, 25 mM Tris HCl pH 7.9, 5mM KCl, 2 mM MgCl₂, and protease

and phosphatase inhibitors, see WCE)). The pellet was resuspended in 1x benzonase buffer containing 125 U of benzonase and was left rolling at 15 rpm at 4°C overnight. The following morning, the benzonase digested sample was incubated with Strep-Tactin® Sepharose® rolling at 15 rpm at 4°C for two hours. All samples were collected and the beads were spun at 50 xg for one minute at 4°C and the supernatant was removed. The beads were washed three times with 1x Immunoprecipitation (IP) buffer (50 mM NaCl, 5 mM KCl, 50 mM Tris HCl pH 7.5, 0.025% NP-40). Proteins were eluted by adding 2.5 mM biotin in 1x IP buffer and shaking the samples for 15 minutes at 4°C.

The samples were processed for MS by in solution digestion. In brief, the samples were reduced by addition of DTT to a final concentration of 5 mM and incubated for 30-60 minutes at RT. Iodoacetic acid (IAA) was then added to prevent reformation of disulphide bridges to a final concentration of 20 mM and incubated for 30-60 minutes at RT. The proteins were then precipitated by the addition of methanol, chloroform, and ddH₂O, in order to remove DTT and IAA. The samples were then spun twice at 7000 xg for 10 minutes. The upper aqueous phase was then removed, leaving the precipitate at the interface. Further methanol was then added before centrifuging again at 21,000 xg for two minutes. All of the supernatant was then removed, leaving a pellet of pure protein. The pellet was then resuspended in 6M Urea and sonicated for two minutes using the Bioruptor (5 sec on, 5 sec off for 10 cycles) to solubilise the protein. The samples were then diluted with ddH₂O to dilute the Urea to below 1M. Trypsin was then added to a concentration of 200 ng/μL and incubated at 37°C overnight. The following morning, formic acid was added to a final concentration of 0.1% to decrease the pH and prevent further trypsin digestion. The protein peptides were then purified using a SEP-PAK column and run on an Orbitrap LTQ Velos Elite Mass Spectrometer. The identified peptides were then matched against the UniProt SwissProt 20150204 database. Cytoscape, using the BiNGO application, was used to draw MS protein maps based on biological process GO pathways.

2.17 Isolation of Proteins on Nascent DNA (iPOND)

iPOND was performed as described in (Sirbu, Couch and Cortez, 2012), with the following modifications. Newly synthesized DNA in HEK293 ($\sim 2 \times 10^8$ cells per condition) was labelled by incubation with 10 μ M EdU for 10 minutes. For thymidine chase experiments, cells were washed in cell culture medium supplemented with 10 μ M thymidine and incubated for 10 minutes in thymidine-containing medium. Chromatin was fragmented into approximately 50-300 bp fragments by sonication with a Bioruptor Plus sonicator (30 seconds ON, 30 seconds OFF for 50 cycles). To isolate proteins associated with EdU-labelled DNA, samples were incubated overnight with streptavidin-coupled agarose beads (Merck Millipore). Aliquots of each extract were kept for use as loading controls. Before Western blotting, DNA slot blots were performed using a Streptavidin-HRP antibody to detect biotin-tagged EdU to ensure equal loading of DNA from wild-type and IREVIP1-depleted cells.

2.18 Rapid Approach to DNA Adduct Recovery (RADAR) DPC Isolation

DNA-protein crosslink (DPC) isolation by RADAR was performed as described in (Vaz *et al.*, 2016). DPCs were detected using a modified rapid approach to DNA adduct recovery (RADAR) assay (Kiianitsa and Maizels, 2013). Approximately, 2×10^6 HEK293 or HEK293 IREVIP1 Δ CRISPR cells were lysed in 1 mL of M buffer (MB), containing 6 M guanidium thiocyanate, 10 mM Tris-HCl (pH 6.8), 20 mM EDTA, 4% Triton X-100, 1% Sarkosyl and 1% dithiothreitol. DNA was precipitated by adding 1 mL of 100% ethanol and washed three times in wash buffer (20 mM Tris-HCl pH 6.8, 150 mM NaCl and 50% ethanol). DNA was solubilised in 1 mL of 8 mM NaOH and quantified using a Nanodrop and deposited onto a nitrocellulose membrane using a slot blotting apparatus.

2.19 DNA Fibre Assay

The DNA fibre assay was performed as described previously (Lessel *et al.*, 2014). Either siNS or siREVIP1 depleted HEK293 cells were labelled with 30 μ M of CldU (Sigma, C6891) for 30 min, washed 3x with warm PBS, then labelled with 250 μ M IdU (Sigma, 17125) for an additional 30 min. After labelling with IdU, cells were treated with ice-cold PBS to inhibit DNA replication. Cells were lysed in 200 mM Tris-HCl pH 7.4, 50 mM EDTA and 0.5% SDS, and DNA fibres were spread onto glass slides and fixed with 3:1 methanol and acetic acid overnight. The next day, DNA was denatured with 2.5 M HCl, blocked with 1% BSA and stained with anti-rat and anti-mouse 5-bromo-2'-deoxyuridine (BrdU) antibodies that specifically recognise either CldU (Abcam-Ab6326, dilution 1:500) or IdU (BD- 347580, dilution 1:100). Anti-rat Cy3 (dilution 1:300, Jackson Immuno Research, 712-116-153) and anti-mouse Alexa-488 (dilution 1:300, Molecular Probes, A11001) were used as secondary antibodies. Microscopy was performed using a Leica DMRB microscope with a DFC360FX camera. The lengths of the CldU- and IdU-labelled tracts were measured using ImageJ software.

3. RESULTS 1: IDENTIFICATION OF IREVIP1, A PREVIOUSLY UNCHARACTERISED PROTEIN

p97 is a ubiquitin dependent molecular segregase involved in virtually all processes within the cell (Meyer and Wehl, 2014). This includes, but is not limited to, protein quality control (endoplasmic reticulum associated degradation; ERAD), cell cycle control, membrane fusion, transcription and ribosome associated-degradation (RAD), DNA damage repair and replication (Chromatin associated degradation; CAD), and apoptosis (Hänzelmann, Buchberger and Schindelin, 2011; Vaz, Halder and Ramadan, 2013; Meyer and Wehl, 2014). In order to do this, p97 employs a plethora of co-factors, or adaptor proteins, which confer pathway specificity to p97 (Hänzelmann, Buchberger and Schindelin, 2011). There are at least 30 different adaptors responsible for directing p97 to specific pathways and substrates (Vaz, Halder and Ramadan, 2013; Meyer and Wehl, 2014). Currently it is thought that there are three main adaptor complexes which bind p97 mutually exclusively: p97^{UFD1-NPL4}, p97^{p47}, and p97^{UBXD1}. This is the starting point for a hierarchical system of adaptor binding, whereby different 'accessory' adaptor proteins further direct p97 specificity towards substrates and/or provide enzymatic activities (Hänzelmann and Schindelin, 2011; Meyer, Bug and Bremer, 2012; Vaz, Halder and Ramadan, 2013). The p97^{UFD1-NPL4} complex has been shown to be of importance for CAD in the response to DSBs (53BP1 access (Acs *et al.*, 2011), Ku removal (van den Boom *et al.*, 2016), and removal of K48-Ub conjugates (Meerang *et al.*, 2011)) and during DNA replication (degradation of CDT1 (Franz *et al.*, 2011; Raman *et al.*, 2011) and MCM7 (Maric *et al.*, 2017)).

Therefore, in order to gain a more in depth understanding of the role of p97 in CAD, in particular the interactome of p97 in response to IR-induced DNA damage, our laboratory carried out a SILAC based mass-spectrometry (MS) (**Figure 3-1 panel A and C**). The experiment was carried out using HEK293 cells expressing a p97 trapping mutant (p97 E578Q-Strep-HA, ATPase inactive). The cells were control treated or exposed to 10 Gy of IR and analysed at 30 minutes

(early) or 8 hours (late) after IR. Prior to p97 Strep-tag pull down, the cells were fractionated into cytoplasmic, nuclear, and chromatin fractions, therefore allowing the location of specific p97 interactions to be mapped within the cell. As anticipated, many known interactors of p97 were identified in the MS, therefore confirming its success, and furthermore many were shown to change their dynamics in response to IR. In particular in the nucleus, the

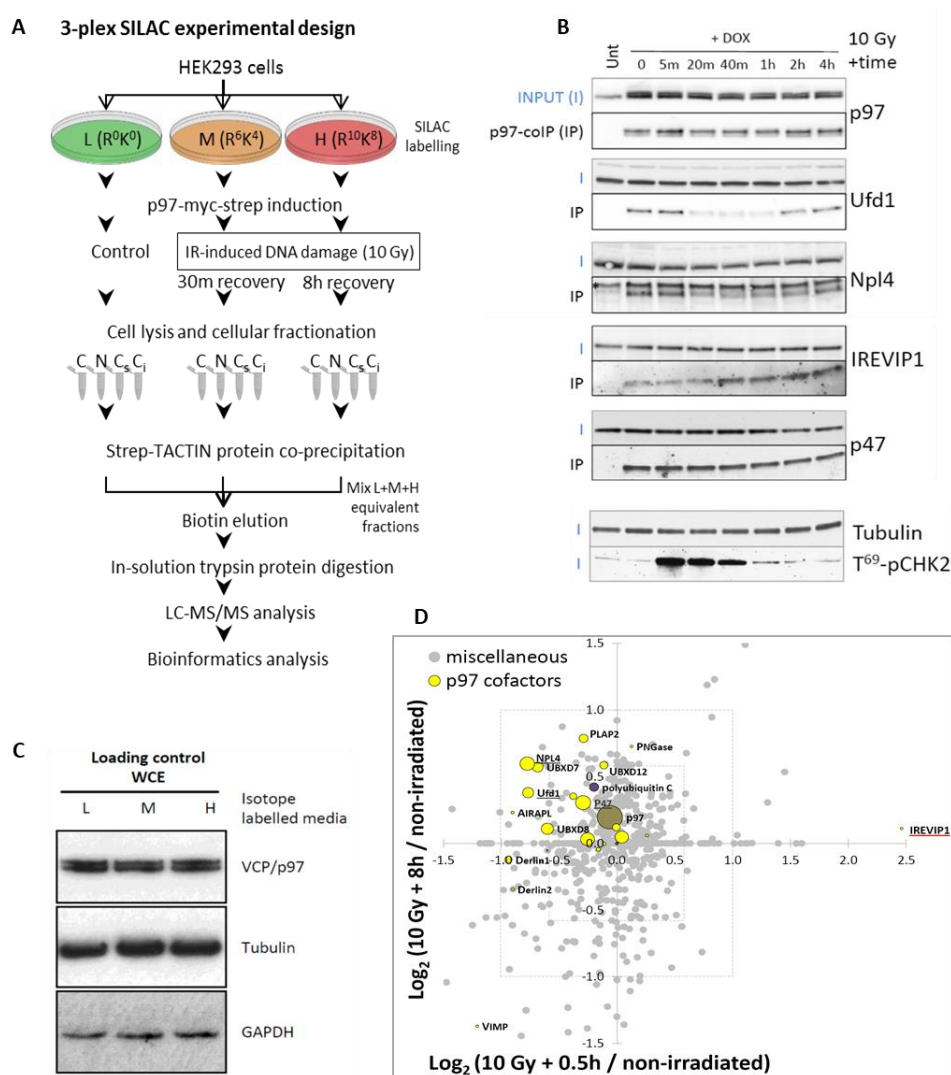


Figure 3-1 p97EQ SILAC MS in response to IR

A) Experimental outline of SILAC MS. HEK293 p97 FlpIn cells with either light, medium, or heavy labelled amino acids were incubated with Doxycycline to induce p97EQ expression. Cells were exposed to 10 Gy IR and left to recover for 30 minutes (medium) or 8 hours (heavy), or control treated (light), prior to lysis and fractionation. The light, medium, and heavy equivalent fractions were then mixed and a CoIP over streptactin beads was carried out. Samples were then processed for MS analysis. **B)** Western blot controls for the p97EQ Co-IP and confirming the MS interactions **C)** Western blot loading controls for light, medium, and heavy samples **D)** Graphical representation of nuclear interaction dynamics of known p97 interactors and IREVIP1 with p97 and 30 minutes and 8 hours. This data was produced by Dr I Torrecilla.

p97 adaptors NPL4, UFD1, and UBXD8, as well as polyubiquitin C, were seen to decrease their interaction with p97 30 minutes after IR. However, their interaction with p97 had increased 8 hours post-IR. Furthermore, a small increase in the amount of nuclear p97 was seen after 8 hours (**Figure 3-1 panels B and D**).

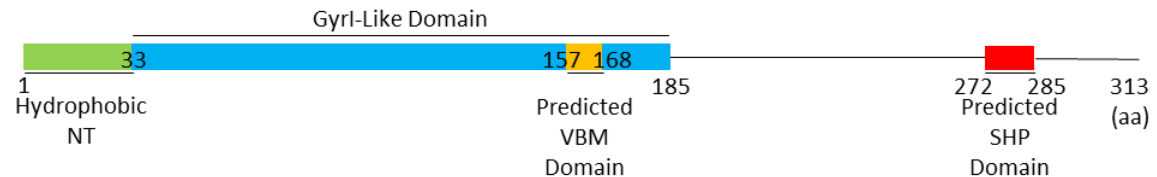


Figure 3-2 Domain schematic of IREVIP1
IREVIP1 was identified as having a hydrophobic leucine rich N-terminal domain, a Gyrl-like domain which homology to the bacterial protein Gyrl, and putative VBM and SHP domains for interaction with p97.

However, of particular note was a previously unidentified protein which, 30 minutes after IR, showed a striking 2.5 fold increase in its interaction with p97, which returned back to almost normal levels by 8 hours (**Figure 3-1 panels C and D**). This protein, here referred to as IREVIP1 (Ionising Radiation Enhanced VCP Interacting Protein 1), is previously uncharacterised. *In silico* analysis revealed a high conservation across eukaryotes, hydrophobic N-terminal domain, a Gyrase Inhibitory-Like (Gyrl-Like) domain, and two predicted p97 interaction domains (VBM and SHP) (**Figure 3-2, Figure 3-3 panels A, B, and C**). The Gyrl-Like domain was intriguing as it showed a degree of homology with the bacterial gyrase inhibitor, Gyrl (Nakanishi *et al.*, 1998) (**Figure 3-3 panel D**). The presence of two predicted p97 domains suggested that IREVIP1 could potentially be a new, uncharacterised, adaptor protein of p97. Analysis of the protein sequence using TMpred suggested that the hydrophobic N-terminal may function as a transmembrane domain, whereas the C-terminal was suggested to be disordered when analysed using the disorder prediction tool Protein DisOrder prediction System (PrDOS) (**Figure 3-3 panels E and F**). As IREVIP1 was identified as interacting with p97 in the nucleus, and their interaction was increased upon damage induction, it was hypothesised that IREVIP1 could be a new p97 adaptor involved in maintaining genome stability. As proteins involved in the DNA damage response are frequently mutated in cancer (Hanahan and Weinberg, 2011), cBioPortal was used to observe

the frequency of mutations, amplifications or deletions relating to IREVIP1 across many different cancers (**Figure 3-4**) (Cerami *et al.*, 2012). An increase in alteration frequency of IREVIP1 was observed, particularly in neuroendocrine prostate cancer (18.7% having amplifications) and kidney renal clear cell carcinoma (11.4% having deletion, 0.4% having mutations). This data would suggest that cancer cells are in some way modifying the function of IREVIP1 in order to become cancerous, or in order to better survive, therefore further highlighting the importance of IREVIP1 for normal cellular functions.

A

1

Human	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Chimp	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Pig	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Dog	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Rat	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Mouse	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Xenopus	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Zebrafish	MSDILLIGLGLALLLLLLTLLPAGYSGL	30
Score	***	

Human	KEWE---WKWRGLIVEAIDTQVDEGADTMSDSSSVSLEVSGSRETSAATISPGA	247
Chimp	KEWE---WKWRGLIVEAIDTQVDEGADTMSDSSSVSLEVSGSRETSAATISPGA	247
Pig	QALQPKWLVRRVEEAAVTRSDGADTMSDSSSVSLEVSGSRETSAATISPGA	250
Dog	KEWE---RKSRGAEASDQMDGADTMSDSSSVSLEVSGSRETSAATISPGA	247
Rat	KEWE---RKSRGAEASDQMDGADTMSDSSSVSLEVSGSRETSAATISPGA	246
Mouse	KEWE---RKCREIAEAYDQMDGADTMSDSSSVSLEVSGSRETSAATISPGA	246
Xenopus	KEWE---RKCREIAEAYDQMDGADTMSDSSSVSLEVSGSRETSAATISPGA	246
Zebrafish	KEWE---RKCREIAEAYDQMDGADTMSDSSSVSLEVSGSRETSAATISPGA	246
Score	***	

2

Human	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Chimp	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Pig	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Dog	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Rat	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Mouse	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Xenopus	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Zebrafish	LAGVVSAGSPPIRNVTVAAYKFMGLYGTGRFLFTECSISPKLRSIAVYDNP	85
Score	***	

Human	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Chimp	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Pig	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	305
Dog	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Rat	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Mouse	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Xenopus	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Zebrafish	SSRGWDDGDRSEHSYSESGASGSSFEELDIEGGPLGSRRLDPTGPLETKWL	302
Score	***	

Human	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Chimp	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Pig	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Dog	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Rat	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Mouse	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Xenopus	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Zebrafish	MVPPDKRCVAGSIIISGEESSPPELIDLYQKFGKVFSPAPSHVVVTAFFPYTH	140
Score	***	

Human	WEPTAEKKEKE 313	599
Chimp	WEPTAEKKEKE 313	278
Pig	WEPTAEKKEKE 313	6
Dog	WEPTAEKKEKE 313	425
Rat	WEPTAEKKEKE 313	159
Mouse	WEPTAEKKEKE 313	*
Xenopus	WEPTAEKKEKE 313	*
Zebrafish	WEPTAEKKEKE 313	*
Score	***	

3

Human	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Chimp	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Pig	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Dog	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Rat	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Mouse	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Xenopus	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Zebrafish	ILSIWLAARRVHPALDVIYKERRICAYPRLEIYQEDQIHFMCLARQGDFFVYBEM	195
Score	***	

Human	224	224
Chimp	224	224
Pig	224	224
Dog	224	224
Rat	224	224
Mouse	224	224
Xenopus	224	224
Zebrafish	224	224
Score	***	

h--Fp--G-G--hg

D

Human	59	59
Chimp	59	59
Pig	59	59
Dog	59	59
Rat	59	59
Mouse	59	59
Xenopus	59	59
Zebrafish	59	59
Score	***	

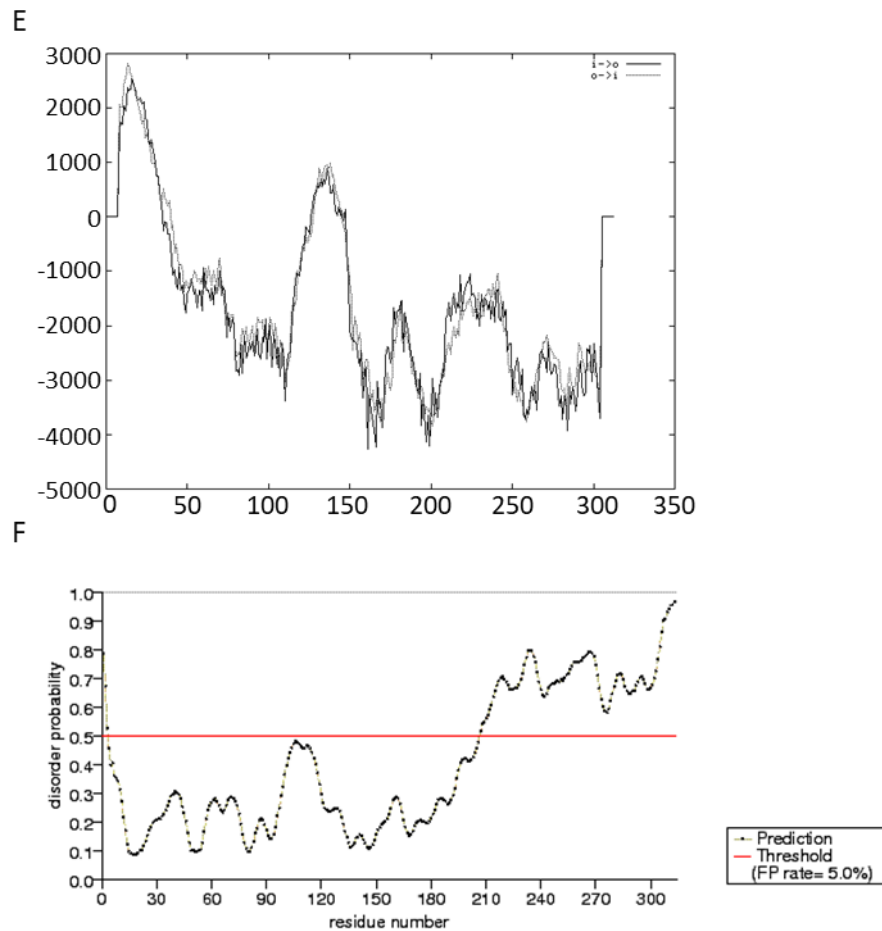


Figure 3-3 REVIP1 sequence alignment and protein characteristics

A) The IREVIP1 protein sequence was aligned with homologous sequences in other organisms using Clustal Omega. The red boxes denote where sequence region of interest. 1 – The leucine rich hydrophobic domain is maintained between species. 2 – This region is conserved between species, as well as being conserved in the bacterial protein Gryl (see D), and falls within the Gryl-like domain. 3 – This region is highly conserved between species and resembles a VBM domain characterised in other protein, with the E being highly conserved (see B). 4 – This region is highly conserved between species and resembles a SHP domain characterised in other proteins, especially the G-G residues (see C). **B)** Alignment of the predicted VBM domain in IREVIP1 with other known proteins. **C)** Alignment of the predicted SHP domain in IREVIP1 with other known proteins. **D)** Alignment of IREVIP1 with a homologous region of Gryl. See Table 2-4 for uniprot accession numbers. Known protein domains were previously aligned in Buchberger, Schindelin and Hänzelmann, 2015. **E)** A hydrophobicity plot of IREVIP1 produced using TMPred (Hofmann and Stoffel, 1993). **F)** A predicted disorder plot of IREVIP1 produced using Protein DisOrder prediction System (PrDOS) (Ishida and Kinoshita, 2007).

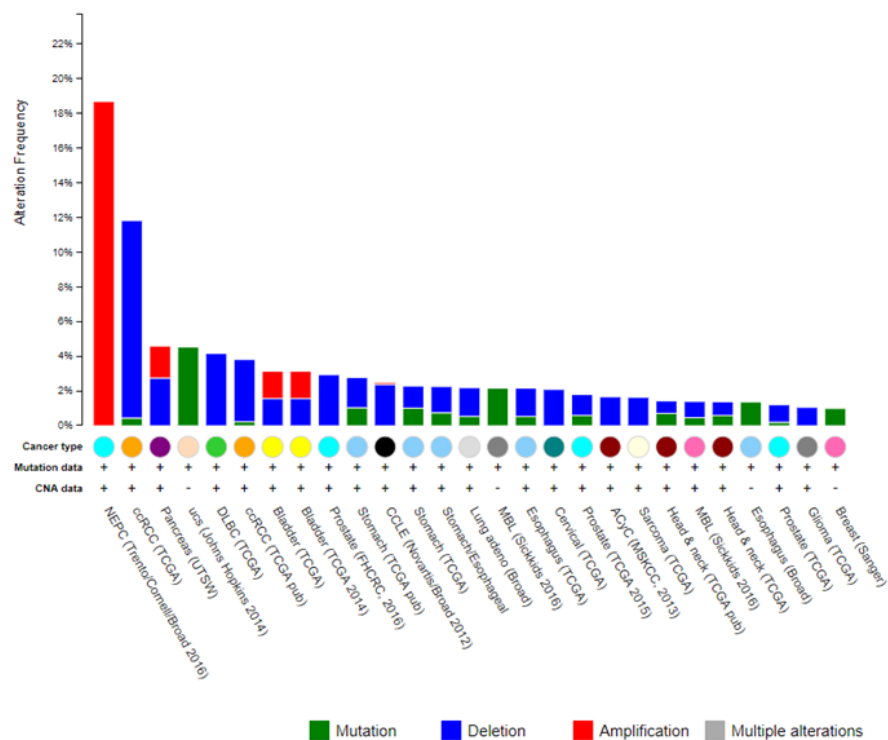


Figure 3-4 cBioPortal cross cancer alteration summary for IREVIP1
cBioPortal was used to generate a query for mutations, deletions, and amplifications in IREVIP1 across 167 large scale cancer genomics data sets. Produced using cBioPortal (Cerami *et al.*, 2012)

3.1 IREVIP1 Interacts With Key DNA Replication, Nuclear Pore, and Chromatin Topology Proteins

In order to further understand the processes in which IREVIP1 might be involved, and which complexes IREVIP1 might be found in, a MS analysis was performed to investigate the interactome of IREVIP1. Inducible HEK293T cells stably overexpressing IREVIP1cSSH were induced overnight using doxycycline, followed by cellular fractionation into soluble (cytosol and nucleosol) and chromatin pools. Co-IP was performed over streptactin beads under native conditions and physiological salt concentrations, before processing for MS. Upon analysis for biological function using the bINGO app for cytoscape (Maere, Heymans and Kuiper, 2005), the soluble fraction was observed to be particularly enriched for proteins involved in DNA replication, DNA repair, and ubiquitin-dependent protein degradation (**Figure 3-6 panel A**). Proteins involved in DNA replication include MCM (3-7), ORC5, PCNA, FEN1, PRIM1, PRIM2, RPA1, RPA2, RFC 1-5, and polymerase α and δ , highlighting proteins involved in both leading and lagging strand DNA synthesis. DNA repair proteins include MSH2, MSH3, MSH6 involved in mismatch repair, Cul4, XPB, TFIIH1-4 involved in nucleotide excision repair, XRCC1 and Ligase III involved in base excision repair, and Ku70, Ku80 and DNAPKcs involved in NHEJ. The analysis also highlighted numerous proteasome subunits, the p97 adaptor UBXD8, as well as p97. The chromatin fraction showed a very similar array of identified proteins to the soluble fractions (**Figure 3-6 panel B**). However, there were several additional key proteins identified here, namely Topoisomerase I, Topoisomerase II α , and numerous Nuclear Pore Complex (NPC) proteins (nucleoporins; NUPS).

In order to have more confidence in the proteins identified via MS, the Co-IP over IREVIP1cSSH was repeated using whole cell lysates and in the presence of ethidium bromide, to remove interactions mediated via DNA (Lai and Herr, 1992). Interactions were then confirmed by Western blot (**Figure 3-5**). As expected, the Co-IPs confirmed that IREVIP1cSSH physically forms

a complex with key DNA replication factors, namely polymerase δ , MCM7, and, PCNA, as well as Topoisomerase II α and Topoisomerase I. In addition, UBXD8, a previously known p97 adaptor protein, which was also identified in the IREVIP1 MS, was confirmed by the Co-IP to be in complex with IREVIP1. Interestingly NPL4 was also identified by the Co-IP, however, it was not seen in the IREVIP1 MS in either the soluble or chromatin fraction. Conversely, an interaction with UFD1 was not seen. Furthermore, the interaction with p97 was observed, however only very weakly under these conditions. Interestingly, the interaction with the Nuclear Pore Complex proteins was quite substantial and highlighted proteins of several different molecular weights, which reflects the multitude of NUPs seen in the MS.

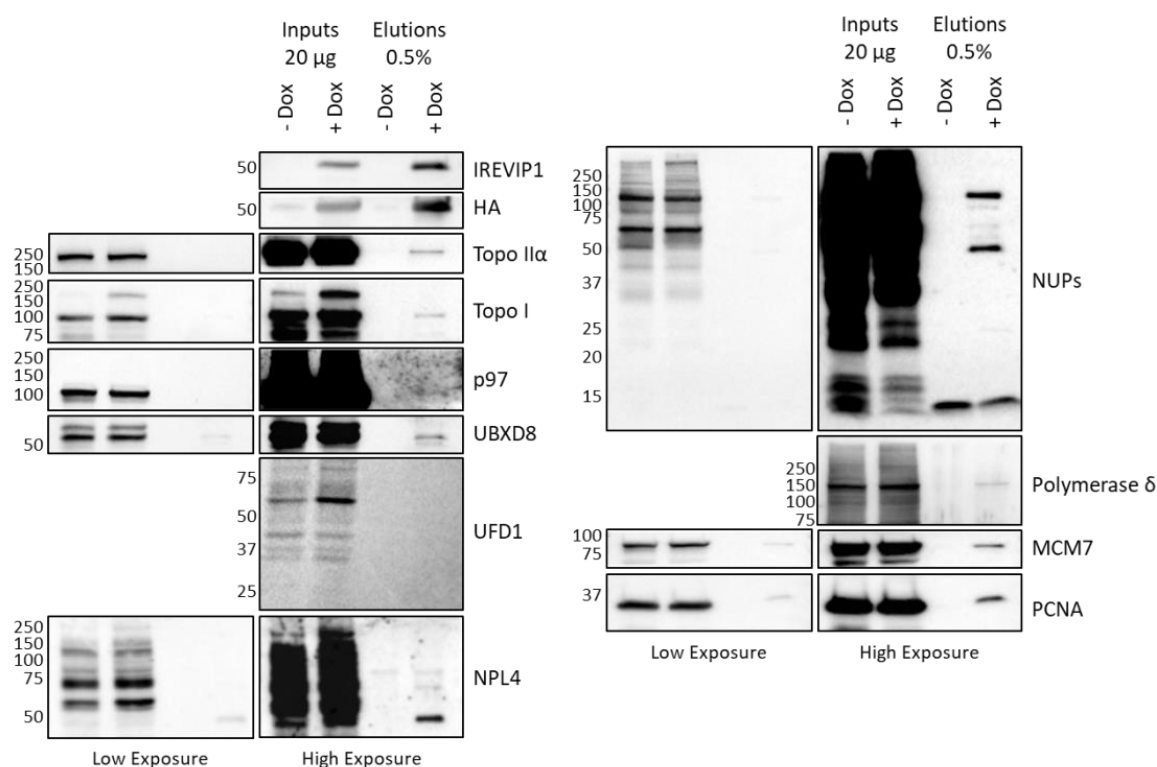


Figure 3-5 IREVIP1 MS interactions can be confirmed by Western blot

HEK293 cells were treated with Doxycycline overnight to induce IREVIP1cSSH expression. Cells were then lysed and all fractions pooled as a whole cell extract. A CoIP was performed in the presence of EtBr over streptactin beads and input and elutions were analysed by Western blot for proteins identified in the MS data sets. For certain antibodies, both low and high exposures are shown where eluted proteins were weakly seen in comparison to the input.

[illegible]

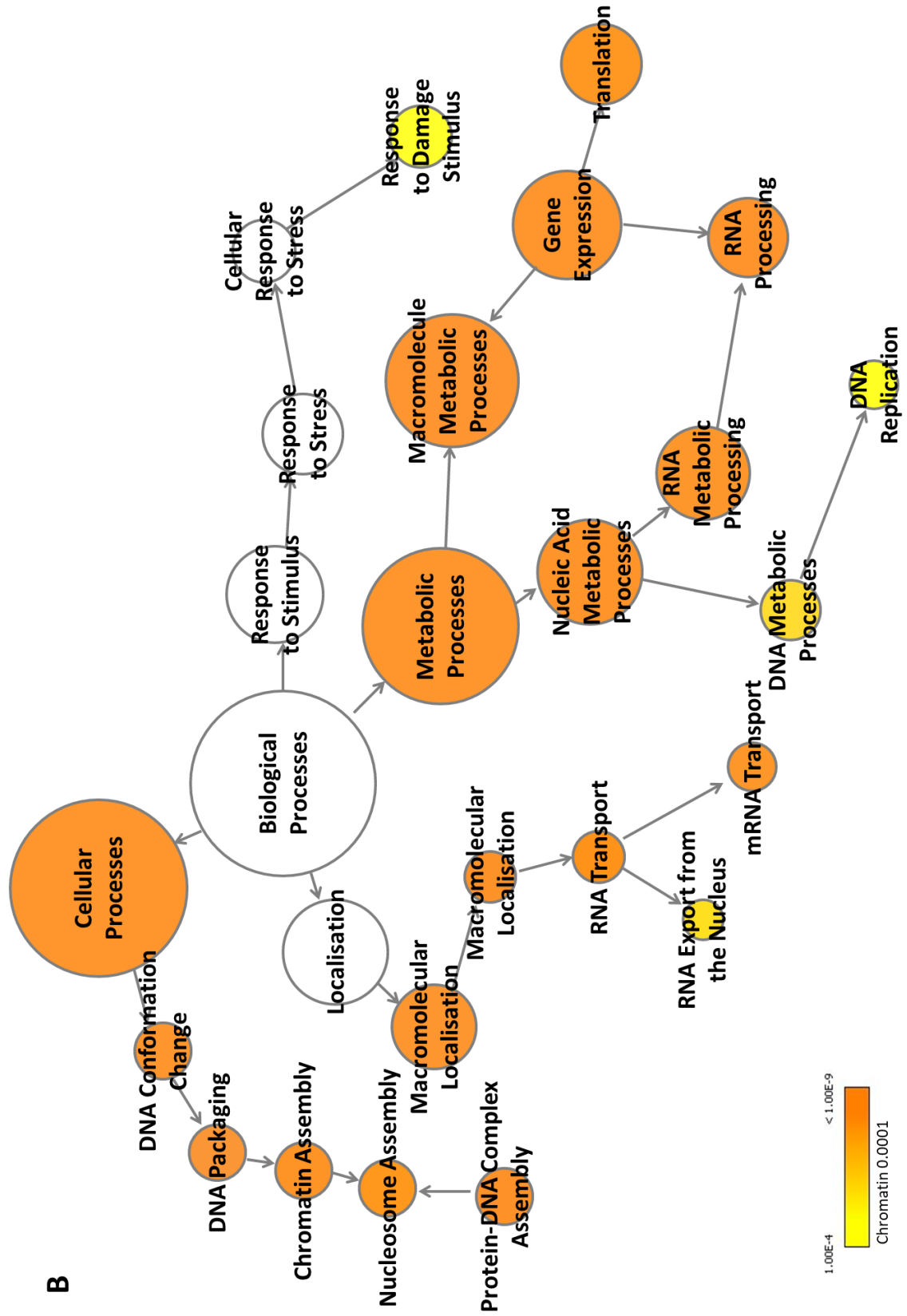


Figure 3-6 MS of IREVIP1 interacting proteins reveals its multi-pathway involvement

IREVIP1cSSH expression in HEK293 FlpIn cells was induced by addition of Doxycycline overnight. Cells were lysed and fractionated into soluble (cytoplasm and nuclear) and chromatin fractions prior to Co-IP over streptactin and sample processing for MS. Data sets were then normalised to non-induced samples and sorted by the number of significant sequences detected. All hits with more than two significant sequences were entered into Cytoscape and analysed using the BiNGO plugin to generate a gene ontology map of biological networks, based on the GO_Biological_Process file. A) Soluble fraction biological process gene ontology map. B) Chromatin fraction biological process gene ontology map. This significance level was set to < 0.00001 for both fractions. The colour gradient represents the P-value for the probability that the genes identified belong to a functional category shared by the genes in the reference set.

3.2 IREVIP1 Is Localised at the Nuclear Membrane by its Hydrophobic N-Terminal

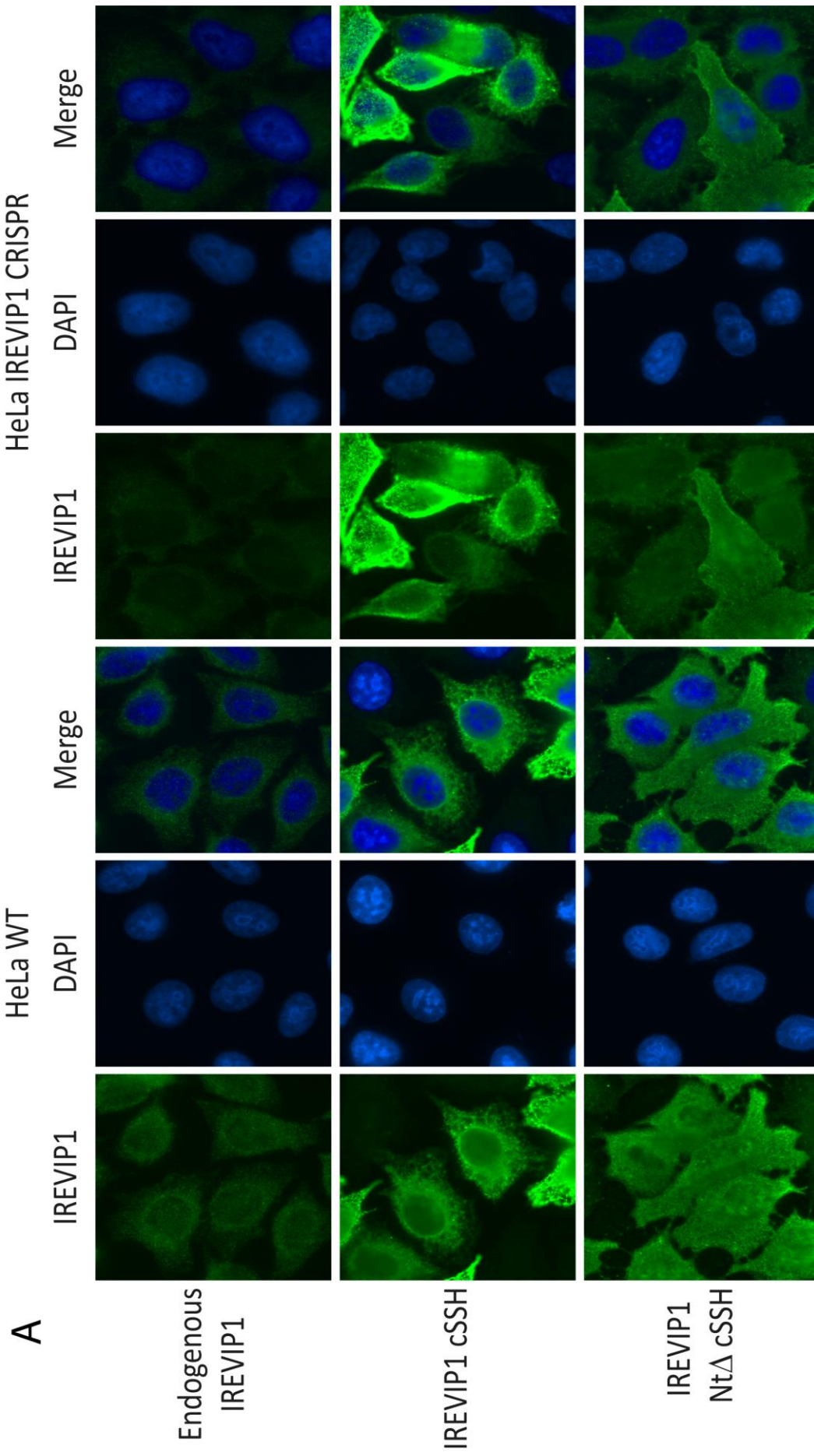
Domain

As shown by MS and Co-IP, IREVIP1 has been seen to interact with several NUPs (**Figure 3-5**). Therefore, it was hypothesised that IREVIP1 might be localised in the region of the nuclear membrane as well. Furthermore, Lamin B1 and Lamin B2, as well as pre-Lamin A/C, were pulled down by IREVIP1 in the chromatin fraction, whereas only Lamin B1 was seen in the soluble fraction (data not shown), further suggesting a role for IREVIP1 at the nuclear membrane.

Therefore, to test this hypothesis, HeLa cells seeded onto coverslips were depleted for IREVIP1 using two different siRNA sequences. Depleted cells were then fixed, permeabilised, and blocked prior to staining with a commercial IREVIP1 antibody and DAPI, followed by fluorescent imaging. IREVIP1 could be seen to be specifically located at the nuclear rim. This nuclear rim staining subsequently disappeared upon depletion with both siRNA sequences (**Figure 3-7 panel B**). Some non-specific staining could be seen across the cell, often including some nucleoli and ER staining, however, this was still observable post IREVIP1 depletion. Differences in the observed nucleoli staining in some cells and experiments could be due to experimental variability, such as permeabilisation variability and antibody penetration into the cell, the effectiveness of staining, the lot of antibody, or even the batch of cells. However, as this did not disappear with siRNA depletion, confirmed by Western blot, it could be hypothesised that this is non-specific staining. It would be prudent to validate a second antibody, or use endogenous tagging, to corroborate this. The nuclear rim staining was also reduced when analysing IREVIP1

CRISPR knock-out (IREVIP1 Δ) HeLa cells. IREVIP1 could be seen as a rim around the nucleus in HeLa WT cells, whereas this rim was not visible in HeLa IREVIP1 Δ CRISPR cells. Some background staining is still visible despite IREVIP1 knocked out, further suggesting that it could indeed be non-specific staining and that more work needs to be done to optimise staining and visualisation techniques, as well as colocalisation with known nuclear membrane proteins, such as lamins (**Figure 3-7 panel A**).

Furthermore, it was also hypothesised that the hydrophobic N-terminal domain of IREVIP1 could be responsible for IREVIP1 localisation to the nuclear membrane. To test this hypothesis, IREVIP1cSSH WT and IREVIP1cSSH NT Δ (N-terminal deletion) were transiently overexpressed overnight in both HeLa WT and HeLa IREVIP1 Δ cells. Upon expression of IREVIP1cSSH WT in both HeLa WT and HeLa IREVIP1 Δ cells, IREVIP1cSSH could be seen to be located at the nuclear rim in all cells overexpressing IREVIP1cSSH. However, overexpression of IREVIP1 did also result in some staining which could be localised to the ER. This could be due to transient plasmid based overexpression resulting in very high levels of IREVIP1 expression, as shown by Western blot (**Figure 3-7**). This could be further investigated by using CRISPR to tag the endogenous protein, allowing more accurate localisation. This could also be coupled with co-staining with nuclear rim proteins such as Lamin to confirm locality. It could also be the case that IREVIP1 is localised throughout the cell, with the nuclear rim being only one, or possibly the main, location of the protein, which would be wise to assess by means such as cellular fractionation. In support of nuclear localisation, upon expression of IREVIP1cSSH NT Δ in both HeLa WT and HeLa IREVIP1 Δ cells, IREVIP1cSSH NT Δ was seen to be located throughout the cell in a uniform manner, with no nuclear rim localisation (**Figure 3-7 panel A**). This suggests that the N-terminal hydrophobic domain of IREVIP1 is necessary for localisation to the nuclear rim. All depletions and over-expressions were confirmed by WB using a non-commercial IREVIP1 antibody. Endogenous IREVIP1 can be seen as the lower band of a doublet slightly larger than 37 kDa, showing substantial IREVIP1 depletion with siRNA and in the HeLa IREVIP1 Δ cells (**Figure 3-7 panel C**).



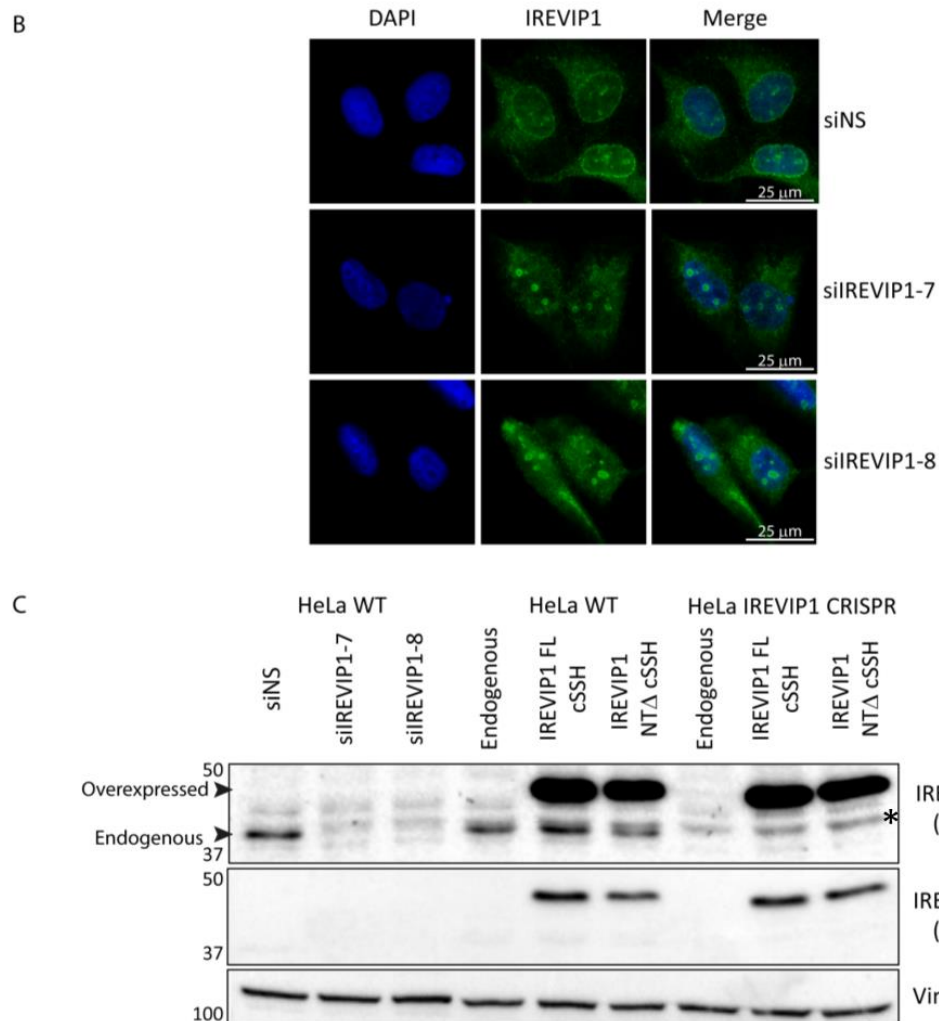


Figure 3-7 IREVIP1 is located at the nuclear rim via its hydrophobic N-terminal domain

A) IREVIP1 cSSH or IREVIP1 NTΔ cSSH was overexpressed in HeLa WT or HeLa IREVIP1Δ CRISPR cells prior to fixing and staining with an endogenous IREVIP1 antibody and DAPI. Overexpression of IREVIP1 NTΔ cSSH abolished the localisation of IREVIP1 to the nuclear rim. **B)** Depletion of IREVIP1 in HeLa cells with two siRNA sequences indicates the specificity of IREVIP1 nuclear rim staining. **C)** IREVIP1 is depleted in HeLa cells with both siRNA sequences and control non-silencing siRNA (luciferase). IREVIP1 cSSH and IREVIP1 NTΔ cSSH are overexpressed to similar levels. IREVIP1 is not present in IREVIP1Δ CRISPR cells (lower band of endogenous doublet in IREVIP1 overexpressing cells). * indicates a non-specific band.

3.3 Discussion

Chromatin serves as a platform for a wide range of protein interactions governing many different processes from transcription, DNA replication and repair to mitosis. In order for these processes to be carried out, there needs to be a degree of chromatin-associated protein turnover to allow chromatin dynamics, thereby maintaining genome integrity (Vaz, Halder and Ramadan, 2013). p97 is a molecular segregase which removes proteins from chromatin in a process termed Chromatin Associated Degradation (CAD). In order to carry out this function, p97 requires adaptor proteins to confer pathway specificity. For the nuclear functions of p97, in particular for DNA damage repair and replication, the p97^{UFD1-NPL4} core complex is responsible for these actions (Vaz, Halder and Ramadan, 2013). However, the process of CAD is still poorly understood.

A MS experiment was carried out to further investigate this, which identified IREVIP1 (**Figure 3-1**). IREVIP1 has homology to the bacterial DNA gyrase inhibitory GyrI and was found to form a complex with key DNA repair and replication proteins, as well as Topoisomerase I and II α and NUPs (**Figure 3-5**). The weak interaction with p97 could be explained by the substrate processivity of p97. The p97 MS in which IREVIP1 was first identified was performed using a substrate trapping mutant of p97 which prolongs the time that p97 is in complex with binding partners (Ye, Meyer and Rapoport, 2003; Meerang *et al.*, 2011). This could suggest that the interaction between IREVIP1 and p97 is quite transient in nature. Also of interest was the identification of NPL4, by WB, as being in complex with IREVIP1, whereas UFD1 was not. This was unexpected as NPL4 and UFD1 are usually found in complex (Vaz, Halder and Ramadan, 2013). NPL4 possesses a UBD p97 interaction domain, whereas UFD1 has a SHP domain (Yeung *et al.*, 2008). Therefore, UFD1 and IREVIP1, which also has a predicted SHP domain, could be in competition for the same p97 binding site. However, p97-UFD1-NPL4 bind in a ratio of 6:1:1, therefore potentially leaving five other SHP binding motifs available for IREVIP1 binding,

therefore competitive binding is unlikely (Hänzelmann and Schindelin, 2016). Otherwise, the levels of UFD1 could simply be below the detection limit of the antibody, especially seeing as when the UFD1-NPL4 complex was first identified it was reported that NPL4 cannot bind to p97 in the absence of UFD1 (Meyer *et al.*, 2000). Therefore, this would be interesting to re-test using a different antibody against UFD1 and to potentially test complex formation between p97-UN and IREVIP1 in vitro. Thus, IREVIP1 interacts with proteins across DNA damage repair, DNA replication, DNA topology regulation, and nuclear envelope proteins, suggesting a role for IREVIP1 in maintaining genome stability in relation to DNA processes occurring around the nuclear membrane. The presence of UBXD8 in the IREVIP1 Co-IP was unexpected, as UBXD8 is usually associated with ERAD and Lipid Droplets (Olzmann and Kopito, 2011; Olzmann, Richter and Kopito, 2013; Meyer and Wehl, 2014). Therefore, this could just be an artefact and could have been carried along with p97, or due to the fact that IREVIP1 is localised to the nuclear membrane, which is contiguous with the ER (Mattaj, 2004). However, the interaction did appear to be quite robust, nevertheless, this could be down to the quality of the antibody.

These findings nicely link DNA replication and transcription as processes which often occur together in regions of DNA located at the nuclear membrane (Bermejo *et al.*, 2009; Mekhail and Moazed, 2011; Ibarra and Hetzer, 2015). Actively transcribed DNA can be tethered to the NPC to facilitate the export of mRNA from the nucleus to the cytoplasm for translation, in a process referred to as gene gating (Bermejo *et al.*, 2011). Therefore, as the transcription machinery approaches the tethering point at the membrane and the DNA is opened, DNA torsion will build up in the section of DNA ahead of the region. In order to modulate this tension, Topoisomerase II is needed (Bermejo *et al.*, 2011). This can be the case with DNA replication as well, whereby if a replication fork is approaching a torsional block, or transcription machinery travelling in the opposite direction, DNA torsion will build up in the connecting section of DNA and Topoisomerase I is needed to release this (Bermejo *et al.*, 2011). However, there is a large amount of crossover between the different Topoisomerases (Bermejo *et al.*, 2009; Pommier *et*

al., 2016). IREVIP1 was found to be located at the nuclear rim via its N-terminal domain (**Figure 3-7**). Returning to the Gyrl-like domain identified in IREVIP1, showing homology to the bacterial DNA gyrase inhibitory Gyrl (**Figure 3-2**), this may suggest that IREVIP1 plays a role in the regulation of Topoisomerases, alongside p97, relating to processes causing DNA torsion around the nuclear rim.

In order to confirm these as direct interactions, they had to be shown *in vitro*. Therefore, the next two chapters will cover the purification of IREVIP1 and the *in vitro* interaction between p97-IREVIP1 and Topoisomerase I and Topoisomerase II α .

4. RESULTS 2: PURIFICATION OF PROTEINS FOR *IN VITRO* STUDIES

IREVIP1 was previously shown to interact with p97 in cells (**Figure 3-1, Figure 3-5**), and predicted SHP and VBM p97 interaction domains were both identified *in silico* (**Figure 3-2**). Therefore, it was hypothesised that IREVIP1 is a potential new p97 adaptor. In order to confirm this, the p97-IREVIP1 interaction had to be shown *in vitro* and whether either of the predicted p97 interaction domains were responsible for the interaction. To be able to do this IREVIP1 first has to be purified.

4.1 Extensive Construct Optimisation was Required in Order to Purify IREVIP1

4.1.1 IREVIP1 N-Terminal GST

Full length IREVIP1 was initially cloned into the vector pGEX-4t-2 with an N-terminal GST tag and transformed into BL21 *E. coli* after sequencing (**Figure 4-2 panel A**). The expression of IREVIP1nGST was then tested by inducing expression using 1 mM IPTG upon reaching an OD₆₀₀ of 0.6 and incubating at 37°C for three hours, or either 25°C or 4°C overnight (**Figure 4-1 panel A**), as lower temperatures can improve solubility (Costa *et al.*, 2014). Pelleted bacteria were resuspended in lysis buffer, sonicated, and then centrifuged to collect the soluble and pellet fractions (**Table 4-1**). These were then analysed separately by WB for IREVIP1 expression using an IREVIP1 antibody. IREVIP1nGST was insoluble at 37°C and 25°C with most of the protein found in the pellet fraction. About half of the protein produced at 4°C was soluble, however expression was very low overall (**Figure 4-1 panel A**). In an attempt to improve the amount of IREVIP1nGST produced, bacteria were grown at 4°C for either 24 h or 48 h and then analysed for IREVIP1nGST. This extended induction time did increase the amount of IREVIP1nGST; however, the protein was almost exclusively insoluble, being seen only in the pellet fraction (**Figure 4-1 panel B**).

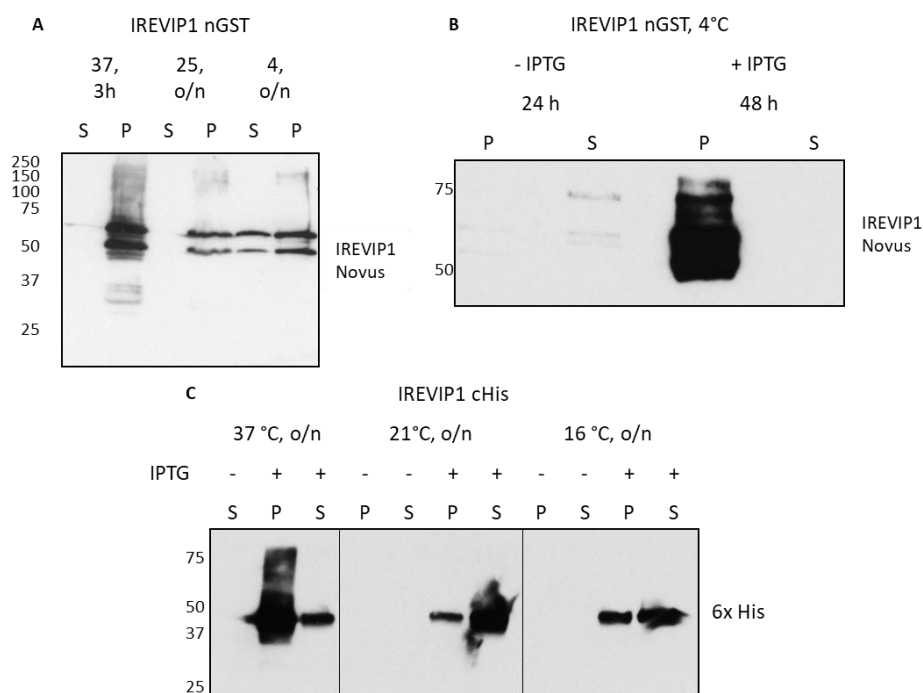


Figure 4-1 IREVIP1 nGST is poorly soluble

A) IREVIP1 nGST expression in BL21 *E. coli* was tested at 37°C for three hours, or 21 or 4°C overnight. Soluble and pellet fractions were examined by Western blot using the indicated antibody. Mild solubility was seen at 4°C. **B)** IREVIP1 nGST expression at 4°C for 24 or 48 hours was tested and examined by Western blot using the indicated antibody. IREVIP1 nGST showed high expression after 48 hours but was entirely insoluble. **C)** IREVIP1 cHis expression was tested overnight in Rosetta *E. coli* at 37, 21 and 16°C. IREVIP1 cHis was highly insoluble at 37°C. Solubility was greatest at 21°C.

4.1.2 IREVIP1 C-terminal His

Following this, a different approach was taken where IREVIP1 was cloned into pET21a(+) with a C-terminal polyhistidine (6xHis) tag and was expressed in Rosetta *E. coli* (**Figure 4-2 panel B**). Expression was tested at 37°C, 21°C, and 4°C overnight in the presence and absence of 1 mM IPTG (**Table 4-1**). Without IPTG, IREVIP1cHis was not detectable by WB with an antibody against 6xHis. Expression was seen with IPTG in all cases. At 37°C, a large amount of IREVIP1cHis expressed was insoluble and seen in the pellet fraction. However, at 21°C and 15°C, at least 50% of the expressed protein was soluble, with the best solubility and expression being seen at 21°C (**Figure 4-1 panel C**). Initially, IREVIP1cHis purification optimisation was begun in batch using Ni-NTA beads (Qiagen), however the proportion of IREVIP1cHis binding to the beads was low. At least 50% of IREVIP1cHis was lost in the flow-through. In an attempt to improve this, 5 mL HisTrap FF Crude columns for the ÄKTA FPLC system were tested as they provided a larger volume of beads and binding capacity.

Construct	IPTG	Induction		Lysis Buffer	Result	Modification
IREVIP1 nGST	0.5 mM OD ₆₀₀ 0.6	37°C, 3 h 25°C, o/n 4°C, o/n		10% sucrose, 50 mM Tris HCl pH 8.0, 150 mM NaCl, 0.5% NP-40, 1 mM DTT, 1x cOmplete, 1 mg/mL lysozyme	Insoluble	Change Tag and Tag position
IREVIP1 cHis	1 mM OD ₆₀₀ 0.6	37°C, o/n 21°C, o/n 16°C, o/n		10% sucrose, 50mM Tris HCl pH 8.0, 150 mM NaCl, 0.5% NP-40 1x cOmplete 1 mg/mL lysozyme	Soluble at 21 °C, Capture efficiency is low	Try ÄKTA system with 5 mL HisTrap FF columns
IREVIP1 cHis	1 mM OD ₆₀₀ 0.5	21°C, o/n		500 mM NaCl, 20 mM Tris HCl pH 7.0, 5 mM Imidazole, 1 mM PMSF, 1 mg/mL lysozyme	Yields low, Protein forms aggregates/ oligomers	Remove hydrophobic N- terminal domain
IREVIP1NTΔ cHis	1 mM OD ₆₀₀ 0.6	37°C 21°C 16°C	2 h 4 h 6 h 8 h o/n	500 mM NaCl, 20 mM Tris HCl pH 7.0, 5 mM Imidazole, 1 mM PMSF, 1 mg/mL lysozyme	Protein yield better at 21 °C, 6 h, Purification yield still low, protein is lost in the ÄKTA system filters	Purify in batch using Qiagen Ni- NTA Beads. Try native and denaturing.
IREVIP1NTΔ cHis	1 mM OD ₆₀₀ 0.6	21°C, 6 h		800 mM NaCl, 20 mM Tris HCl pH 7.0, 5 mM Imidazole, 1% Triton X-100, 1 mM PMSF, 1 mg/mL lysozyme	Sufficient yields from one step, batch, His purification	Currently in use

Table 4-1 IREVIP1 purification optimisation

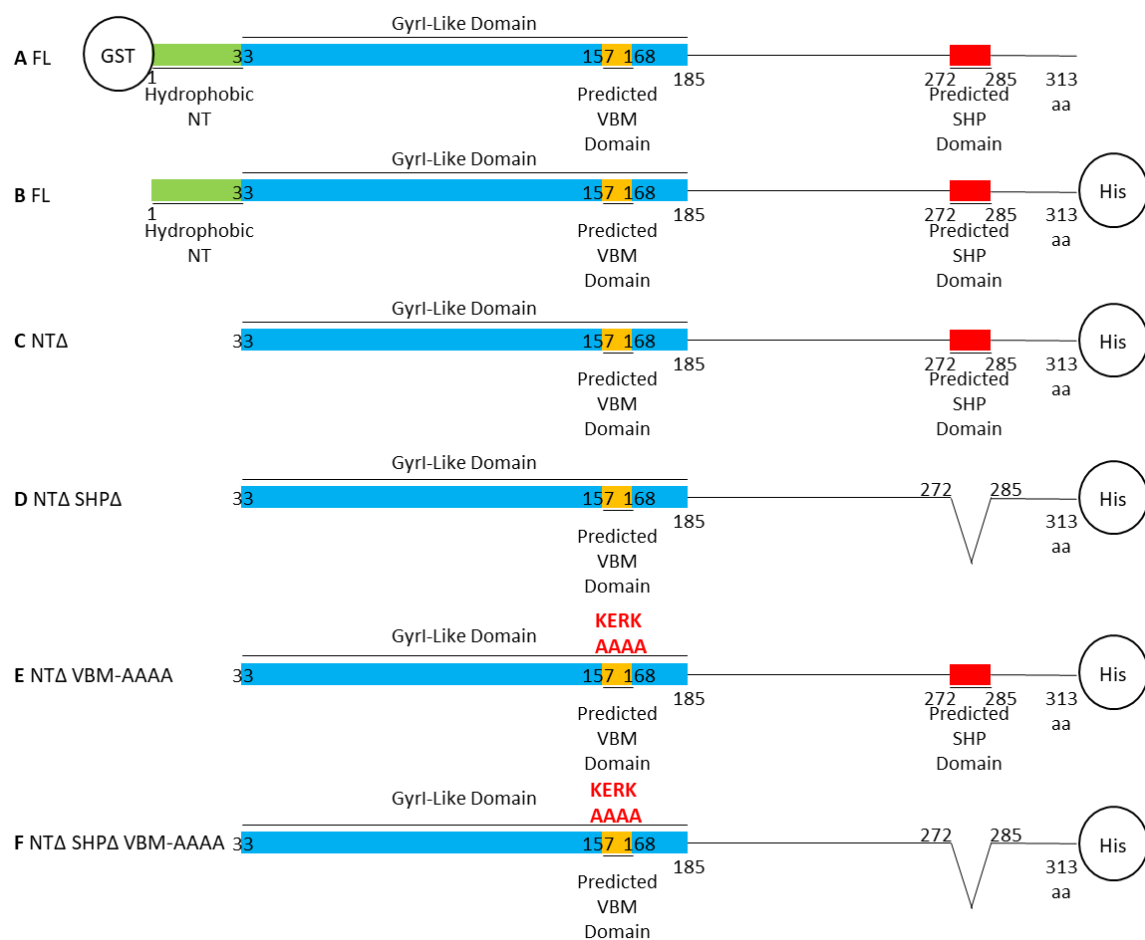


Figure 4-2 Schematics of IREVIP1 constructs

Schematics of IREVIP1 constructs of protein purification. A) Full length IREVIP1 with an N-terminal GST tag. B) Full length IREVIP1 with a C-terminal 6xHis tag. C) IREVIP1 N-terminal deletion with a C-terminal 6xHis tag. D) IREVIP1 N-terminal and SHP domain deletion with a C-terminal 6xHis tag. E) IREVIP1 N-terminal deletion and VBM mutant with a C-terminal 6xHis tag. F) IREVIP1 N-terminal and SHP domain deletion and VBM mutant with a C-terminal 6xHis tag.

Unfortunately, this did not increase the proportion of binding or elution yields, but did provide a more convenient purification method (**Table 4-1**). To try and improve IREVIP1cHis solubility and its binding to the resin, the NaCl concentration was increased from 150 mM to 500 mM (Nordlund *et al.*, 2008). Furthermore, the purified IREVIP1cHis was unable to enter a native acrylamide gel, despite having an isoelectric point (pI) of 4.7, and accumulated at the stacking/separating gel interface (**Figure 9-1**) (**Table 4-1**). This suggested that IREVIP1cHis was forming some form of high molecular weight oligomer or aggregate.

4.1.3 IREVIP1 C-terminal His and Hydrophobic N-terminal Removal

IREVIP1 has a hydrophobic N-terminal domain (**Figure 3-2, Figure 4-3**). Therefore to try and improve solubility and reduce the high molecular weight forms, IREVIP1 was amplified without its N-terminal domain (25 aa) and cloned into pET21a(+) with a C-terminal His tag (IREVIP1 cHis NTΔ) (**Figure 4-2 panel C**). The online tool TMpred (Hofmann and Stoffel, 1993) was used to identify the hydrophobic region and confirm that the NTΔ construct was within reasonable hydrophobicity limits (**Figure 4-3 panel B**). This construct was transformed into Rosetta *E. coli* prior to expression testing. Upon reaching an OD₆₀₀ of 0.6, cells were either grown at 37°C, 21°C and 4°C for 0, 2, 4, 6, or 8 hrs and overnight in the absence or presence of IPTG to induce IREVIP1 NTΔ cHis expression (**Figure 4-3 panel D**). Prior to each collection of +/- IPTG induced samples the OD₆₀₀ was read as a measure of how well the cells were growing. This suggested that induction of IREVIP1 is toxic at 37°C, showing a large drop off in cell density from between four and six hours post induction with IPTG, and significantly less growth than cells without IPTG at all time points (**Figure 4-3 panel C**). Cells grown at 21°C in the presence of IPTG showed a much similar pattern of growth to those grown without IPTG, suggesting that this is much less toxic to the cells. This was also the case with cells grown at 16°C, however the overall cell density was lower even in the absence of IPTG, therefore overall protein yields could be lower. Consequently, 21°C was selected as the growth temperature of IREVIP1 NTΔ cHis expression.

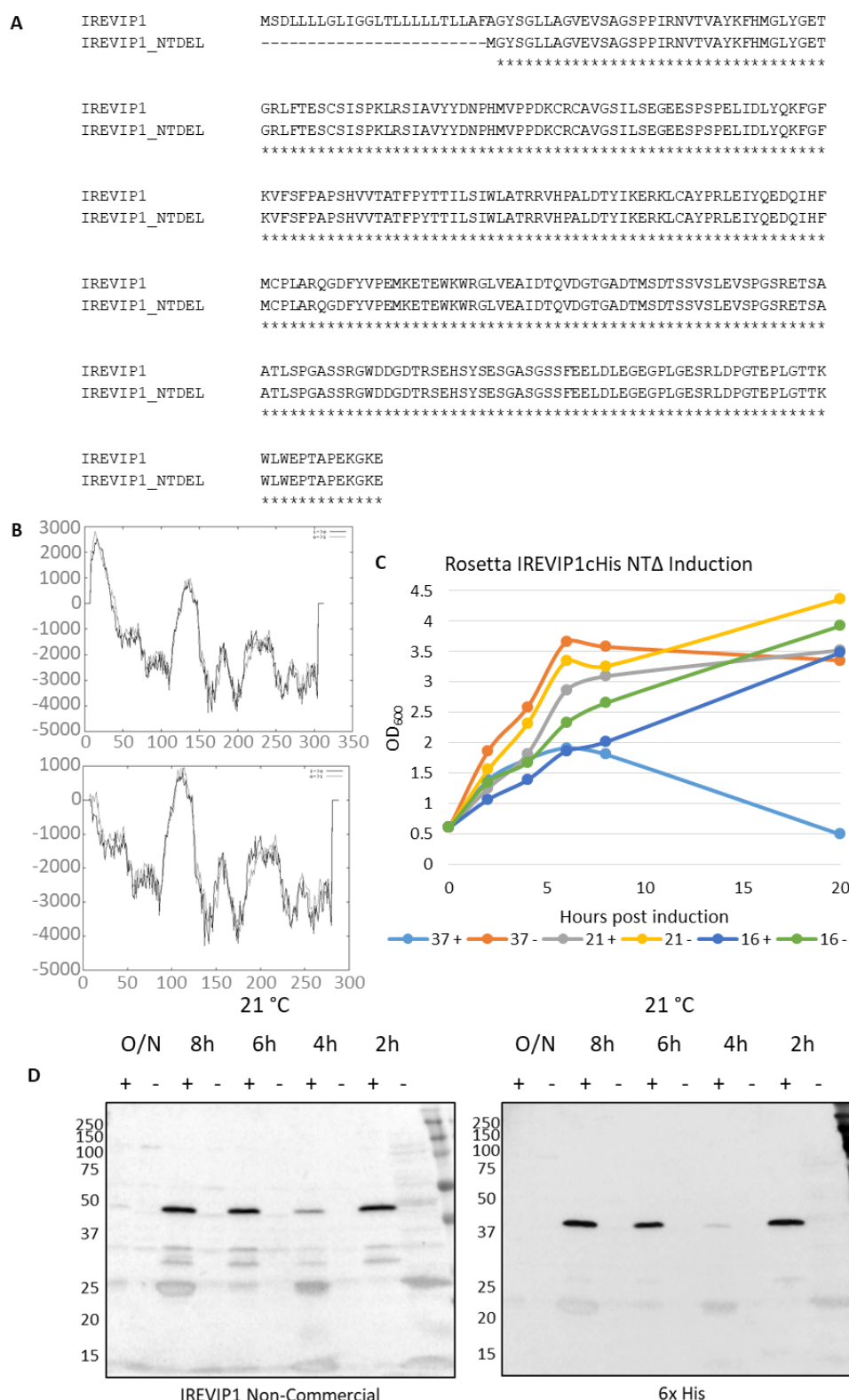


Figure 4-3 IREVIP1 NTA cHis has greater solubility

A) Diagram showing the removal of the first 25 amino acids of IREVIP1. **B)** TMpred hydrophobicity plots of IREVIP1 before (upper panel) and after (lower panel) removal of the hydrophobic N-terminal. **C)** IREVIP1 NTA cHis was expressed in Rosetta *E. coli* cells at 37, 21 and 16°C. Upon collection at 2, 4, 6, 8, and 20 hours, the OD₆₀₀ was read as measure of cell density. This was displayed as a graphical representation and induced and non-induced cultures compared. **D)** IREVIP1 NTA cHis expression was analysed by Western blot at the indicated times by IREVIP1 and 6xHis antibodies.

The expression was also confirmed by WB, using antibodies against IREVIP1 and 6xHis. Samples were loaded relative to the culture OD₆₀₀ at the time of harvesting, showing even expression across the different time point (**Figure 4-3 panel D**). The OD₆₀₀ for the cells grown at 21°C peaked at around 6 h post induction, therefore this time point would yield the greatest protein production (**Figure 4-3 panel C**). IREVIP1 NTΔ cHis showed improved protein expression yields, however purification of IREVIP1 NTΔ cHis using the ÄKTA still proved problematic (**Table 4-1**). A large proportion of IREVIP1 NTΔ cHis was still not binding to the column, and was therefore lost in the flow-through (**Figure 9-2**). Furthermore, before loading samples for size exclusion chromatography, the sample needs to be concentrated. Interestingly, the same amount of IREVIP1 was seen before and after concentration, suggesting that large amounts of protein were lost during concentration (**Figure 9-3 panel B**). IREVIP1 NTΔ cHis also did not behave as expected with either Sephadex 75 (**Figure 9-3**) or Sephadex 200 (**Figure 9-4**) size exclusion columns. In both of these columns, IREVIP1 NTΔ cHis was found to be present in the dead volume, suggesting that it is forming a structure larger than 600 kDa. Once the samples had been run through an ion exchange MonoQ column, the protein yield was additionally reduced (**Figure 9-5**). In order to run a MonoQ column directly after the HisTrap column, desalting columns were needed to reduce the amount of NaCl in the sample from 500 mM to 50 mM, a process which also resulted in a loss of IREVIP1 NTΔ cHis (Data not shown). Furthermore, IREVIP1 NTΔ cHis purification fractions were also analysed by denaturing and native PAGE. Once again IREVIP1 NTΔ cHis was found to enter denaturing gels, but not native gels (**Figure 9-6**). As the using the ÄKTA system was not proving to be successful, the batch method of purification was returned to (**Table 4-1**).

Purification of IREVIP1 NTΔ cHis in batch was more successful, especially after increasing the NaCl concentration to 800 mM and addition of 1% Triton X-100 to the lysis and wash buffers (**Table 4-1**). Triton X-100 was added to remove any hydrophobic interaction with contaminants, including membranes, and to disrupt any microaggregates (Itaya, Brett and Smith, 2012). This

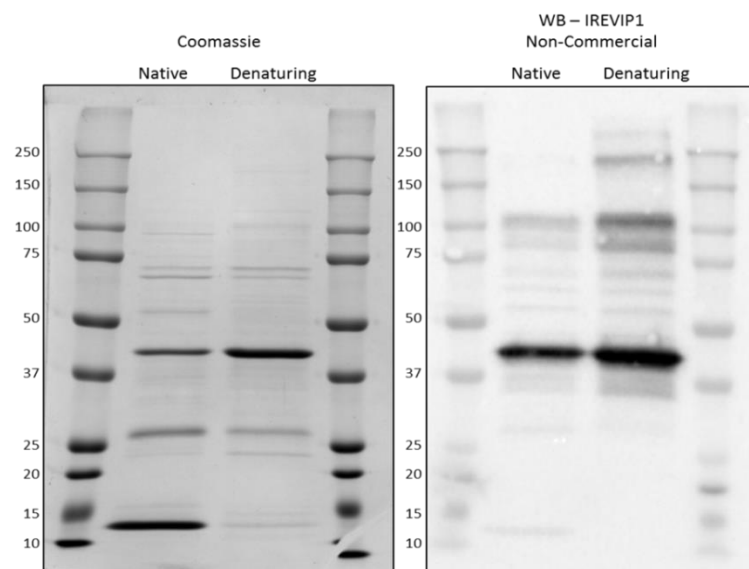


Figure 4-4 IREVIP1 N Δ cHis can be purified in batch under native and denaturing (8 M urea) conditions
 Expression of IREVIP1 was induced in Rosetta *E. coli* using 1 mM IPTG upon reaching an OD₆₀₀ of 0.6, and incubated at 21°C for 6 hours. The bacteria were collected and lysed in either native or denaturing conditions (8M Urea) prior to incubation with NiNTA beads for one hour, followed by washing, elution, dialysis, and concentration. IREVIP1 N Δ cHis was refolded on the beads prior to elution by gradually reducing the urea concentration in the washes.

method produced highly purified IREVIP1 N Δ cHis at the correct size, as seen on denaturing gels, as well as several other bands both larger and smaller, all of which were recognised by a homemade IREVIP1 antibody (**Figure 4-4**).

In order to be more certain that these bands were indeed IREVIP1, a denaturing purification of IREVIP1 N Δ cHis using 8 M urea was performed in batch (**Figure 9-7**). This was carried out also using 800 mM NaCl and 1% Triton X-100. Refolding of IREVIP1 N Δ cHis was carried out on the beads by gradually reducing the concentration of urea in the wash buffer from 8 M to 0 M. Upon analysis by Coomassie and WB, the band pattern of denatured/re-folded IREVIP1 N Δ cHis appeared to be very similar to the native purified IREVIP1 N Δ cHis (**Figure 4-4**); therefore adding confidence to the thought that the other molecular weight bands are indeed IREVIP1 N Δ cHis.

4.1.4 Mutants of IREVIP1 were Purified for Subsequent Biochemical Analysis

To be able to confirm that IREVIP1 is indeed a p97 adaptor, the interaction domain would need to be identified. To investigate this, IREVIP1 NTΔ cHis was used as template for mutagenesis to create a SHP domain deletion mutant (272-285) (SHPΔ), a VBM KERK-AAAA mutant (VBM*), and a double SHPΔ/VBM* mutant (**Figure 4-2 panels D, E, and F, respectively**).

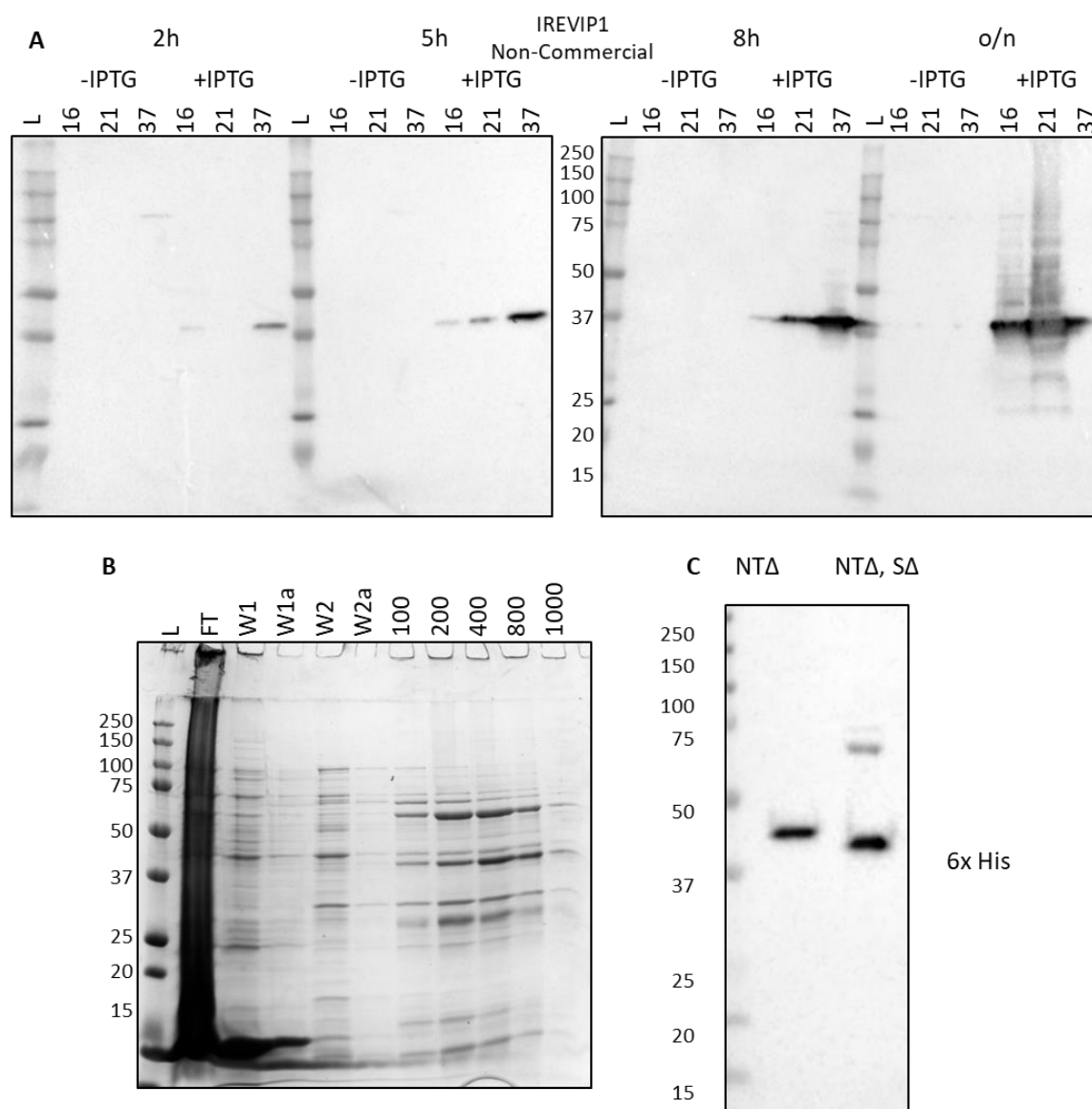


Figure 4-5 Purification of IREVIP1 NTΔ SHPΔ cHis

A) IREVIP1 NTΔ SHPΔ cHis expression was tested at 37, 21, and 16°C for 2, 5, and 8 hours, and overnight. Expression was analysed by WB using a non-commercial IREVIP1 antibody. **B)** Expression of IREVIP1 NTΔ SHPΔ cHis was induced at an OD₆₀₀ of 0.6 using 1 mM IPTG, and incubated at 37°C for 6 hours. The bacteria were then collected and lysed prior to incubating with Ni-NTA beads, followed by washing (1/1a = 5 mM Imidazole, 2/2a = 50 mM Imidazole no Triton X-100), elution (100mM – 1M Imidazole), dialysis, and concentration. Purification fractions

were analysed by Coomassie. C) The concentrated protein was analysed by WB alongside IREVIP1 NTΔ cHis using the indicated antibody.

The SHPΔ mutant was created by mutagenesis using the IREVIP1 NTΔ cHis as a template to delete the predicted SHP domain. Expression was tested in the presence or absence of 1 mM IPTG for 2 h, 5 h, 8 h, or overnight at 37°C, 21°C, and 4°C. Expression was not seen in the absence of IPTG and good expression was seen from 5 h of induction at 37°C. The SHPΔ mutant appeared to be degraded after overnight expression at 37°C and multiple higher and lower molecular weight products were seen at 21°C and 16°C, therefore induction for 6 h at 37°C was chosen (**Figure 4-5 panel A**). The SHPΔ purification was performed as before, with Imidazole and no-Triton X-100, followed by Imidazole elution (See figure legend or details) (**Figure 4-5 panel B**). The elutions were analysed by coomassie prior to pooling and overnight dialysis, followed by protein concentration. Protein concentration was determination by Coomassie against known BSA standards (**Figure 4-5 panel C**). The product was very similar to IREVIP1 NTΔ cHis, although slightly smaller in molecular weight, as shown by WB and Coomassie (**Figure 4-5 panel C, Figure 4-6**). The final SHPΔ product was a similar yield to IREVIP1 NTΔ cHis (**Figure 4-6**). A large upper band of the same size as BSA was seen in the SHPΔ mutant but not in the IREVIP1 NTΔ cHis. During the concentration of the SHPΔ mutant, the centricon centrifugal filter was pre-blocked with BSA, suggesting that this band is likely to be BSA.

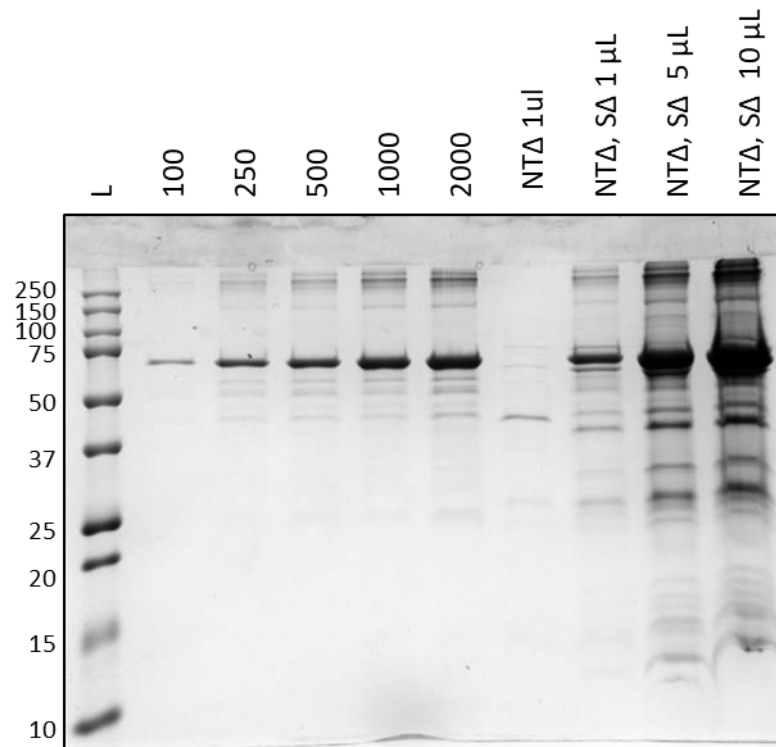


Figure 4-6 IREVIP1 NTA SHPΔ cHis is similar to IREVIP1 NTA cHis

Both proteins were analysed by Coomassie against a known BSA gradient in order to determine protein concentration. Relative protein concentration was determined using ImageJ. IREVIP1 NTA SHPΔ cHis has an additional higher molecular weight band compared to IREVIP1 NTA cHis.

The VBM* mutant was created using two rounds of site-directed mutagenesis. Expression was tested in the presence and absence of IPTG for 2 h, 6 h, or overnight. Expression was seen after 6 h of IPTG induction at all temperatures, with expression at 21°C giving slightly higher expression than at 16°C or 37°C (**Figure 4-7 panel A**). Purification of the VBM* was carried out in batch as previously described. Purification fractions were analysed by Coomassie and WB, where the washes and elutions were tested as a pool, using both IREVIP1 and 6xHis antibodies (**Figure 4-7 panel B and C**). The final VBM* pool was similar to the purified IREVIP1 NTA cHis, in both purity and molecular weight (**Figure 4-9**). BSA was not used to pre-block the centricon membrane during concentration, reflected by the absence of the large upper band seen in the SHPΔ mutant.

The SHPΔVBM* double mutant was created using site-directed mutagenesis from the IREVIP1 SHPΔ cHis construct. Expression was tested in the presence and absence of IPTG for 2 h, 6 h, or

overnight. As with the VBM*, 6 h at 21°C was chosen as the optimal expression condition (**Figure 4-8 panel A**), and purification was carried out as described earlier. Purification fractions were analysed by Coomassie and by WB using both IREVIP1 and 6xHis antibodies, however with pooled washes and elutions (**Figure 4-8 panels B and C**). The SHPΔVBM* double mutant showed a similar level of purity to the other IREVIP1 purified proteins (**Figure 4-9**).

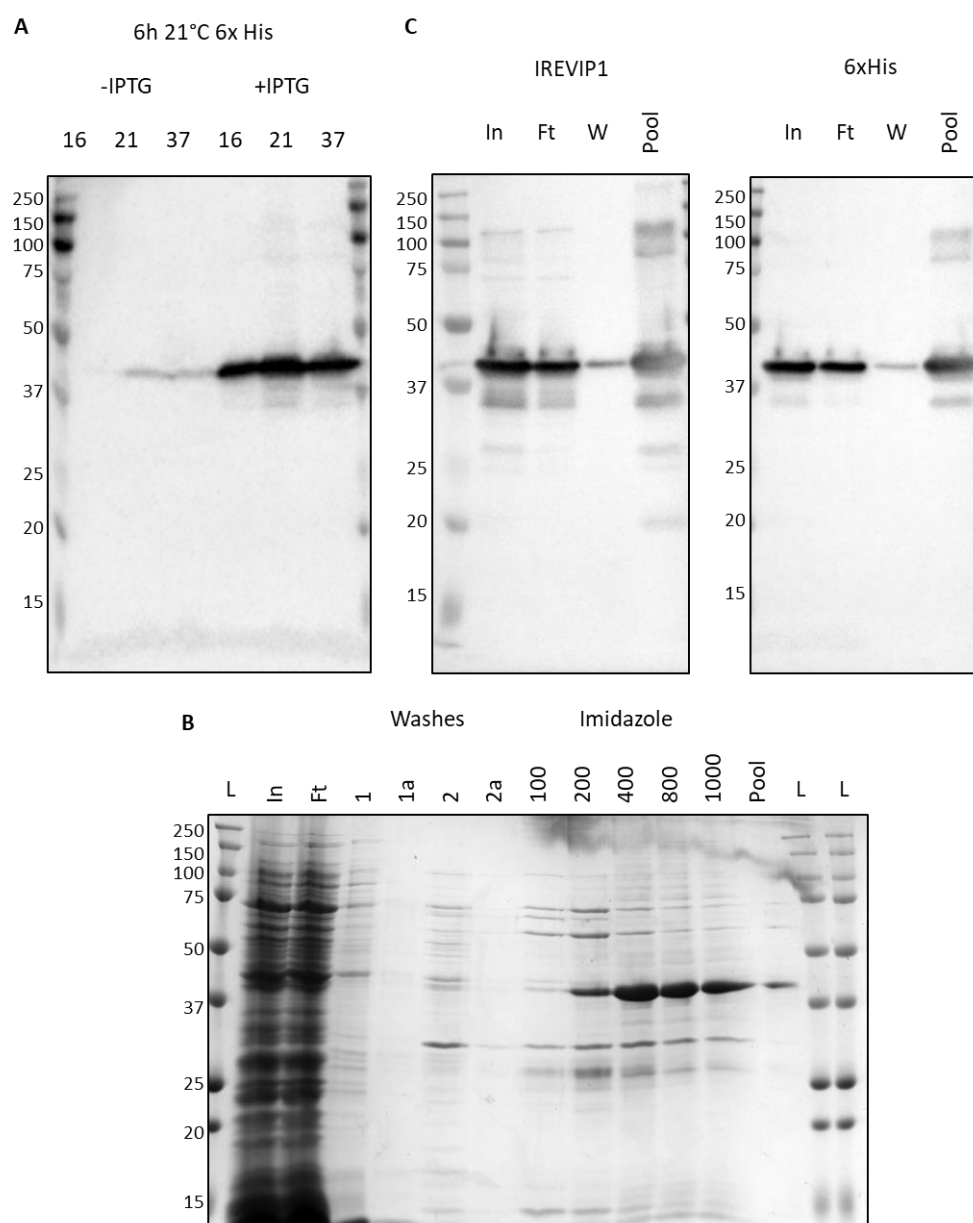


Figure 4-7 Purification of IREVIP1 NTD VBM* cHis

A) IREVIP1 NTD VBM* cHis expression was tested at 37, 21, and 16°C for 2, 5, and 8 hours, and overnight. Expression was analysed by WB using a non-commercial IREVIP1 antibody. Expression was optimal at 21°C for 6 hours. **B)** Expression of IREVIP1 NTD SHPΔ cHis was induced at an OD₆₀₀ of 0.6 using 1 mM IPTG, and incubated at 21°C for 6 hours. The bacteria were then collected and lysed prior to incubating with Ni-NTA beads, followed by washing (1/1a = 5 mM Imidazole, 2/2a = 50 mM Imidazole, no Triton X-100), elution, dialysis, and concentration. Purification fractions were analysed by Coomassie. **C)** The input, flow through, pooled washes and concentrated protein were analysed by WB using non-commercial IREVIP1 and 6xHis antibodies.

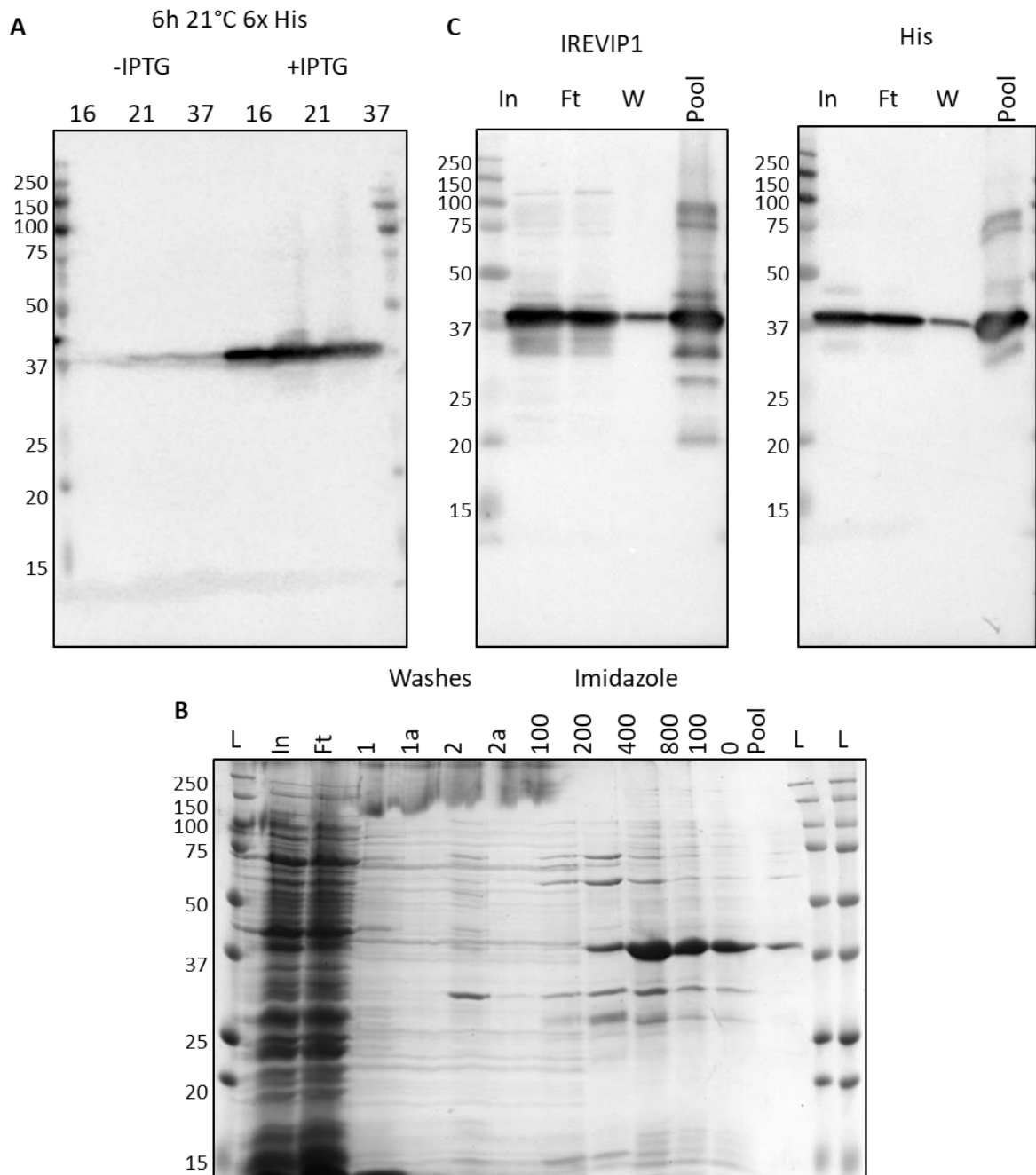


Figure 4-8 Purification of IREVIP1 NΔ SHPΔ VBM* cHis

A) IREVIP1 NΔ SHPΔ VBM* cHis expression was tested at 37, 21, and 16°C for 2, 5, and 8 hours, and overnight. Expression was analysed by WB using a non-commercial IREVIP1 antibody. Expression was optimal at 21°C for 6 hours. **B)** Expression of IREVIP1 NΔ SHPΔ cHis was induced at an OD₆₀₀ of 0.6 using 1 mM IPTG, and incubated at 21°C for 6 hours. The bacteria were then collected and lysed prior to incubating with Ni-NTA beads, followed by washing (1/1a = 5 mM Imidazole, 2/2a = 50 mM Imidazole, no Triton X-100), elution, dialysis, and concentration. Purification fractions were analysed by Coomassie. **C)** The input, flow through, pooled washes and concentrated protein were analysed by WB using non-commercial IREVIP1 and 6xHis antibodies.

All four purified IREVIP1 proteins (NT Δ , SHP Δ , VBM*, and SHP Δ VBM*) were analysed together by Coomassie and by WB using both IREVIP1 and 6xHis antibodies (**Figure 4-9**). When compared to known BSA standards IREVIP1 purified proteins were of similar concentrations, with the SHP Δ mutant appearing to be the lowest. All four proteins were of a similar purity, with only the SHP Δ mutant showing the presumed BSA band (**Figure 4-9 panel A**). When examined by WB using both IREVIP1 and 6xHis antibodies, all four proteins showed a similar banding pattern, the strongest of which were recognised by both IREVIP1 and 6xHis antibodies (**Figure 4-9 panel B**). Both mutants with SHP Δ showed a marked reduction in molecular weight in comparison to the NT Δ construct. The VBM* mutation also appeared to slightly reduce the molecular weight, best seen in the comparison of the SHP Δ and SHP Δ VBM* mutants (**Figure 4-9 panel B**).

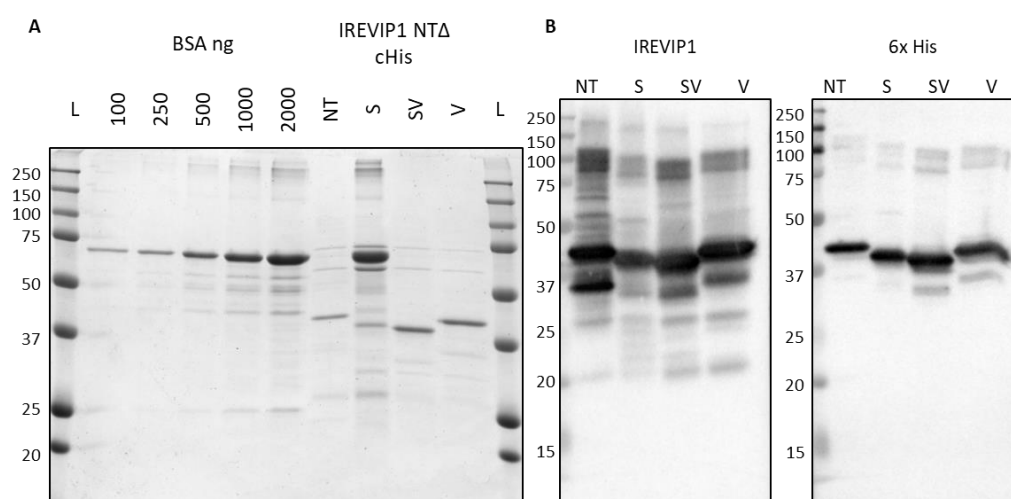


Figure 4-9 Protein quantification and comparison of all four purified IREVIP1 cHis proteins

A) All four proteins were run alongside known BSA standards to determine protein concentration. The Coomassie stained gel was imaged and analysed using ImageJ. **B)** All four purified proteins were analysed by WB using non-commercial IREVIP1 and 6xHis antibodies.

p97 was purified for use in interaction studies with IREVIP1. p97 with a C-terminal S•tag was cloned into the pET21a(+) vector with a stop codon introduced before the 6xHis tag in the vector. The expression of p97cS•tag was tested at 2 h, 5 h, 8 h, and overnight at 16°C, 21°C, and 37°C. Samples were analysed by WB using a p97 antibody and expression was observed from 5 h onwards at 37°C. Overnight induction resulted in the degradation of p97 at 37°C, whereas good expression was seen at 21°C and 16°C. Degradation products were seen at all time points and

temperatures (**Figure 4-10 panel A**). Therefore, induction for 6 h at 37°C was used for p97 expression. Purification was carried out using S-protein agarose (Merck) using the same conditions as IREVIP1, however without imidazole. Washes were carried out using lysis buffer without Triton X-100 for the final two washes. p97 was eluted from the S-protein agarose using citric acid at pH 2 which was immediately neutralised with 1 M Tris pH 11, followed by overnight dialysis and concentration (**Figure 4-10 panel B**). Purification fractions were analysed by Coomassie and protein concentration was determined against known BSA standards (**Figure 4-10 panel C**). Purified p97 was reasonably pure as seen by Coomassie, with lower molecular weight bands presumed to be the p97 degradation products seen during expression testing.

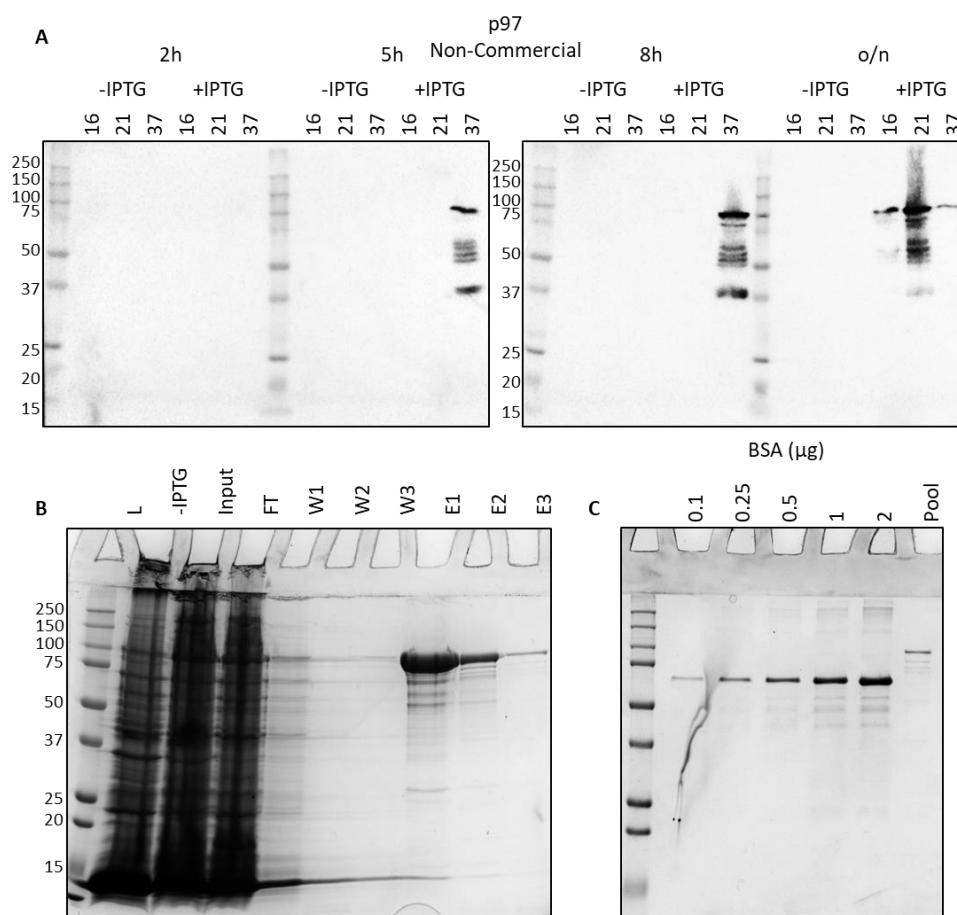


Figure 4-10 Purification of p97cS•Tag

A p97cS•Tag expression was tested at 37, 21, and 16°C for 2, 5, and 8 hours, and overnight. Expression was analysed by WB using a non-commercial p97 antibody. Expression was optimal at 37°C for 6 hours. B) Expression of p97cS•Tag was induced at an OD₆₀₀ of 0.6 using 1 mM IPTG, and incubated at 37°C for 6 hours. The bacteria were then collected and lysed prior to incubating with S-protein agarose, followed by washing (1/2 = non Triton X-100), elution, neutralisation, dialysis, and concentration. B and C) Purification fractions were analysed by Coomassie and quantified against known BSA standards

4.2 Discussion

To investigate whether IREVIP1 is an adaptor of p97, IREVIP1 was purified, along with mutants of the predicted p97 interaction domains. Several different approaches were investigated before soluble IREVIP1 was successfully purified. Initially, IREVIP1nGST was tested in the *E. coli* BL21 (DE3) mutant strain, which enhances membrane protein synthesis, and was found to be insoluble (**Table 4-1**) (Costa *et al.*, 2014). GST is a commonly used solubility enhancer partner, however the mechanism by which fusion partners improve solubility remains elusive (Costa *et al.*, 2014). The GST tag was positioned at the N-terminal of IREVIP1 as this can often increase protein solubility. In comparison to other solubility tags GST performs quite poorly (Costa *et al.*, 2014), with some even stating that it does not enhance solubility (Waugh, 2005). GST fusion proteins are often expressed at high levels in *E. coli*, which can result in protein aggregation. This can be counteracted by growth at lower temperatures, although this reduces overall protein yields (Costa *et al.*, 2014). The protein can be purified from inclusion bodies; however this results in the unfolding of GST and the fusion protein, therefore potentially destroying function. This requires refolding prior to purification; however this may not happen correctly depending on the protein structure (Kimple, Brill and Pasker, 2013). GST is a relative large tag (26 kDa) which may also affect the function of the fusion protein. The relatively large size of GST also puts a metabolic burden on the cell, on top of that of the recombinant protein (Waugh, 2005; Kimple, Brill and Pasker, 2013; Costa *et al.*, 2014). GST is a homodimer, which can introduce complications for purification and makes it incompatible for purification of oligomeric proteins (Waugh, 2005). The affinity purification tag polyhistidine (6xHis) allows purification via immobilised Ni^{2+} , Co^{2+} , Cu^{2+} , Zn^{2+} , Ca^{2+} , or Fe^{3+} ions coupled to beads or resin for Immobilised Metal Affinity Chromatography (IMAC) (Waugh, 2005; Kimple, Brill and Pasker, 2013; Costa *et al.*, 2014). Ni(II)-Nitrilotriacetic acid agarose (Ni-NTA, Qiagen) is a commonly used, cost-effective, resin with high binding capacity (Waugh, 2005). The imidazole ring of each histidine forms a

coordination bond with the immobilised transition metal. Proteins can be eluted by the addition of free imidazole (Bornhorst and Falke, 2000). The 6xHis tag is a small, low metabolic burden, N- or C-terminal tag, making it less likely to interfere with its fusion protein; therefore correct folding is more likely and proteins are rarely targeted to inclusion bodies (Waugh, 2005; Kimple, Brill and Pasker, 2013; Costa *et al.*, 2014). Furthermore, protein expression in *E. coli* using the 6xHis tag can often produce a reasonably pure protein, up to 95%, after just one purification step (Bornhorst and Falke, 2000; Kimple, Brill and Pasker, 2013). Thus, IREVIP1cHis was tested, therefore avoiding the hydrophobic N-terminal domain and increasing the chances of correct folding and solubility. Furthermore, the *E. coli* BL21 DE3 strain derivative, Rosetta, was used to allow for potential rare codon tRNA usage (Costa *et al.*, 2014). A shortage of these rare codon tRNAs can result in translational stalling and premature translation termination, therefore resulting in poor protein expression (Sørensen, Sperling-Petersen and Mortensen, 2003). Another step taken to improve solubility, stability, and purity was by increasing the NaCl concentration from 150 mM to 500 mM (Nordlund *et al.*, 2008). These changes improved solubility and purity slightly, however, batch purification capture efficiency was still low and elution concentrations poor.

To try and improve this, the ÄKTA FPLC system was tested using a 5 mL HisTrap FF column, as this provided a larger protein binding capacity. Unfortunately, this did not improve the proportion of IREVIP1cHis captured (**Table 4-1**). This could suggest that some of the IREVIP1cHis proteins expressed have their tag obscured, either by aggregation or tertiary structure, therefore preventing binding to the beads (Crowe, Masone and Ribbe, 1995; Kneusel *et al.*, 1998). IREVIP1cHis purification fractions were run on native and denaturing gels to determine the molecular state of the purified IREVIP1. IREVIP1cHis was able to enter denaturing SDS-PAGE gels, however, IREVIP1cHis was unable to enter native PAGE gels (**Figure 9-1**). The stacking and separating gels are pH 6.8 and pH 8.8 respectively, therefore IREVIP1cHis, with a pI of 4.7, was negatively charged; therefore IREVIP1cHis should run through the native gel (Costa *et al.*, 2014).

This suggested that IREVIP1cHis was forming some kind of high molecular weight structure, whether this was naturally occurring oligomers or undesirable protein aggregates. Therefore, the properties of IREVIP1 itself were looked at in an attempt to improve purification.

IREVIP1 has a highly hydrophobic N-terminal domain (**Figure 4-3**). It is known that single pass membrane proteins will often form large aggregates which can subsequently be removed by centrifugation, however microaggregates and oligomers can remain (Itaya, Brett and Smith, 2012). Therefore, IREVIP1 was re-cloned without the first 25 aa, then on referred to as IREVIP1 NTΔ cHis, as this region had a TMPred score of 2816, where scores over 500 are considered significant (Hofmann and Stoffel, 1993). This region did not overlap with any of the identified domains of interest, therefore removing it was not expected to have any effect on biochemical interaction studies. After expression testing, it was found that induction of IREVIP1 NTΔ cHis with IPTG was most successful when cells were grown at 21°C for 6 h (**Figure 4-3**). However, upon purification using the ÄKTA system a large proportion of IREVIP1 NTΔ cHis was still not binding to the column, despite the binding capacity almost certainly not being reached (>40 mg/mL media, GE Healthcare) (**Figure 9-2**). In addition to affinity purification, size exclusion and ion exchange chromatography were used. Prior to size exclusion, the protein sample requires concentration, therefore a centricon centrifugal filter unit with a cut off of 10 kDa was used. This allows everything which is larger than 10 kDa to be concentrated without also concentrating any buffers. Conversely, analysis of IREVIP1 by SDS-PAGE before and after concentration did not appear to show an increase in the amount of protein loaded, suggesting that protein has been lost (**Figure 9-3 panel B**). This could either be from binding to the centricon membrane, or the formation of aggregates upon concentration, which are then removed during centrifugation prior to column loading. Additionally, IREVIP1 NTΔ cHis was found in the dead volume in both Superdex 75 and Superdex 200 columns (**Figure 9-3, Figure 9-4**). Superdex 75 size exclusion columns should separate everything between 3 and 70 kDa, with everything larger than 70 kDa passing through the column first in the 'dead volume' (**GE Healthcare**). The same was true of

Superdex 200 size exclusion columns, however, their range spans 10 – 600 kDa (**GE Healthcare**). Thus, size exclusion chromatography can give a view of the aggregation state of a protein, adding weight to the hypothesis that IREVIP1 is forming high molecular weight structures (Itaya, Brett and Smith, 2012). The loss of protein therefore meant that there were very low levels of protein left after Ion Exchange Chromatography (**Figure 9-5**). This could also suggest that the high molecular weight IREVIP1 NTΔ cHis was somehow being lost within the ÄKTA system itself. In further support of the high molecular weight hypothesis, IREVIP1 NTΔ cHis purification fractions were still unable to enter native gels (**Figure 9-6**).

To circumvent this problem, batch purification was returned to. Histidine is an infrequently naturally used amino acid (<2 %), but some endogenous proteins may contain two or more in a row, therefore allowing them to co-purify with the 6xHis tagged protein. However, this problem is more evident when purifying from mammalian system than from *E. coli* (Bornhorst and Falke, 2000). To try and reduce the high molecular weight forms of IREVIP1 NTΔ cHis and to further remove contaminants, the NaCl concentration was increase to 800 mM and 1% triton X-100 was added to the buffers, as these can often disrupt microaggregates and reduce non-specific binding (Bornhorst and Falke, 2000; Itaya, Brett and Smith, 2012). Batch purification with these altered conditions allowed the purification of IREVIP1 NTΔ cHis at a measurable concentration (**Figure 4-4**). Some other bands were also present alongside the purified form of IREVIP1 NTΔ cHis but were recognised by both IREVIP1 and 6xHis antibodies, suggesting that they were other forms of IREVIP1 NTΔ cHis (**Figure 4-4**).

To further confirm this, IREVIP1 NTΔ cHis was purified under denaturing conditions to remove all non-specific interactions with the resin and IREVIP1 (**Figure 4-4, Figure 9-7**). IREVIP1 NTΔ cHis was refolded whilst still bound to the resin, after washing, by gradually reducing the urea concentration (Bornhorst and Falke, 2000). Analysis of this, in comparison to IREVIP1 NTΔ cHis purified under native condition, by both Coomassie and WB using IREVIP1 and 6xHis antibodies,

suggested that the bands observed at molecular weights other than that predicted for IREVIP1, are indeed IREVIP1. The banding pattern between the two, using all forms of identification, was highly similar.

IREVIP1 was predicted to have two known p97 interacting domains; a VBM domain and a SHP domain (**Figure 3-2**). IREVIP1 mutants for either domain were therefore purified, along with a double mutant. All IREVIP1 mutants were produced from the IREVIP1 NTΔ cHis construct, therefore, in addition to the domain mutations, they were also lacking the N-terminal region. The IREVIP1 SHP mutant (SHPΔ) was performed by deleting the amino acids 272-284 spanning the predicted SHP domain. The VBM* mutant was performed by two rounds of site-directed mutagenesis to mutate K160A, E161A, R162A, and K163A, either in the IREVIP1 NTΔ cHis backbone (VBM*), or the derivative IREVIP1 SHPΔ cHis backbone (SHPΔVBM*). All mutants were purified using the same final procedure as IREVIP1 NTΔ cHis. The banding patterns for all of the mutants were very similar to that of IREVIP1 NTΔ cHis as analysed by WB and Coomassie, suggesting that the proteins were behaving in a similar manner (**Figure 4-9**). The main difference was that in the SHPΔ purification there was a higher molecular weight band at around the same size as BSA which was not identified by either the IREVIP1 or 6xHis antibodies. During the purification of the SHPΔ mutants, the centricon centrifugal filter unit was pre-blocked with BSA to try and reduce the binding of IREVIP1 to the membrane and improve concentration times. This could therefore mean that this band is BSA, as it is also a similar size. This should not, however, be a problem for any downstream interaction studies as the beads used are also pre-blocked with BSA prior to usage. All four IREVIP1 constructs showed higher molecular weight bands which were identified by both IREVIP1 and 6xHis antibodies. As they are present in all constructs, it is unlikely that they are forming due to structural changes due to domain mutations. Therefore, it is likely that they are either part of a natural oligomerisation process, which would be interesting to investigate in cells, or are aggregates formed as a result of bacterial expression and purification. As the proteins are used in this potentially aggregated or

oligomerised state in all *in vitro* assays and various protein-protein interactions were observed it would be interesting to speculate that this is the natural state of the protein. This oligomerisation could be a mechanism by which to sequester IREVIP1 in the cell until it is needed, which can then be reversed upon the presence of an interacting protein.

The main distinction between the different purified proteins was the size of the band corresponding to the predicted molecular weight of IREVIP1. IREVIP1 has been reported to have a predicted size of 34.189 kDa (Uniprot (Bateman *et al.*, 2017)), however it is observed as running between 37 and 50 kDa (**Figure 4-9**). The SHPΔ mutant runs at a smaller size than IREVIP1 NTA cHis, running just above 37 kDa. The VMB* mutant appeared to be slightly smaller as well, running just higher than the SHPΔ mutant. The VBM* further reduced the size of the SHPΔ in the double mutant. This could be due to the deletion of 12 aa to produce the SHPΔ mutant, which could impact upon the protein structure (Carroll and Setlow, 2005). However, deletion of the 25 aa N-terminal domain did not seem to have any impact upon the molecular weight, in comparison to full length IREVIP1, as observed by SDS-PAGE, suggesting that the N-terminal domain does not play a role in the folding of the rest of the protein. One way in which this could be checked is by using circular dichroism (CD) spectroscopy, which is a frequently used method to discern secondary structure (Itaya, Brett and Smith, 2012).

Additionally, p97cS●Tag was purified for use in later *in vitro* studies. The S-peptide (aa 1-20) and S-protein (aa 21-124) are two enzymatically inactive parts of the enzymatically active ribonuclease S (Kim and Raines, 2008). The two parts come together to form a tight association, therefore minimising the binding of off target molecules (Raines *et al.*, 2000). Furthermore the S●tag is small and uncharged thus unlikely to interfere with protein folding or function (Raines *et al.*, 2000). Purified p97 ran at the predicted size and was of high purity, and therefore could be used in interaction studies with IREVIP1 and mutants. Therefore, the next chapter will cover the biochemical investigation of the IREVIP1-p97 interaction.

5. RESULTS 3: *IN VITRO* CHARACTERISATION OF IREVIP1

p97 is a key component of the Ubiquitin Proteasome System (UPS), which, with a hierarchical network of adaptor proteins (cofactors), regulates a multitude of processes within the cell (Hänzelmann and Schindelin, 2011; Hänzelmann, Buchberger and Schindelin, 2011; Meyer and Wehl, 2014). To carry out this function, and to confer substrate specificity, p97 requires different adaptor proteins for different processes, of which at least 30 have been identified thus far (Meyer and Wehl, 2014). Adaptor proteins can bind either to the N- or C- terminal of p97 via a wide range of p97 binding motifs (UBX, PUB, SHP, VIM, and VBM) (Yeung *et al.*, 2008; Meyer, Bug and Bremer, 2012). Three main mutually exclusive core complexes, p97- UFD1-NPL4, p97-p47, and p97-UBXD1 (Hänzelmann, Buchberger and Schindelin, 2011; Vaz, Halder and Ramadan, 2013), potentially induce a conformational change in p97 which then allows for the subsequent binding of additional cofactors which can add another layer of control and could potentially influence subcellular localisation and substrate specificity (Hänzelmann, Buchberger and Schindelin, 2011). IREVIP1 had been identified as being in complex with p97 *in vivo* by MS (**Figure 3-6**) and Co-IP (**Figure 3-5**) and was predicted to have both SHP and VBM p97 interacting domains (**Figure 3-2**). Therefore, it was tested *in vitro* whether IREVIP1 and p97 could interact directly, and whether the interaction domain could be identified. Furthermore, interaction with Topoisomerase I and II α was tested as potential substrates.

5.1 IREVIP1 Interacts with p97 *in vitro*

To investigate the possibility that IREVIP1 is a new p97 adaptor, their direct interaction was tested *in vitro*. S-protein agarose was blocked overnight at 4°C with 5% BSA. The following day the beads were washed twice and resuspended in 137 mM NaCl, 20 mM Tris HCl pH 7.0 (low salt buffer). p97cS•tag was then incubated with the beads at 4°C for two hours, followed by three five minute washes. IREVIP1 N Δ cHis was then added to either blank beads or p97cS-tag coated beads and was incubated for two hours at 4°C. Beads were then washed three times for five

minutes each in low salt buffer with 0.5% Triton X-100. The beads were eluted and samples were then analysed by WB alongside input samples using both 6xHis (IREVIP1) and p97 antibodies. p97cS-tag was seen to bind well to the S-protein agarose. IREVIP1 was observed to bind marginally to the blank S-protein agarose, but was seen to bind much more to the p97 coated beads (**Figure 5-1 panel A**), thus confirming that IREVIP1 and p97 interact directly *in vitro*.

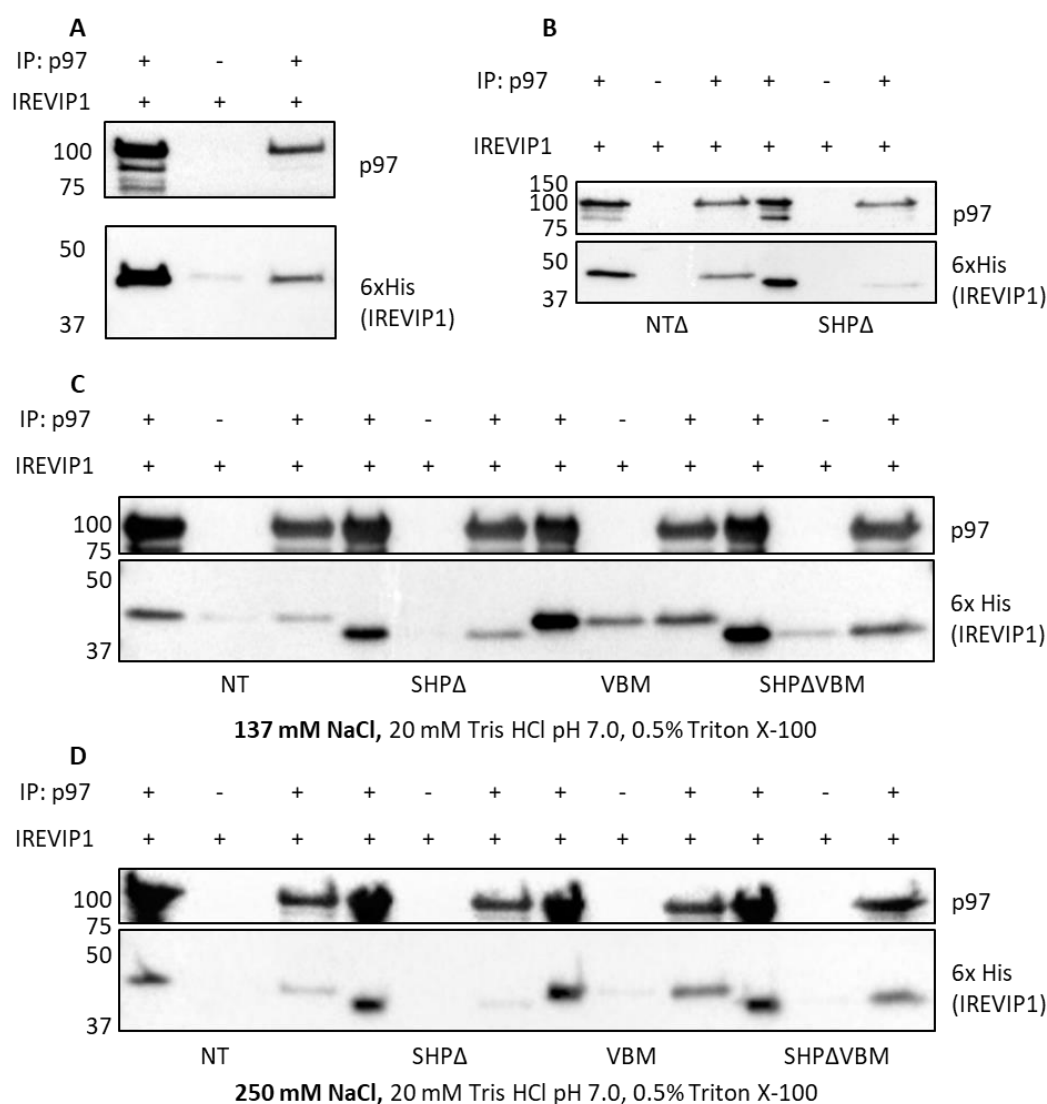


Figure 5-1 IREVIP1 interacts with p97 via its SHP domain in a salt dependent manner
S-protein agarose was pre-blocked with 5% BSA overnight before incubating with or without p97cS•Tag. Blank beads or p97-coated beads were then incubated with various IREVIP1 cHis purified proteins and eluted by boiling. Samples were analysed by WB with the indicated antibodies. A) IREVIP1 NTΔ cHis was observed to bind p97 beads to a greater extent than blank beads. B) SHPΔ prevents IREVIP1 from binding to p97. C) IREVIP1 mutants appear to have greater binding to p97 at 137 mM salt, however VBM* and SHPΔVBM* mutants had high background. D) Increasing the salt concentration to 200 mM abolished binding of the SHPΔ mutant to p97, but not the VBM* or SHPΔVBM* mutants. Inputs represent 100% of p97 incubated with the beads, and 20% of IREVIP1.

5.2 IREVIP1 Binds p97 via its SHP Domain

To assess whether either of the predicted p97 binding domains are responsible for the interaction between IREVIP1 and p97, the purified IREVIP1 mutants (SHPΔ, VBM*, and SHPΔVBM*) were tested. p97cS●tag was incubated with pre-blocked S-protein agarose, washed, and incubated with either IREVIP1 NTΔ cHis, SHPΔ, VBM*, or SHPΔVBM*. The beads were then washed with low or high (250 mM NaCl) salt buffer with 0.5% Triton X-100. IREVIP1 NTΔ cHis bound well to p97 when washed in both low and high NaCl, with high NaCl washes showing slightly reduced binding, however this also reduced background binding (**Figure 5-1 panels C and D**). Upon washing with low salt buffer, it was observed that all of the IREVIP1 mutants bound more strongly to p97 than IREVIP1 NTΔ cHis, with the VBM* and SHPΔVBM* mutants also showing high background levels (**Figure 5-1 panel C**). However, upon using high salt washes, the binding of the SHPΔ mutant to p97 was strongly abolished, whereas the VBM* and SHPΔVBM* were still strongly bound to p97, despite the background levels being abolished (**Figure 5-1 panel D**). This indicates that the VBM* mutation is causing IREVIP1 to be sticky and have an increased affinity for p97. This could explain why the SHPΔVBM* mutant is still binding to p97. Therefore, these results suggest that the direct interaction of IREVIP1 and p97 is influenced by the SHP domain, however the involvement of the VBM domain cannot be ruled out with these experiments, especially as some residual binding remains with the SHPΔ mutant.

5.3 IREVIP1 Interacts with Topoisomerase IIα and Topoisomerase I

As IREVIP1 has a predicted GyrI-like domain, which is homologous to the bacterial DNA gyrase inhibitor SbmC, it was hypothesised that IREVIP1 could be interacting directly with mammalian Topoisomerases. As this interaction had previously been seen *in vivo*, it was tested *in vitro*. To this end, IREVIP1 NTΔ cHis was incubated with HisPur™ Ni-NTA Magnetic Beads (Thermo Scientific), in low salt buffer, at 4°C for two hours, followed by low salt washes. IREVIP1 coated beads were then incubated with either Topoisomerase IIα or Topoisomerase I (Inspiralis)

followed by low salt washes with 0.5% Triton X-100. Proteins were eluted and analysed by WB using 6xHis (IREVIP1) and Topoisomerase II α or Topoisomerase I antibodies respectively. IREVIP1 bound well to the beads with no binding observed to the blank beads (**Figure 5-2 panels A and B**). Topoisomerase II α bound to the blank beads to some degree, however, this was greatly increased with IREVIP1 coated beads (**Figure 5-2 panel A**). Background binding of Topoisomerase I to the beads was not observed, and Topoisomerase I was seen to bind highly to IREVIP1 coated beads (**Figure 5-2 panel B**). These results, therefore, suggest that both Topoisomerase II α and Topoisomerase I both interact with IREVIP1 directly *in vitro*.

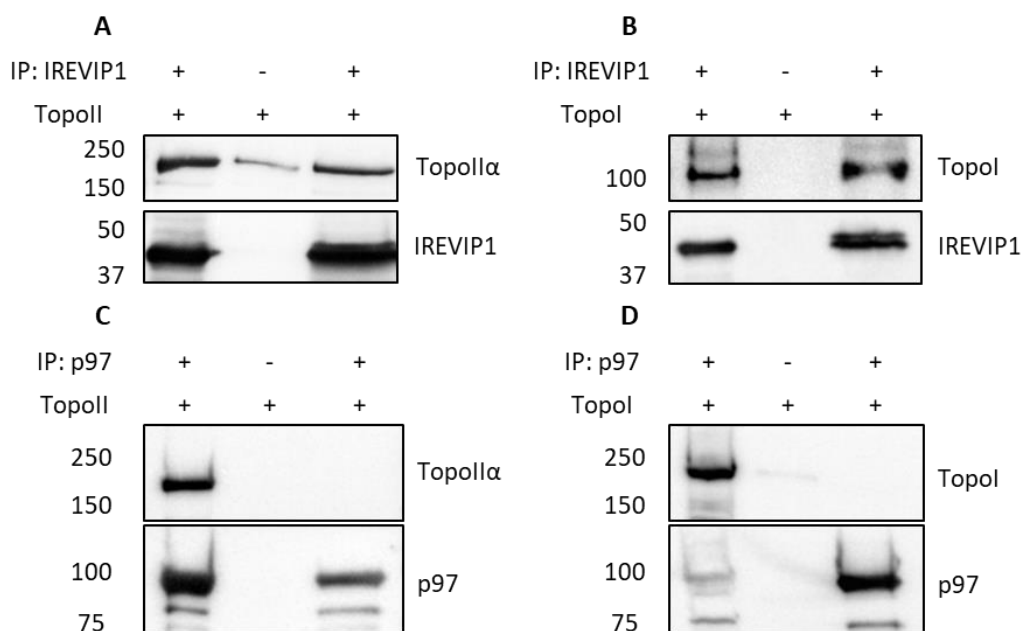


Figure 5-2 IREVIP1, but not p97, interacts with Topoisomerase I and II α
IREVIP1 cHis was coupled to Ni-NTA beads for two hours, followed by washing, and incubation of either A) Topoisomerase II α with IREVIP1 coated or blank beads, or B) Topoisomerase I with IREVIP1 coated or blank beads, for two hours. Beads were eluted by boiling and analysed by WB using the indicated antibodies. p97 cS•Tag was coupled to S-protein agarose for two hours, followed by washing, and incubation of either C) Topoisomerase II α with p97 coated or blank beads, or D) Topoisomerase I with p97 coated or blank beads, for two hours. Proteins were eluted by boiling and analysed by WB using the indicated antibodies. Inputs represent A) 100% of IREVIP1 and 20% of Topoisomerase II α incubated, B) 100% of IREVIP1 and 40% of Topoisomerase I incubated, C) 100% of p97 and 20% of Topoisomerase II α incubated, and D) 100% of p97 and 40% of Topoisomerase I incubated.

5.4 p97 is Unable to bind Topoisomerase II α or Topoisomerase I Directly

As IREVIP1 was shown to bind p97 *in vitro* via its SHP domain, and is therefore likely an adaptor protein, it was hypothesised that IREVIP1 is able to bridge p97 and Topoisomerases in an

adaptor-like fashion. To further strengthen this idea, it was important to determine whether p97 can bind directly to either of the two Topoisomerases. Therefore, it was assessed whether p97 can bind either Topoisomerase directly in the absence of adaptors. p97cS-tag was coupled to S-protein agarose, washed, and incubated with either Topoisomerase II α or Topoisomerase I. The beads were then washed, eluted, and analysed by WB using specific antibodies for each protein. Slight background binding was observed for Topoisomerase I, however, neither Topoisomerase II α nor Topoisomerase I were capable of binding p97 coated beads (**Figure 5-2 panels C and D**). Together, these data suggest that p97 cannot bind Topoisomerases directly.

5.5 p97 Requires IREVIP1 to Bridge the Interaction with Topoisomerase II α and Topoisomerase I

Having ruled out that neither Topoisomerase can bind directly to p97 *in vitro*, yet both displayed a strong interaction with IREVIP1, it was exciting to speculate that the presence of IREVIP1 would enable the formation of a complex containing p97, IREVIP1 and either Topoisomerase.

To see whether IREVIP1 could bridge Topoisomerases to p97-coated beads, an *in vitro* triple interaction experiment was carried out. p97cS-tag was coupled to S-protein agarose, followed by washes. IREVIP1 and either Topoisomerase II α /Topoisomerase I were added to blank beads or p97 coated beads, or Topoisomerase II α /Topoisomerase I was added to p97 coated beads in the absence of IREVIP1. The proteins were incubated before washing, elution and analysis by WB using p97, 6xHis (IREVIP1) and Topoisomerase II α or Topoisomerase I antibodies. When washing the beads with low salt buffer, p97 was clearly binding to the beads (**Figure 5-3 all panels**). IREVIP1 was binding to the p97 coated beads and only a slight background binding to the blank beads could be detected (**Figure 5-3 all panels**). Topoisomerase II α was observed binding to the blank beads, and this increased marginally when incubated with p97 coated beads in the absence of IREVIP1. Excitingly, in the presence of both IREVIP1 and p97, much more Topoisomerase II α was pulled down in the reaction, suggesting that IREVIP1 is able to bridge the

interaction between p97 and Topoisomerase II α (**Figure 5-3 panel A**). However, upon washing with high salt buffer, the increased interaction of Topoisomerase II α was not observed (**Figure 5-3 panel B**). Similar results could be observed when the experiment was performed with Topoisomerase I. When washing with low salt buffer, Topoisomerase I had some residual binding to p97 in the absence of IREVIP1, however this interaction was greatly increased upon addition of IREVIP1, therefore suggesting that IREVIP1 can bridge the interaction between p97 and Topoisomerase I (**Figure 5-3 panel C**). Furthermore, this increased interaction was abolished when washing with high salt buffer (**Figure 5-3 panel D**). This reduced interaction at high salt could be due to loss of ionic interactions between the Topoisomerases and IREVIP1. However, the background interaction between Topoisomerases and p97 is not lost at high salt (**Figure 5-3 panels B and D**). This could be due to non-specific interactions between Topoisomerases and p97 occurring to a greater extent than when p97 is coated with IREVIP1 and potentially blocked from interacting non-specifically with Topoisomerases. Therefore, it would be interesting to carry out this pull down reciprocally.

Together, these data demonstrate that, *in vitro*, IREVIP1 is capable of facilitating the bridging of p97 and Topoisomerases.

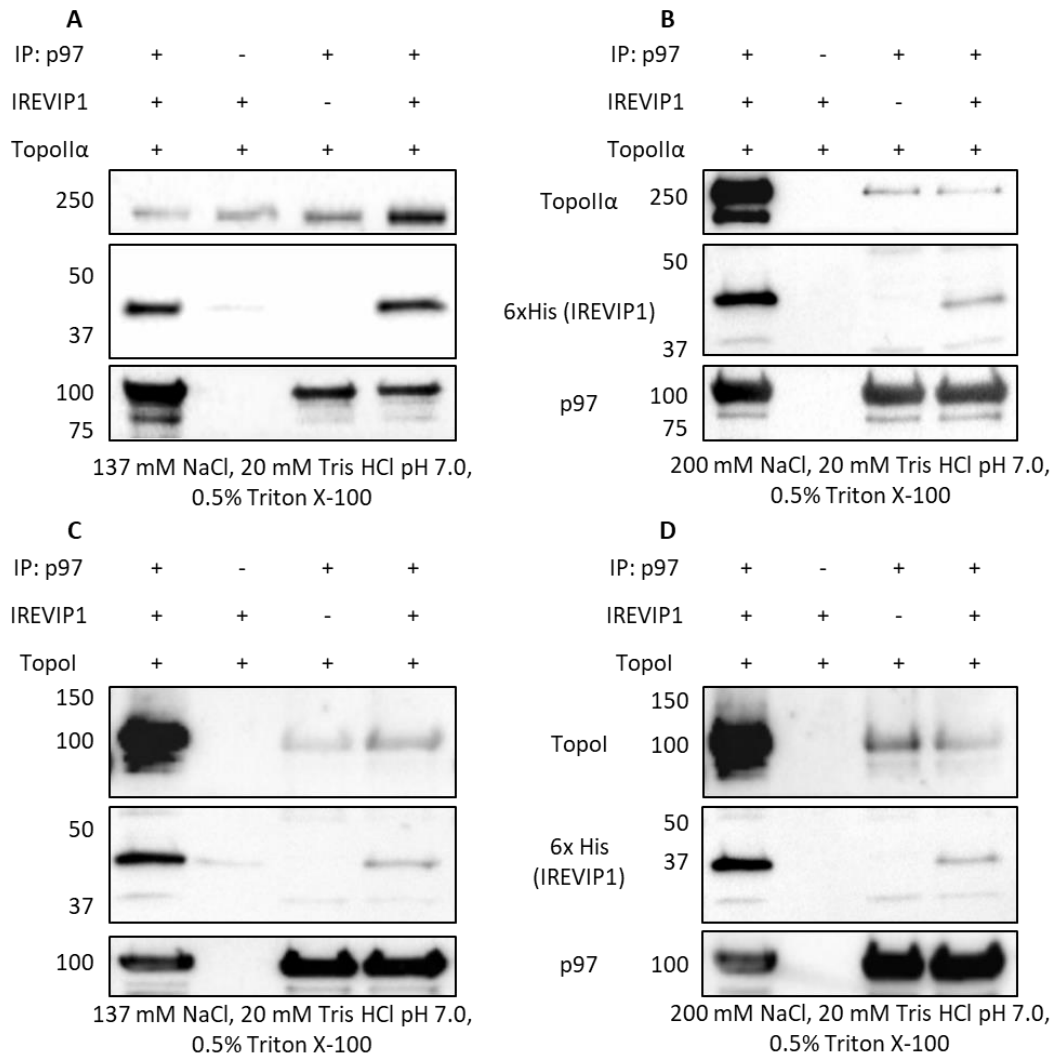


Figure 5-3 IREVIP1 is required to bridge the interaction between p97 and Topoisomerase I or IIα
p97cS•tag was coupled to BSA blocked S-protein agarose for two hours. A) Either Topoisomerase IIα and IREVIP1 were incubated with blank beads, or p97 coated beads, or Topoisomerase IIα was incubated with p97 coated beads in the absence of IREVIP1 in buffer containing 137 mM NaCl or B) 200 mM NaCl. C) Either Topoisomerase I and IREVIP1 were incubated with blank beads, or p97 coated beads, or Topoisomerase I was incubated with p97 coated beads in the absence of IREVIP1 in buffer containing 137 mM NaCl or D) 200 mM NaCl. Inputs represent A, B) 50% of p97, 20% of IREVIP1, 50% of Topoisomerase IIα incubated, C, D) 50% of p97, 20% of IREVIP1, 30% of Topoisomerase I incubated.

5.6 IREVIP1 is Capable of Binding Both Double Stranded and Single Stranded DNA

Topoisomerases bind DNA, therefore the ability of IREVIP1 to also bind DNA was tested by an electromobility shift assay (EMSA). IREVIP1 was pre-incubated with either double stranded (DS) or single stranded (SS) fluorescently labelled probes prior to running on a 0.5% TBE agarose gel for analysis. The binding sequence for SP1, a well-described transcription factor (Dyran and Tjian, 1983), was used as the DNA probe, and purified SP1 was used as a positive control (**Figure 5-4**). IREVIP1 was able to bind both the DS and SS probe to a greater extent than BSA, therefore suggesting that IREVIP1 was able to bind DNA (**Figure 5-4 panel A**). It would also be interesting to see whether the IREVIP1 SHPΔ and VBM* mutants are still capable to bind DNA. This experiment was performed using IREVIP1 purified under native conditions, which could contain co-purified contaminants. Hence, to rule out that these contaminants might be responsible for the binding, IREVIP1 purified under denaturing conditions was also tested. The denatured and refolded IREVIP1 was able to bind DNA (**Figure 5-4 panel B**) supporting the idea of IREVIP1 being capable of binding directly to DNA. Furthermore, this result also suggests that the denatured/refolded IREVIP1 has refolded in the correct conformation.

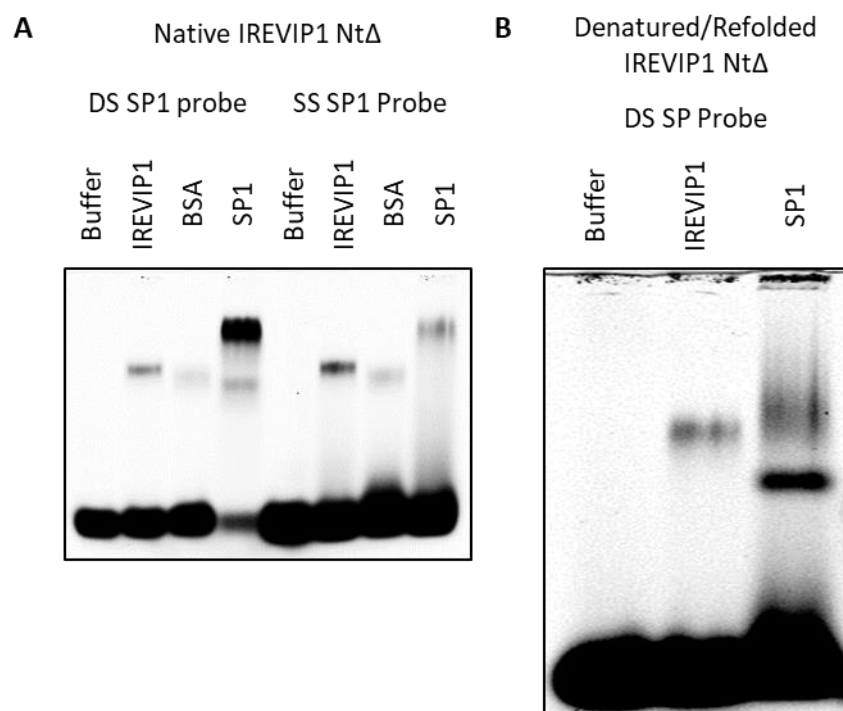


Figure 5-4 IREVIP1 is capable of binding single stranded and double stranded DNA

IREVIP1, storage buffer, or BSA in storage buffer was pre-incubated with either dsDNA or ssDNA labelled probes for 30 minutes at 37°C. Samples were run on a 0.5% agarose gel at 4°C and imaged using the LI-COR Odyssey Imaging system. Purified SP1 was used as a control as the SP1 consensus binding sequence was used as the probe. IREVIP1 purified under both native (A) and denaturing (B) conditions was able to bind to both dsDNA and ssDNA probes.

5.7 Discussion

IREVIP1 was identified *in vivo* during a p97 MS, and potential p97 interacting domains were identified (**Figure 3-1, Figure 3-2**). Therefore, this interaction was tested *in vitro* using purified IREVIP1 and p97, as well as IREVIP1 p97 interaction domain mutants to identify the interaction motif (**Figure 5-1**). As IREVIP1 was first identified in complex with an ATPase deficient mutant of p97, it would be interesting to test whether the binding of IREVIP1 to p97 is affected by the ATP-bound state of p97. This could be achieved by using non-hydrolysable forms of ATP and assessing the binding of IREVIP1 and p97. This could further be a means to explain why wild-type p97 was pulled down poorly by IREVIP1cSSH and would be of interest to test with non-hydrolysable ATP in cells (**Figure 3-5**). IREVIP1 and p97 were found to interact *in vitro*, and the SHP domain was observed to be required for p97 binding, therefore suggesting that IREVIP1 is indeed a new p97 adaptor protein.

IREVIP1 possesses a Gyrl-like domain with homology to the bacterial protein Gyrl, which in *E. coli* is responsible for DNA gyrase inhibition (Nakanishi *et al.*, 1998). Furthermore, as IREVIP1 has been shown to interact with both Topoisomerase I and Topoisomerase II α *in vivo* by MS and Co-IP, it was hypothesised that IREVIP1 might interact directly with them. Pull-down experiments with purified proteins confirmed a direct interaction between IREVIP1 and both Topoisomerase I and Topoisomerase II α (**Figure 5-2 panels A and B**), but not between p97 and tested Topoisomerases (**Figure 5-2 panels C and D**). However, p97 was able to pull down both Topoisomerase I and II α in the presence of IREVIP1, suggesting that IREVIP1 is an adaptor that targets p97 to Topoisomerases and most probably to Topo-ccs (**Figure 5-3**).

In addition to a direct interaction with p97 and Topoisomerases, IREVIP1 also binds both dsDNA and ssDNA (**Figure 5-4**). The GC content of the probes used was quite high (72.73 %), with two runs of GGGG. Therefore, these probes could potentially be forming secondary DNA structures, such as G-quadruplexes (Bochman, Paeschke and Zakian, 2012; Yadav *et al.*, 2014). These structures are commonly found at promoter regions. Active transcription through GC rich regions can result in the formation of R-loops, where the transcribed RNA forms a hybrid with the template DNA strand (Yadav *et al.*, 2014). This R-loop formation can also cause the formation of G-quadruplexes in the non-template strand, where this combination is referred to as a G-loop. (Yadav *et al.*, 2014). Furthermore, the negative supercoiling introduced behind replication forks can promote the formation of R-loops if it is not relieved by Topoisomerase I (Massé and Drolet, 1999). R-loops and G-loops can cause a barrier to replication forks and therefore induce replication stress arising from the collision of the replication fork with the R-loop, resulting in genomic instability (Tuduri *et al.*, 2009; Yadav *et al.*, 2014). Therefore, recognition of these structures by IREVIP1 could aid in its recruitment to Topo-ccs located in regions enriched with R-loops. However, this speculative hypothesis must be further explored.

p97 binds the heterodimer UFD1-NPL4 in a 6:1 ratio via their SHP and UBX domains respectively. It is as yet unclear whether these two domains bind the same p97 N-terminal domain, or adjacent N-terminal domains (Hänzelmann and Schindelin, 2016; Le *et al.*, 2016). This would therefore leave the remaining N-terminal domains available for further cofactors to bind, therefore allowing another level of regulation which is consistent with hierarchical cofactor binding (Hänzelmann and Schindelin, 2011, 2016; Hänzelmann, Buchberger and Schindelin, 2011). The p97 adaptor SPRTN was identified as being the missing link for the processing of Topo-ccs prior to TDP1 or TDP2 degrading the remaining peptide (Vaz *et al.*, 2016). However, as SPRTN is also involved in the degradation of other DNA bound proteins (Vaz *et al.*, 2016), how is SPRTN specifically targeted to Topo-ccs? Both SPRTN and IREVIP1 bind to p97 via their SHP domains and could therefore bind different p97 N-terminal domains (Davis *et al.*, 2012; Mosbech *et al.*, 2012; Stingle *et al.*, 2016; Ramadan *et al.*, 2017). That would then allow IREVIP1 to be involved in the specific recruitment of p97 or even p97-SPRTN complexes to Topo-ccs. However, it is so far unknown whether IREVIP1 has any direct proteolytic or regulatory function of its own independently of SPRTN, such as SPRTN's involvement in TLS (Davis *et al.*, 2012; Mosbech *et al.*, 2012) and DPC processing (Vaz *et al.*, 2016).

Therefore, in light of these results, it was then investigated whether IREVIP1 plays a role in maintaining genomic stability in unchallenged conditions and in response to exogenous DNA damage.

6. RESULTS 4: IREVIP1 AND GENOME STABILITY

p97 has previously been implicated in chromatin associated degradation (CAD) in order to allow progression of DNA repair and replication (Acs *et al.*, 2011; Franz *et al.*, 2011; Meerang *et al.*, 2011; Raman *et al.*, 2011; Davis *et al.*, 2012; Mosbech *et al.*, 2012; Stingle *et al.*, 2016; van den Boom *et al.*, 2016; Vaz *et al.*, 2016; Maric *et al.*, 2017). This is often carried out in response to a myriad of ubiquitin signals with the help of one of the core adaptor complexes; UFD1-NPL4 (Vaz, Halder and Ramadan, 2013). IREVIP1 was previously shown here to dynamically interact with p97 in response to IR (**Figure 3-1**), and was confirmed to interact directly with p97 via its SHP domain (**Figure 5-1**). IREVIP1 was also found to be in complex with key DNA repair and replication proteins (**Figure 3-5**), and was shown to be able to bridge the interaction between p97 and Topoisomerases *in vitro* (**Figure 5-3**). In light of this, it was hypothesised that IREVIP1 might be involved in maintaining genomic stability via the regulation of Topoisomerases dealing with DNA replication and transcription related DNA topology. Therefore, the role of, and reliance on, IREVIP1 in these processes was investigated.

6.1 IREVIP1 Depletion Causes Activation of the DNA Damage Response Signalling Pathway

HEK293 cells were either depleted using non-silencing siRNA or two different siRNA sequences targeting IREVIP1 for 72 hrs. Cells were then collected, lysed, and analysed by Western blot using antibodies against key DDR markers (**Figure 6-1**). The IREVIP1 antibody revealed good knockdown with either siRNA sequence. Several key DDR markers were activated upon IREVIP1 depletion in the absence of exogenous damage. Histone H2AX was seen to have greatly increased levels of phosphorylation (γ H2AX) upon IREVIP1 depletion compared to siNS conditions. Both CHK1 and CHK2 had increased levels of phosphorylation upon IREVIP1 depletion in comparison to siNS conditions, with total CHK2 levels remaining stable. However, total CHK1 protein levels were reduced upon IREVIP1 depletion. Both ATM and ATR also showed

increased levels of phosphorylation upon IREVIP1 depletion (**Figure 6-1**). IREVIP1 is therefore important for genome stability by preventing endogenous DNA damage.

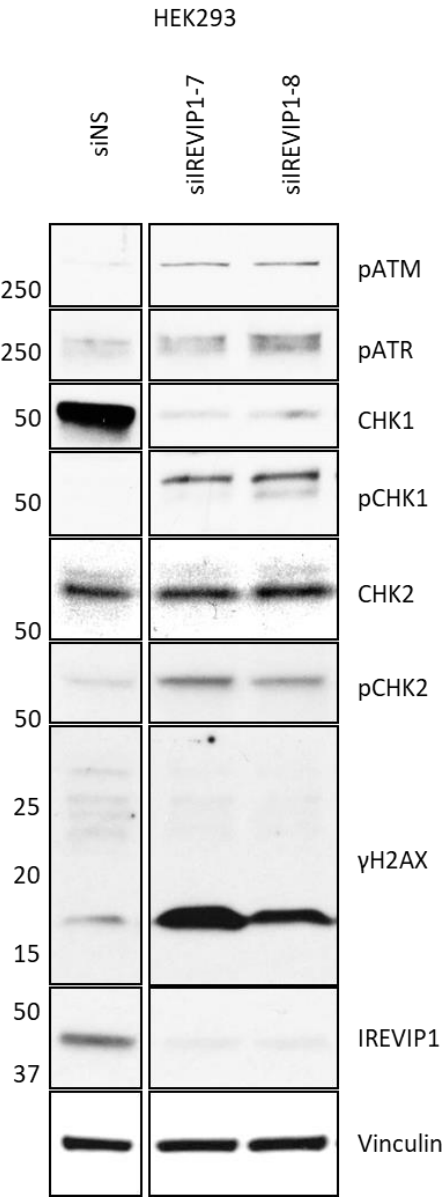


Figure 6-1 IREVIP1 depletion causes activation of the DNA damage response
HEK293 cells were depleted for 72 hours using non-silencing siRNA (luciferase) and two different siIREVIP1 siRNA sequences. Cells were collected and lysed using 1x RIPA buffer with added protease and phosphatase inhibitors followed by sonication. Samples were then analysed by WB using indicated antibodies. IREVIP1 depletion was observed to cause activation of DDR proteins.

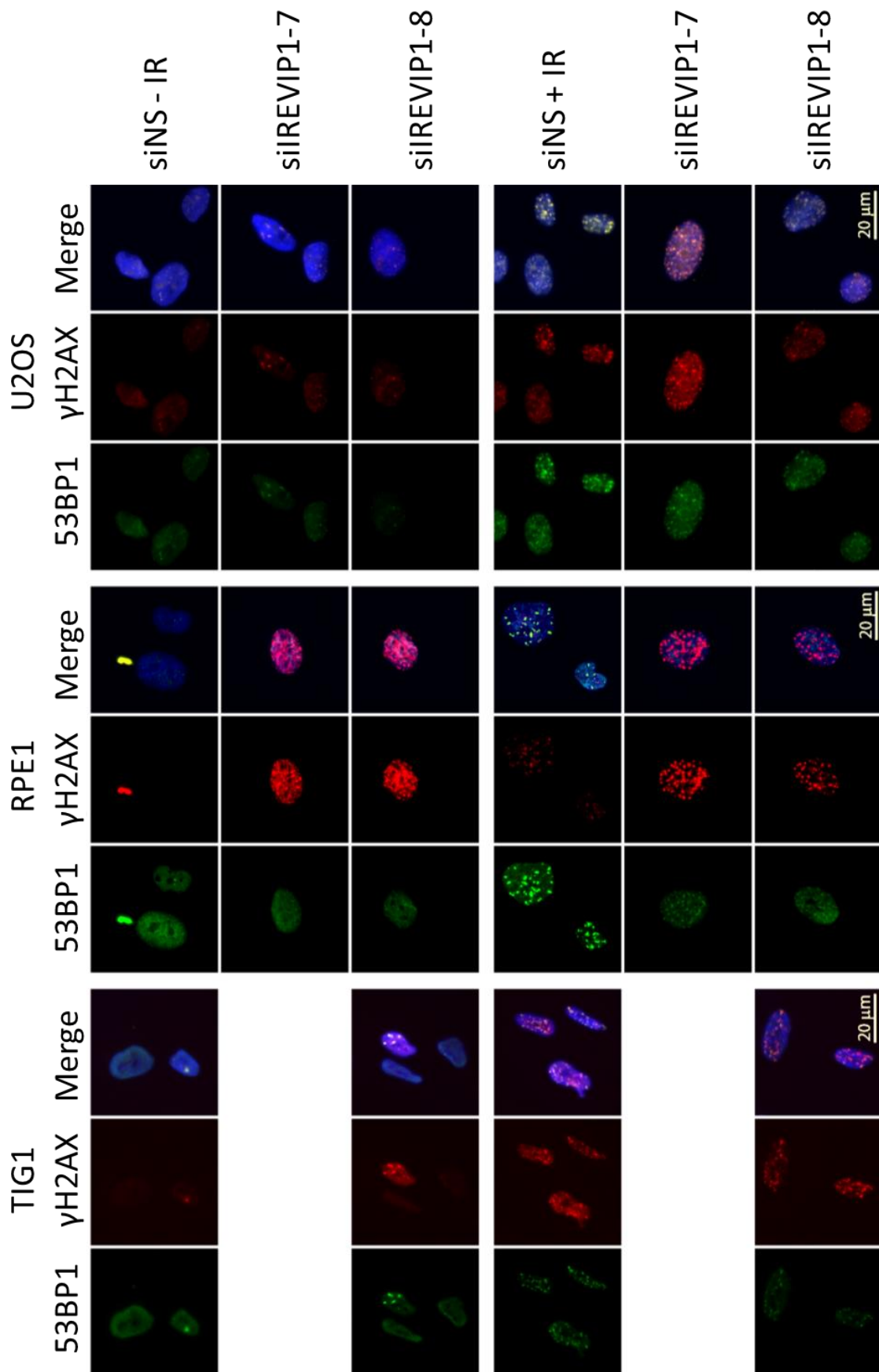
6.2 IREVIP1 Depletion Results in DNA Damage, Causing Histone H2AX Phosphorylation, but Reduces 53BP1 Recruitment, Both With and Without Exogenous DNA Damage

Three different cell lines (TIG1 primary fibroblasts, RPE1 retinal epithelial cells, and U2OS osteosarcoma cells) were depleted for 72 hrs using either two different IREVIP1 siRNA sequences or control siRNA. Cells were then either treated with 2 Gy of IR and left to recover for two hours, or left untreated. Cells were then fixed, permeabilised and blocked with 5% BSA, followed by immuno staining for γ H2AX, 53BP1, and DNA (DAPI), prior to imaging (**Figure 6-2 panel A**). The number of foci per cell were then counted using ImageJ and plotted as the average number of foci with error bars of the SEM from three replicates (**Figure 6-2 panel B**). The siRNA depletions were analysed by WB using a homemade IREVIP1 antibody (**Figure 6-2 panel C**).

TIG1 cells were unable to be imaged upon depletion with the IREVIP1-7 sequence as very few cells survived. With the IREVIP1-8 sequence in non-irradiated condition, an increase in γ H2AX foci was seen, however a reduction in 53BP1 foci formation was observed. After IR, the number γ H2AX foci were unchanged in comparison to siNS treated cells, however 53BP1 foci were reduced in number (**Figure 6-2 panels A and B**).

RPE1 cells showed an increase in γ H2AX foci number in the absence of exogenous damage upon IREVIP1 depletion with either sequence. However, no increase was seen in the number of 53BP1 foci with siIREVIP1-7 in comparison to siNS, and only a small decrease in 53BP1 was seen with siIREVIP1-8. After treatment with IR, both IREVIP1 siRNA sequences showed an increase in the number γ H2AX foci and a decrease in 53BP1 foci number compared to siNS (**Figure 6-2 panels A and B**).

A



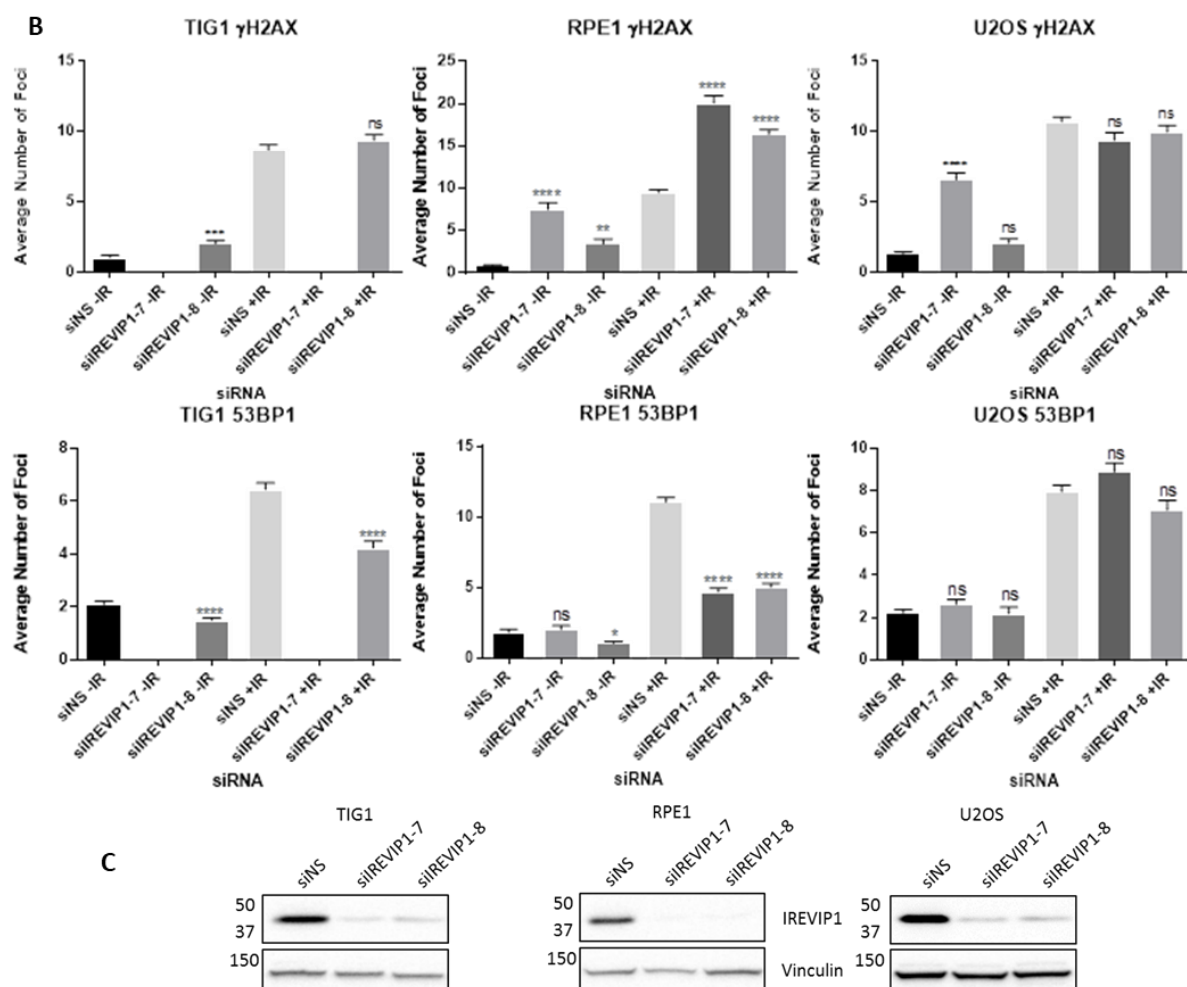


Figure 6-2 IREVIP1 depletion causes varied DNA damage response in a cell type dependent manner

Either TIG1, RPE1, or U2OS cells seeded on to cover glasses were depleted for 72 hours using either non-silencing siRNA (luciferase) or two different IREVIP1 siRNA sequences. A) Cells were fixed and stained on the cover glasses using the indicated antibodies prior to microscopy analysis. B) At least 100 cells were counted per condition, and the number of foci per cell were counted via a consistent automated system using ImageJ. The average number of foci per cell was plotted with the SEM using GraphPad Prism, and the statistical significance calculated using a one-way ANOVA. * <0.05, ** <0.01, *** <0.001, **** 0.0001, n=3. C) WBs for IREVIP1 depletion. IREVIP1 depletion induces the DDR with varied 53BP1 recruitment.

U2OS cells, in the absence of exogenous damage, displayed an increased number of γ H2AX foci with siIREVIP1-7, but not with siIREVIP1-8. Neither IREVIP1 siRNA sequence showed a difference in the number of 53BP1 foci compared to siNS. After treatment with IR, neither γ H2AX nor 53BP1 showed a difference in foci number upon IREVIP1 depletion compared to siNS. However, the background level of γ H2AX and 53BP1 were generally higher in this cell line (**Figure 6-2 panels A and B**).

In summary, IREVIP1 depletion appears to result in DNA damage in the absence of exogenous damage, resulting in Histone H2AX phosphorylation, but in some cases not 53BP1 recruitment. In response to exogenous damage, such as IR, H2AX is phosphorylated, yet often 53BP1 recruitment is reduced. However, in the cancerous cell line U2OS, the DNA damage response is largely similar to siNS cells

6.3 IREVIP1 Depletion Causes Cell Cycle Defects

TIG1, RPE1, HEK293 and U2OS cells were treated with either non-silencing or two different IREVIP1 siRNAs for 72 hrs. Cells were then collected, fixed and permeabilised, followed by EdU, DNA (PI) and phosphorylated Histone H3 (Serine 10) staining. Samples were then analysed using the BD FACScalibur and FlowJo software. The percentage of cells in each cell cycle phase was assessed and plotted with error bars of the SEM from three replicates (**Figure 6-3 upper panel**). The EdU incorporation in siIREVIP1 cells was plotted as a percentage relative to siNS cells with error bars of the SEM from three replicates (**Figure 6-3 lower panel**). Both TIG1 and RPE1 cells were arrested in G1 and G2/M, showing a reduction in both EdU positive S-phase and mitotic cells. HEK293 cells showed mild G1 arrest with a reduction in EdU S-phase cells. U2OS cells were arrested in G2 with a reduction in G1 cells; however, the number of EdU S-phase cells remained similar. All cell lines showed a reduction in total EdU incorporation relative to non-silenced cells upon IREVIP1 depletion (**Figure 6-3 lower panel**).

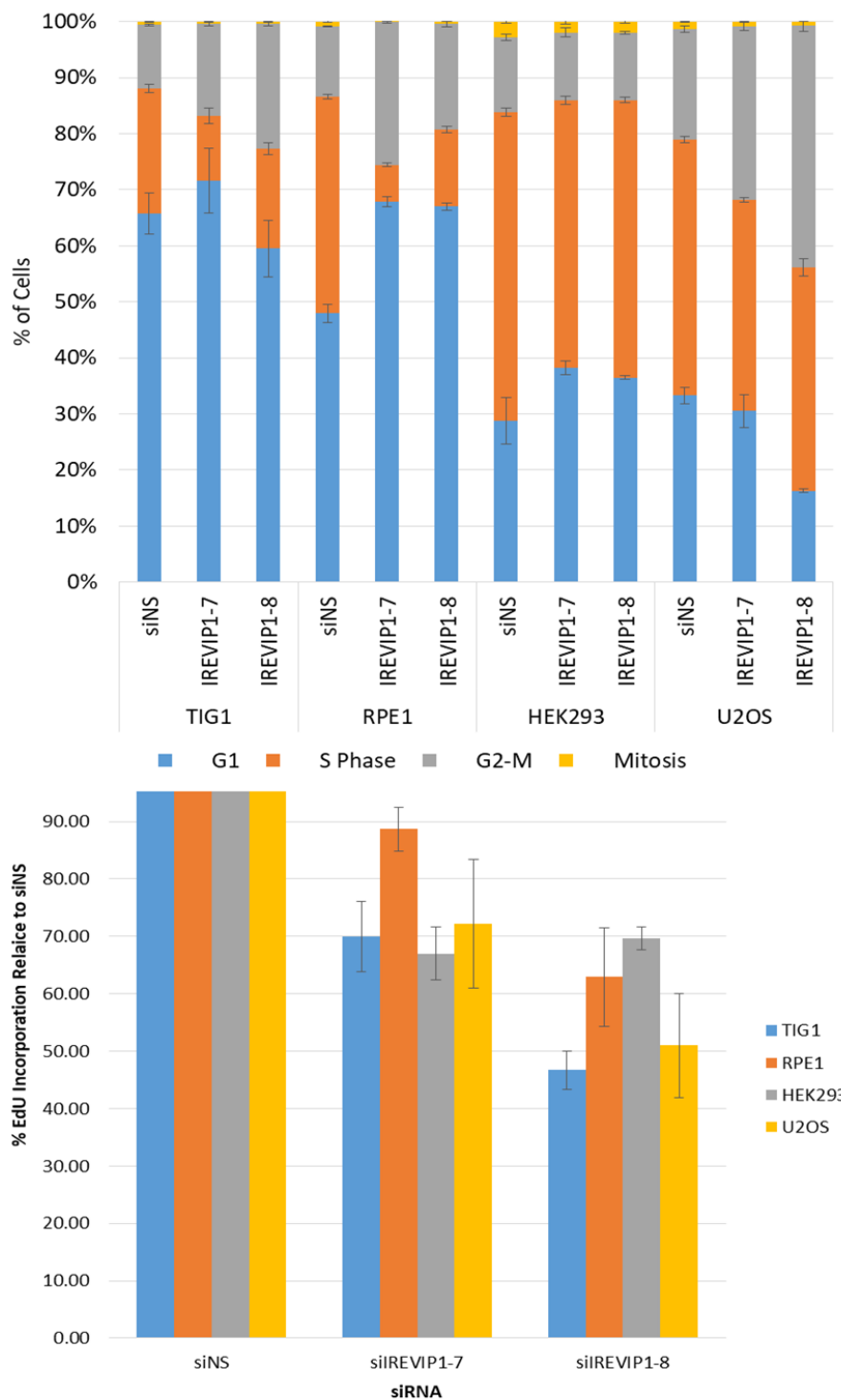


Figure 6-3 IREVIP1 is required for normal cell cycle progression

Either TIG1, RPE1, HEK293, or U2OS cells were depleted for 72 hours with non-silencing siRNA (luciferase) or two different IREVIP1 siRNA sequences. Cells were incubated with EdU for 30 minutes prior to collection. Cells were subsequently fixed, and permeabilised prior to staining with Alexa 647 by click chemistry (EdU, S phase), 488 Histone H3Ser10 antibody (Mitosis), and propidium iodide (DNA content). Cell cycle profiles were analysed using FlowJo. The percentage of cells in each cell cycle phase (upper panel) and the EdU incorporation relative to siNS cells (lower panel) was plotted with the SEM using excel, n=3. IREVIP1 depletion causes cell cycle arrest and reduces EdU incorporation.

6.4 IREVIP1 is a Part of the Replication Fork

As the above result indicate that absence of IREVIP1 induces replication stress, it was of interest to understand if IREVIP1 is part of the replication fork. To address this hypothesis, isolation of proteins on nascent DNA (iPOND) was performed.

HEK293 cells were incubated with or without EdU for 10 minutes followed by either collection or incubation with thymidine for a further 10 minutes to allow the replication machinery to progress past the region of EdU incorporation (**Figure 6-4 panel B**). Proteins were then crosslinked using formaldehyde, collected, and permeabilised. Click reaction was then carried out to couple biotin to the incorporated EdU. Subsequently, cells were lysed and a streptavidin Co-IP was carried out. Samples were analysed by WB using the desired antibodies. Polymerase δ , MCM3 and PCNA, as well as p97, SPRTN and IREVIP1, were observed after pull down in the presence of EdU, however they were not seen after thymidine chase (**Figure 6-4 panel A**). This indicates, that IREVIP1 is part of the replication fork. Upon IREVIP1 siRNA depletion for 72 hrs, an increase in Topoisomerase I was observed at the replication fork (**Figure 6-4 panel C**) indicating that IREVIP1 plays a role in avoidance of Topoisomerase accumulation at the replication fork.

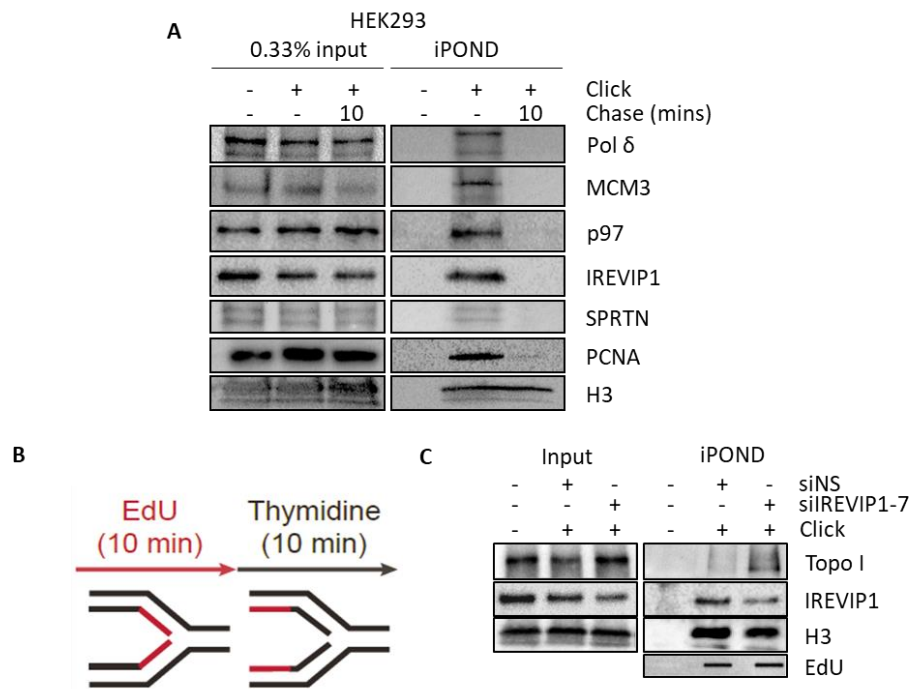


Figure 6-4 IREVIP1 is part of the replication fork and prevents Topoisomerase I accumulation

A) EdU was incorporated into HEK293 cells for 10 minutes, followed by 10 minutes of thymidine incorporation for the chase samples. Cells were collected and lysed, followed by chromatin fragmentation. A streptavidin pull down was performed to capture biotin-coupled EdU. Samples were analysed by WB using the indicated antibodies. IREVIP1 was found to be present at the replication fork. **B)** iPOND 'pulse – chase' experimental outline. **C)** HEK293 cells were depleted for 72 hours using non-silencing siRNA (luciferase) or IREVIP1-7 siRNA. iPOND was performed as described in A. IREVIP1 depletion was seen to result in the accumulation of Topoisomerase I on DNA at the replication fork. Performed by J Fielden

6.5 IREVIP1 Depletion Results in Accumulation of Topoisomerase I and II α on DNA

Topoisomerases are well known to become crosslink to DNA leading to the formation of Topo-ccs. p97 has been implicated in the removal of Topoisomerase I-ccs (Lin *et al.*, 2009; Lee *et al.*, 2016). As IREVIP1 can bind to both p97 and Topoisomerase I, it was exciting to speculate that IREVIP1 might play a role in the removal of Topoisomerase I-ccs. To address this, DNA-protein crosslinks by the RADAR method were extracted and analysed.

HEK293 or HEK293 IREVIP1 Δ CRISPR clones #11 and #12 were analysed for DPCs. Cells were lysed under harsh denaturing conditions in buffer containing 6 M guanidium thiocyanate in order to remove all non-covalently attached protein from the DNA. The DNA was precipitated using 100% ethanol and solubilised using NaOH, followed by quantification using Nanodrop. Proteins were then analysed by slot blot with antibodies against Topoisomerase I, or

Topoisomerase II α . dsDNA was used as a loading control for isolation of same amount of DNA. Control cells showed minimal accumulation of covalently attached Topoisomerase I and Topoisomerase II α on DNA (Topo1-ccs, Topo2-ccs), however in either IREVIP1 Δ CRISPR clones both Topoisomerase I and Topoisomerase II α are cross-linked to DNA (**Figure 6-5**).

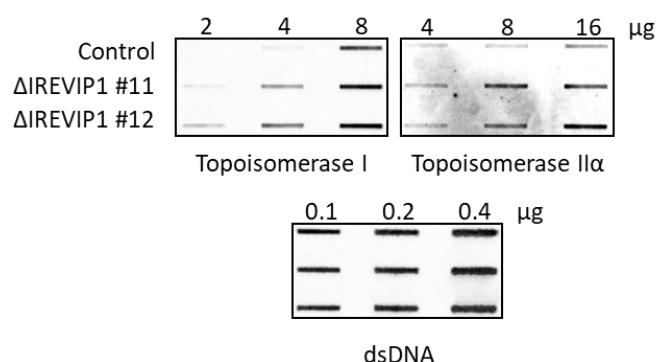


Figure 6-5 Topoisomerase I and Topoisomerase II α DPCs accumulate in the absence of IREVIP1
HEK293 WT or HEK293 IREVIP1 Δ CRISPR cells were lysed in stringent denaturing lysis buffer followed by DNA precipitation and solubilisation. Samples were analysed by slot blot using the indicated antibodies. dsDNA was used as a loading control. Cells lacking IREVIP1 accumulated a greater amount of Topoisomerase I and Topoisomerase II α crosslinked to DNA. Performed by J Fielden

6.6 IREVIP1 is required for efficient replication fork progression

As an absence of IREVIP1 was seen to cause an increase in the number of Topoisomerase-ccs (**Figure 6-5**), and that DPCs can cause replication fork stalling (Duxin *et al.*, 2014), it was speculated that IREVIP1 might be important for the unperturbed progression of the DNA replication fork. Therefore, this was assessed by the DNA fibre assay in cells containing or depleted for IREVIP1, and also in the absence of Topoisomerase I, Topoisomerase II α , or both Topoisomerases.

HEK293 cells were depleted for 72 hrs using siNS, siIREVIP1-7, siTopoisomerase I, siTopoisomerase II α , or combinations of these. Cells were then incubated with CldU and then IdU for 30 minutes, followed by lysis and spreading on glass slides. After fixing, denaturing and blocking with 1% BSA, they were immunostained with antibodies against CldU and IdU and

analysed by microscopy. The length of CldU and IdU labelled tracts were measured using ImageJ (**Figure 6-6**).

IREVIP1 depletion reduced the average fibre tract length compared to siNS, as did siTopoisomerase I and siTopoisomerase II α to levels comparable with siIREVIP1 (**Figure 6-6**). Interestingly, co-depletion of IREVIP1 and Topoisomerase I was able to rescue fibre tract length to that of siNS, whereas co-depletion of IREVIP1 and Topoisomerase II α did not. Furthermore, co-depletion of IREVIP1, Topoisomerase I, and Topoisomerase II α did not rescue the fibre tract length (**Figure 6-6**).

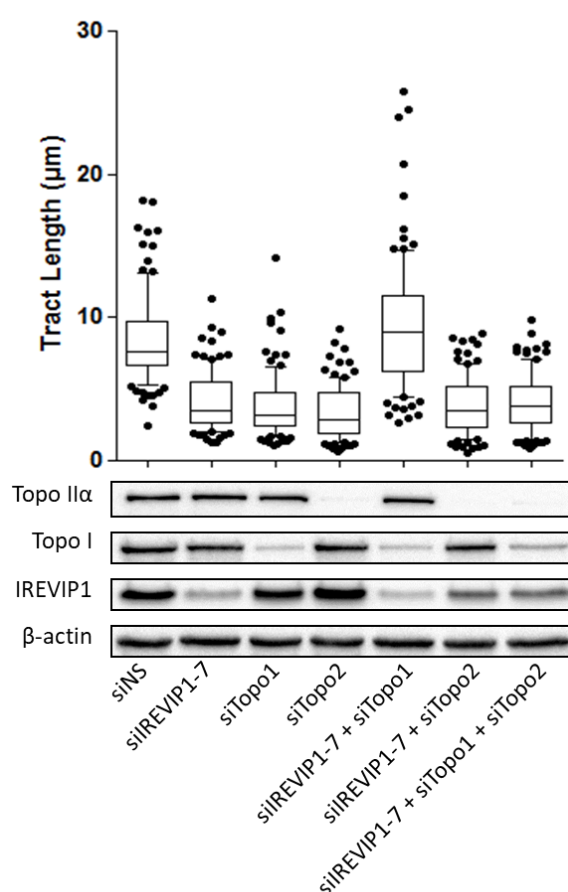


Figure 6-6 Depletion of IREVIP1 results in DNA replication fork stalling, which can be rescued by Topoisomerase I, but not Topoisomerase II α , co-depletion

HEK293 cells were depleted for 72 hours with the indicated siRNA and control non-silencing siRNA (luciferase). Cells were sequentially labelled with CldU and IdU prior to collection and stretching of fibres on slides. Slides were subsequently stained with antibodies against CldU and IdU prior to microscopy analysis. The CldU and IdU tract lengths were measured using ImageJ and plotted using GraphPad Prism. IREVIP1, Topoisomerase I and Topoisomerase II α depletion all reduced fibre tract length. Co-depletion of IREVIP1 and Topoisomerase I rescued fibre tract length, while co-depletion of IREVIP1 and Topoisomerase II α did not. Performed by J Fielden

6.7 Discussion

In light of the previous results, the role of IREVIP1 was investigated in the maintenance of genome stability. It was seen in HEK293 cells depleted of IREVIP1 that the ATM-CHK2 and ATR-CHK1 DNA damage signalling pathways were activated (**Figure 6-1**). This would suggest that IREVIP1 depleted cells are experiencing DSBs as well as long stretches of ssDNA, which could originate either from DSB processing or from stalled replication forks (Dai and Grant, 2010; Toledo, Neelsen and Lukas, 2017). If IREVIP1 is somehow involved in regulating Topoisomerases, IREVIP1 depletion could result in an increase in Topoisomerase related DNA damage, including DSBs and damage originating from stalled replication forks (Pommier *et al.*, 2016). However, different cell lines responded differently to IREVIP1 depletion (**Figure 6-2, Figure 6-3**).

HEK293 cells are derived from a neural cell from the complex human embryonic kidney cell culture which was transformed with sheared DNA from Adenovirus 5. A 4 kb fragment was inserted into chromosome 19 which expressed E1A, which interacts with Rb and promotes S-phase entry, and E1B which inhibits p53, thereby allowing G1 to S-phase and G2 to M-phase progression, as well as inhibiting apoptosis (Lin et al. 2014; Berk 2005). Therefore, this could explain why, even though both ATM-CHK2 and ATR-CHK1 pathways are activated, HEK293 cells show very little cell cycle arrest. The mild G1 arrest observed could be triggered by ATM-CHK2-mediated inhibition of CDK2 by preventing its CDC25A-mediated dephosphorylation and subsequent activation required for S-phase progression (Blagosklonny and Pardee, 2002; Dai and Grant, 2010). TIG1 primary fibroblasts show a varied response to IREVIP1 depletion, with siIREVIP1-7 being too toxic to carry out γ H2AX/53BP1 IF analysis. This fits with the cell cycle profile showing a checkpoint arrest in G1 as well as G2 with siIREVIP1-7, and the depletion also being slightly greater than with siIREVIP1-8 (**Figure 6-2 panel C, Figure 6-3**). siRNA sequences can have differing effects within the cell which could be down to off-target effects. If a sequence is targeting mRNA other than the target sequence, the depletion of the target could be affected,

therefore attenuating its effect. Off-target effects could also provide phenotypes not related to the depletion of the target protein. However, in this case use of IREVIP1 Δ CRISPR cells lines produced phenotypes in line with those seen with siRNA against IREVIP1 (**Figure 6-4**, **Figure 6-5**, and **Figure 6-6**). If siIREVIP1-8 is producing a milder phenotype in TIG1 this could mean that DNA damage is accumulating to a lesser extent as more IREVIP1 is still present, allowing cells to progress through the G1-S phase checkpoint, thereby accumulating more DNA damage during S phase, potentially Topoisomerase related, and finally arresting in G2/M. Topoisomerase II α is required for the G2/M transition and inhibiting it is sufficient to arrest cells in G2/M (Taylor and Stark, 2001). Therefore, if IREVIP1 is regulating Topoisomerases, this could be partly responsible for the observed G2/M arrest.

RPE1 cells are retinal pigmented epithelial cells immortalised using the catalytic subunit of human telomerase (hTERT). This allows cells to become immortalised, however does not cause cancerous changes or phenotypic alterations (Lee, Choi, and Ouellette 2004). Therefore, these cells are another good example of 'normal' cells. They behaved similarly, yet with more severe phenotypes, to TIG1 fibroblasts in γ H2AX/53BP1 foci formation (**Figure 6-2**) and cell cycle profile (**Figure 6-3**). This greater severity could be due to the greater degree of IREVIP1 depletion seen in RPE1 cells, and their ability to cope with this increased damage could be aided by their immortalisation. The RPE1 cells have a greater number of cells in S phase to begin with, which could mean that the population of cells has a greater potential to accumulate S phase related DNA damage. RPE1 cells also accumulate more γ H2AX post-IR than TIG1 cells, therefore suggesting that they may be more tolerant towards DNA damage before initiating apoptosis. It has previously been noted in HTC116 cells, which, like TIG1 and RPE1 cells, are p53 wild type, that DNA damage-induced p21 induction can result in G2 (lower p21) or G1 (higher p21) cell cycle arrest (Blagosklonny and Pardee, 2002). Therefore, the greater DNA damage seen in RPE1 cells could increase p21 levels, causing the observed G1 and G2 arrest to a greater degree than TIG1 cells.

One aspect that both TIG1 and RPE1 cells have in common is their reduced 53BP1 recruitment in the absence of IREVIP1, both with and without IR (**Figure 6-2**). This reduction in 53BP1 recruitment could be due to several things. 53BP1 is generally involved in NHEJ throughout the cell cycle, therefore its reduced recruitment could suggest that the HR pathway is being used instead, especially seeing as the cells are largely arrested in G2 where the sister chromatid would be present (Panier and Boulton, 2014). However, it has also been reported that 53BP1 can play a role in HR at heterochromatic regions during G2; therefore this may not be a satisfactory explanation (Kakarougkas, Ismail, Klement, *et al.*, 2013). Another reason could be that if IREVIP1 is regulating an aspect of Topoisomerase function, this could affect the topology of the DNA, making it less accessible to DDR proteins, such as 53BP1. The histone marks needed for 53BP1 recruitment (H4K20me2 and H2AK15Ub) (Hustedt and Durocher, 2017) could either be masked by this altered topology, or the proteins responsible for laying them down could be prevented from doing so. Furthermore, this could also be due to aberrant signalling downstream of γ H2AX (MDC1, RNF8, RNF168) required for sufficient 53BP1 recruitment to DSBs (Panier and Boulton, 2014; Schwertman, Bekker-Jensen and Mailand, 2016; Hustedt and Durocher, 2017). Therefore, it would be interesting to assess the distribution between NHEJ and HR in these cells, as well as the presence of the relevant PTMs, and signalling cascade downstream of γ H2AX, and whether this 53BP1 recruitment issue is just delayed or completely reduced in the absence of IREVIP1. Furthermore, p97 has previously been identified as being involved in the removal of L3MBTL1 from H4K20me2 ubiquitin marks in order to allow the access of 53BP1 to the chromatin (Acs *et al.*, 2011). Thus, investigating whether the lack of 53BP1 recruitment is in anyway related to p97 dependent protein turnover would also be prudent.

U2OS cells are a malignant osteosarcoma cell line which has a highly altered chromosome profile, often in the hypertriploid range (ATCC), yet retain wild type p53 status (Tang, Yuan and Guo, 2014). U2OS performed markedly differently to both TIG1 and RPE1 cells in both γ H2AX/53BP1 foci formation (**Figure 6-2**) and cell cycle profile (**Figure 6-3**). This could most likely

be explained by their highly altered nature. U2OS show a high degree of G2 arrest, which has been shown previously in response to p21 induction (Medema *et al.*, 1998; Taylor and Stark, 2001). Therefore, as the G1-S phase restriction point is often lost in cancer cells independently of p53 (Blagosklonny and Pardee, 2002), this could explain the increase in G2 arrested U2OS cells. The lack of G1 arrest could also mean that cells are progressing through S phase accumulating damage, especially as the percentage of S phase cells is not reduced upon IREVIP1 depletion (**Figure 6-3**). This could suggest an increased tolerance towards DNA damage before initiating apoptosis, as with RPE1 cells.

Altogether, IREVIP1 depletion is causing DNA damage in several different cell types; however the individual response varies depending on the cell type and genetic background.

One thing in common with all four cell lines used is the reduction in EdU incorporation in S phase cells, despite the reduction in cells entering S phase being taken into account (**Figure 6-3**). This would suggest that even if IREVIP1 depleted cells do manage to enter into S phase, they are facing problems during DNA replication. Therefore, the involvement of IREVIP1 with the DNA replication fork and its progression was investigated (**Figure 6-4**).

By using iPOND technology, IREVIP1 was found to be a part of the DNA replication fork, along with p97 and SPRTN (**Figure 6-4 panel A**). This suggests that IREVIP1, along with p97 and SPRTN, is required for the unperturbed progression of the DNA replication machinery. This could give weight to the previous speculation that IREVIP1 is recruited to Topoisomerases, thereby recruiting p97 and SPRTN, resulting in the processing of Topo-ccs and allowing DNA replication fork progression. Indeed, lack of IREVIP1 resulted in the accumulation of Topoisomerase I at the replication fork (**Figure 6-4 panel B**), as well as general accumulation of crosslinked Topoisomerase I and Topoisomerase II α on DNA (**Figure 6-5**). Together, this strongly suggests that IREVIP1 is involved in the removal of Topoisomerases from DNA even in unchallenged conditions.

Interestingly, IREVIP1 depletion reduced fibre tract length, which could be rescued by Topoisomerase I co-depletion, but not by co-depleting Topoisomerase II α or by a co-depletion of both Topoisomerases (**Figure 6-6**). Topoisomerase I is required to relieve the build-up of positive supercoiling ahead of the replication fork with Topoisomerase I trapping causing DNA replication fork stalling (Fachinetti *et al.*, 2010; Pommier *et al.*, 2016). Topoisomerase I-ccs increase in IREVIP1 depleted cells, and co-depletion of Topoisomerase I in IREVIP1-inactivated cells rescues DNA replication fork progression (**Figure 6-6**). This rescue effect could be due to cells still having enough IREVIP1 (approximately 20% of IREVIP1 remained after siRNA-depletion) to deal with the remaining Topoisomerase I-ccs, with still enough Topoisomerase I being left to carry out DNA replication, therefore restoring the equilibrium.

It has previously been reported in mammalian cells that depletion of either Topoisomerase II α or II β does not have any effect on DNA replication initiation or elongation, which could be why depletion of Topoisomerase II α does not rescue fibres upon IREVIP1 depletion (**Figure 6-6**). However, this contradicts what is shown here as depletion of Topoisomerase II α reduced fibre tract length (**Figure 6-6**) (Gómez-González *et al.*, 2011; Pommier *et al.*, 2016). Topoisomerase II α is required for the resolution of DNA catenanes which can be converted from precatenanes arising from the rotation of the replication fork (Fachinetti *et al.*, 2010; Bermejo, Lai and Foiani, 2012; Pommier *et al.*, 2016). The convergence of two DNA replication forks and replication through heterochromatic or centromeric DNA can often prevent the binding of Topoisomerases (Fachinetti *et al.*, 2010; Pommier *et al.*, 2016), requiring tension to be relieved by replication machinery rotation, therefore creating catenanes which are later resolved by Topoisomerase II α . In the absence of Topoisomerase II α , the presence of catenanes can result in the formation of mitosis-related DNA damage from ultra-fine bridges (Pommier *et al.*, 2016), and from the mechanical force required to separate topologically linked chromosomes (Bermejo *et al.*, 2009). This can either result in the failure of cell division and cell death, or restitution and an increase in the DNA copy number, resulting in multipolar mitosis in the absence of p53 (Gómez-González *et*

al., 2011). In budding yeast, it has been reported that Topoisomerase II α depletion requires cytokinesis to cause DNA damage, whereas expression of a catalytically inactive Topoisomerase II α causes DNA damage before mitosis and prevents completion of DNA replication, suggesting that the presence of Topoisomerase II α is needed for G2/M arrest in the presence of catenanes (Baxter and Diffley, 2008). Therefore, as these experiments were carried out in HEK293 cells depleted of Topoisomerase II α , they could escape a G2/M arrest, acquire extensive damage during mitosis, and then bypass the G1 checkpoint to then end up in S phase with damaged chromosomes, causing the fibre tract length problems observed with Topoisomerase II α depletion alone (**Figure 6-6**). This could also explain why Topoisomerase II α depletion does not rescue IREVIP1 depletion, as these problems observed are not resulting from Topoisomerase II-ccs.

Relaxation of positive supercoils by Topoisomerase I is required to reduce the formation of precatenanes, which could explain why depleting both Topoisomerase I and II α cannot rescue fibres upon IREVIP1 depletion, as this would create more damage which is not a result of Topoisomerase-ccs (Pommier *et al.*, 2016).

In summary, IREVIP1 depletion causes endogenous DNA damage, resulting in cell cycle arrest and a reduction in DNA replication fork progression. This appears to be via the processing of Topoisomerase-ccs as the replication fork advances, along with which p97, SPRTN and IREVIP1 were seen to move. Therefore, IREVIP1 appears to be a novel p97 adaptor for the processing of Topoisomerase-ccs at the replication fork, responsible for maintaining genome stability.

7. DISCUSSION

7.1 IREVIP1 is a Novel p97 Adaptor Able to Bind Topoisomerases

DNA Topoisomerases have been widely implicated in the generation of DSBs during gene expression and DNA replication (Lin *et al.*, 2009; Tuduri *et al.*, 2009; Gómez-Herreros *et al.*, 2017). Replication fork stalling is a significant cause of genome instability which often occurs when the replication machinery encounters DNA lesions or firmly bound protein complexes. Replication across actively transcribed genes has also been implicated in genome instability, however this is generally not due to direct collisions between the transcription and replication machinery, but between the machinery and DNA structures formed as a consequence of these process, such as R-loops and DNA topology. Both Topoisomerase I (Tuduri *et al.*, 2009) and Topoisomerase II (Bermejo *et al.*, 2009) have been implicated in the suppression of genomic instability arising from the interface between replication and transcription, particularly at chromosome fragile sites.

Topoisomerases release torsional stress accumulating from unwinding of the DNA duplex by introducing DNA breaks. Topoisomerases are covalently linked to the DNA via an intermediate referred to as the cleavage complex (Topo-ccs). This is usually a transient intermediate, however Topoisomerases can become trapped, thus requiring specialised removal. Such Topoisomerase poisons exist and are used in chemotherapy, which result in the formation of non-resolvable Topo-ccs, resulting in cell death (Burden and Osheroff, 1998; Pommier *et al.*, 1998, 2010; Bandele and Osheroff, 2007).

TDP1, the enzyme responsible for cleaving Topoisomerase I from the 3' end of DNA was discovered by Yang *et al.* in 1996 and TDP2, the human enzyme responsible for cleaving Topoisomerase II from the 5' end of DNA by Ledesma *et al.* in 2009. TDP1 and TDP2 cleave the covalent bond joining the Topoisomerase and the DNA, generating ends which can be repaired by the NHEJ DNA repair machinery, thereby reducing the occurrence of chromosome

translocations, especially during gene transcription, and alleviating the toxicity of Topoisomerase II poisons such as etoposide, thereby maintaining genome stability (Yang *et al.*, 1996; Wu and Liu, 1997; Ledesma *et al.*, 2009; Zeng *et al.*, 2011; Gómez-Herreros *et al.*, 2017). Topoisomerase II α is also important for chromosome condensation and segregation during mitosis as it is responsible for the separation of newly replicated sister chromatids, therefore preventing mitotic derived genomic instability (Guturi *et al.*, 2016). However, it has been reported that TDP2 can contribute to the resolution of Topoisomerase I-ccs in the absence of TDP1, suggesting some cross-talk between processing Topoisomerase I and Topoisomerase II-ccs, as with between the activity of the two Topoisomerases as well (Zeng *et al.*, 2012).

Topoisomerase II α is heavily modified by ubiquitination, SUMOylation, and phosphorylation, and these PTMs regulate Topoisomerase II α localisation, function, and degradation. BRCA1 and RNF168 have been implicated in stimulatory K63-linked ubiquitination of Topoisomerase II α , which is counteracted by the DUB USP10. Ubiquitination by BRCA1, as well as by BMI1/RING1A, and APC/C^{Cdh1}, has also been reported to cause proteasomal degradation of Topoisomerase II α (Guturi *et al.*, 2016). Indeed, the degradation of Topoisomerase II α , but not its orthologue II β , has been linked to the ubiquitin-proteasome pathway, and is both ubiquitin and ATP-dependent. In fact, Topoisomerase II protein levels are reduced in a proteasome dependent manner in the presence of Topoisomerase poisons, resulting in resistance to poisons which can be overcome by proteasome inhibition (Lee *et al.*, 2016).

However, the accumulation of replication-transcription-dependent Topoisomerase I induced DSBs was dependent on the proteasomal degradation of Topo-ccs, exposing SSBs, leading to replication run-off and DSB formation. Concurrently, proteasomal inhibition by MG-132 or knockdown of the 20S proteasome maturation protein conferred resistance to camptothecin Topoisomerase I poisons. As with Topoisomerase II-ccs, this also highlights proteasomal involvement in Topoisomerase I-cc processing (Lin *et al.*, 2009). Therefore, the ubiquitination

status of Topoisomerases and their proteasomal degradation could be of interest when investigating the effectiveness and resistance to Topoisomerase poisons such as etoposides, including epirubicin, and mitoxantrone, and camptothecins, including topotecan and irinotecan, in cancer therapy (Lin *et al.*, 2009; Gao *et al.*, 2014; Lee *et al.*, 2016).

Interestingly, Topoisomerase I and Topoisomerase II Topo-ccs cannot be processed by TDP1 and TDP2 respectively without first undergoing proteolytic processing by the 26S proteasome, due to the relatively tight structure of the catalytic site (Deb  thune *et al.*, 2002; Interthal and Champoux, 2011; Gao *et al.*, 2014).

Recently, SPRTN, a previously characterised p97 adaptor involved in regulating translesion synthesis (Davis *et al.*, 2012; Mosbech *et al.*, 2012), was also found to be responsible for proteolysis of DNA-protein crosslinks, including many different substrates, but of greatest interest were Topo-ccs (Stingele *et al.*, 2016; Vaz *et al.*, 2016; Maskey *et al.*, 2017). This therefore bridges the missing link between Topo-ccs and their removal from DNA by TDP1/2.

The data presented in this thesis identifies and describes a previously uncharacterised, novel, p97 adaptor, which is here referred to as IREVIP1. IREVIP1 appears to play a role in recruitment of p97 to Topoisomerases and is important for preventing DNA damage and genome instability (**Figure 7-1**).

IREVIP1 was first of interest due to the homology of one of its domains to that of the bacterial DNA gyrase inhibitor, Gyrl (Nakanishi *et al.*, 1998), as the regulation of Topoisomerases in mammalian cells is still not fully understood and is of great interest for cancer therapeutics (Burden and Osheroff, 1998; Pommier *et al.*, 1998; Bandele and Osheroff, 2007; Koster *et al.*, 2007; Nitiss, 2009; Maede *et al.*, 2014; Meisenberg *et al.*, 2014; G  mez-Herreros *et al.*, 2017) as well as because of its predicted SHP and VBM p97 interaction domains. It was confirmed that IREVIP1 is able to interact with p97 *in vitro* where this interaction was abolished in the absence of the SHP, but not the VBM domain, as well as *in vivo*, therefore identifying IREVIP1 as a bona

fide p97 adaptor protein. This interaction does not require the presence of other known p97 adaptors *in vitro*, unlike some other adaptors which require the present of one of the three core p97 adaptor complex, therefore forming an adaptor hierarchy (Hänzelmann, Buchberger and Schindelin, 2011). A very distant relative of p97 is present in only a few eubacterial species (Barthelme and Sauer, 2016), therefore it is highly unlikely that p97 or adaptors co-purified with either overexpressed IREVIP1 or p97. However, this would be interesting to also test *in vivo* by depletion of each of the core complexes in turn and observing whether the p97-IREVIP1 interaction remains.

As SPRTN is a protease able to degrade many different types of DPCs, the questions is raised as to how SPRTN is directed to specific DPCs, such as Topo-ccs (Stingele *et al.*, 2016; Vaz *et al.*, 2016). Therefore, as IREVIP1 is able to bridge the p97-Topoisomerase interaction, it would be exciting to speculate that IREVIP1 could be a joint adaptor for the direction of SPRTN towards Topo-ccs. SPRTN binds p97 in the presence of the core adaptors UFD1 and NPL4 via its SHP domain, which is also present in UFD1 (Davis *et al.*, 2012; Mosbech *et al.*, 2012). As p97 is a homohexamer, this in theory leaves four out of the six SHP binding motifs available for additional adaptor binding, thus enabling this hypothesis (Hänzelmann, Buchberger and Schindelin, 2011).

IREVIP1 was observed as being able to bind both Topoisomerase I and Topoisomerase II α . It would be intriguing as to whether IREVIP1 was able to bind other mammalian Topoisomerases such as Topoisomerase II β or Topoisomerase III, however due to the limitations of this thesis it was not possible so far. It would also be prudent to map the binding domains in IREVIP1 responsible for the interaction with Topoisomerase I and Topoisomerase II α , and whether, as could be postulated, this lies within the predicted gyrase inhibitory domain.

7.2 IREVIP1 is Located at the Nuclear Membrane, Where Torsional Stress is Often Exacerbated

IREVIP1 was found to be located to the nuclear membrane via its hydrophobic N-terminal domain, and in complex with various nuclear pore complex proteins by Mass Spectrometry and Co-IP. Actively transcribed genes are often tethered to the nuclear pore complex via THO-TREX-2 (human orthologues THO complex - THOC1/THOC2, TREX complex – THOC4/BAT1, TREX2 complex – SHD1/GANP/ENY2/Q5JVF3) in a process called gene gating (Bermejo *et al.*, 2011; Bermejo, Lai and Foiani, 2012). Although the role of ATR in releasing gene gating has not been shown in mammals, the two residues for ATR-dependent phosphorylation are highly conserved and lie within the same domain (Bermejo *et al.*, 2011). Furthermore, a role for the NUP Nup210 in gene gating has also been reported in mammalian cells as well (D'Angelo *et al.*, 2012). This tethering can add to the generation of DNA torsion by providing a topological barrier. Therefore, Topoisomerases are of particular importance at these structures, in conjunction with the release of transcribed genes from the nuclear pore in a Mec1/ATR and RAD53-dependent manner by checkpoint kinase phosphorylation of nucleoporins in yeast (Bermejo *et al.*, 2011). This coupling of mRNA transcription with nuclear export via the nuclear pore complexes is important for preventing the entanglement of the nascent RNA with the transcription bubble and R-loop formation. Furthermore, this entanglement can prevent the rotation of the RNA polymerase, thereby exacerbating transcription-generated DNA torsion (Bermejo *et al.*, 2009; Bermejo, Lai and Foiani, 2012). Topoisomerase II α in yeast, along with Hmo1 and Htz1, is also important for facilitating the formation of higher-order chromatin organisation in a process termed gene looping, which promotes transcription by forming interactions between spatially separated regulatory elements, whilst at the same time insulating transcription, and transcription generating DNA torsion from other DNA-dependent processes, thus minimising genome instability (Bermejo *et al.*, 2009; Bermejo, Lai and Foiani, 2012). Topoisomerase I also plays a

role in regulating transcription and replication collisions by preventing R-loop formation. Topoisomerase I regulates torsion generated ahead of DNA replication forks. Replication forks are markedly slower in the absence of Topoisomerase I, which is compensated for by firing dormant origins. The accumulation of R-loops in the absence of Topoisomerase I due to the nascent RNA annealing with the negatively supercoiled DNA, results in increase fork stalling and genomic instability (Tuduri *et al.*, 2009). In addition to preventing R-loop formation, Topoisomerase I is necessary for preventing G-quadruplex formation at repetitive DNA sequences, which can also be a significant source of genomic instability as they pose a barrier to DNA replication (Yadav *et al.*, 2014). Therefore, it is intriguing and logical that IREVIP1 might be found in the region where more Topo-ccs might occur, positioning it in a prime location to deal with them and prevent genome instability. It would be thought-provoking to see whether IREVIP1, which cannot locate to the nuclear membrane, is able to rescue the phenotypes seen in cells lacking IREVIP1, or whether this is merely just a convenience of locality which is not necessary for wider IREVIP1 function. However, the physical location of IREVIP1 could be a mechanism by which IREVIP1 is able to direct p97-UN-SPRTN complexes to the region of high Topo-ccs, thereby conferring SPRTN specificity. Furthermore, it would also be intriguing to compare the level of Topo-ccs in both SPRTN and IREVIP1 depleted cells and see how similar this is.

7.3 IREVIP1 is Necessary to Prevent Genome Instability

Concurrently with the knowledge that Topoisomerases are required to prevent genome instability, a lack of IREVIP1 was also perceived to cause the same fate. Indeed, it was observed that upon IREVIP1 depletion, key DDR proteins were activated, such as γ H2AX, pATM, pATR, pCHK1, and pCHK2, resulting in cell cycle arrest across a panel of cell lines. In line with IREVIP1 being in complex with various DNA replication proteins and its absence resulting in the shortening of DNA replication fibres, I would posit that the genomic instability seen in IREVIP1

depleted cells arises from DNA replication. It would also be provocative to test whether there is a similar perturbation of transcription and whether RNA polymerases are stalled in the absence of IREVIP1.

Upon IREVIP1 depletion, variability in the recruitment of 53BP1 was observed between different cell lines. p97 has previously been reported to be responsible for removal of L3MBTL1 from H4K20me2, the signal responsible for the recruitment of 53BP1 to DNA damage (Acs *et al.*, 2011). However, as IREVIP1 has been implicated in the recruitment of p97 to Topo-ccs, it would be interesting to see whether there is any increase in DNA torsion which could potentially mask the histone mark, preventing 53BP1 recruitment. It could be suggested that this is more likely than an involvement of IREVIP1 in the removal of L3MBTL1 as there is no reason to believe that this is the case, especially seeing as it was not identified in the IREVIP1 Mass Spectrometry analysis. It would therefore be advisable to investigate the kinetics of 53BP1 recruitment upon IREVIP1 depletion to determine whether the kinetics are delayed or abolished. Traditionally, 53BP1 is required for NHEJ in the absence of a sister chromatid. However, it has been reported that 53BP1 can be required for HR-mediated repair of DSBs positioned in heterochromatin (Kakaroukas, Ismail, Klement, *et al.*, 2013). Notably, U2OS cells which showed no defect in 53BP1 recruitment were largely arrested in G2. Therefore, it would be wise to assess the distribution between the use of NHEJ and HR within the cell and whether this can be correlated with cell type and 53BP1 recruitment, and even the degree of genomic instability observed within the cell upon IREVIP1 depletion.

This would therefore attractively combine both DNA replication and transcription and their role in DNA repair and genomic instability (Helmrich *et al.*, 2013; Brambati *et al.*, 2015; García-Muse and Aguilera, 2016; Zhang *et al.*, 2016; Marnef, Cohen and Legube, 2017), consequently highlighting the importance of IREVIP1 as a caretaker and illuminating its potential as a therapeutic target to aid in the battle against drug resistance (Chaudhuri *et al.*, 2016), in

particular in reference to Topoisomerase mutations which prevent the action of the drug in trapping Topo-ccs, but where Topoisomerase is still functional (Chrencik *et al.*, 2004; Gongora *et al.*, 2011). However, it may not be of particular use in situations where the activity of Topoisomerases is reduced, thus reducing the number of Topo-ccs in which to trap, as in the case of nitric oxide production in the cell, which inhibits the ATPase activity of Topoisomerase II, decreasing the amount of Topo-ccs and the formation of Topo-cc induced DSBs (Sinha, Kumar and Mason, 2017).

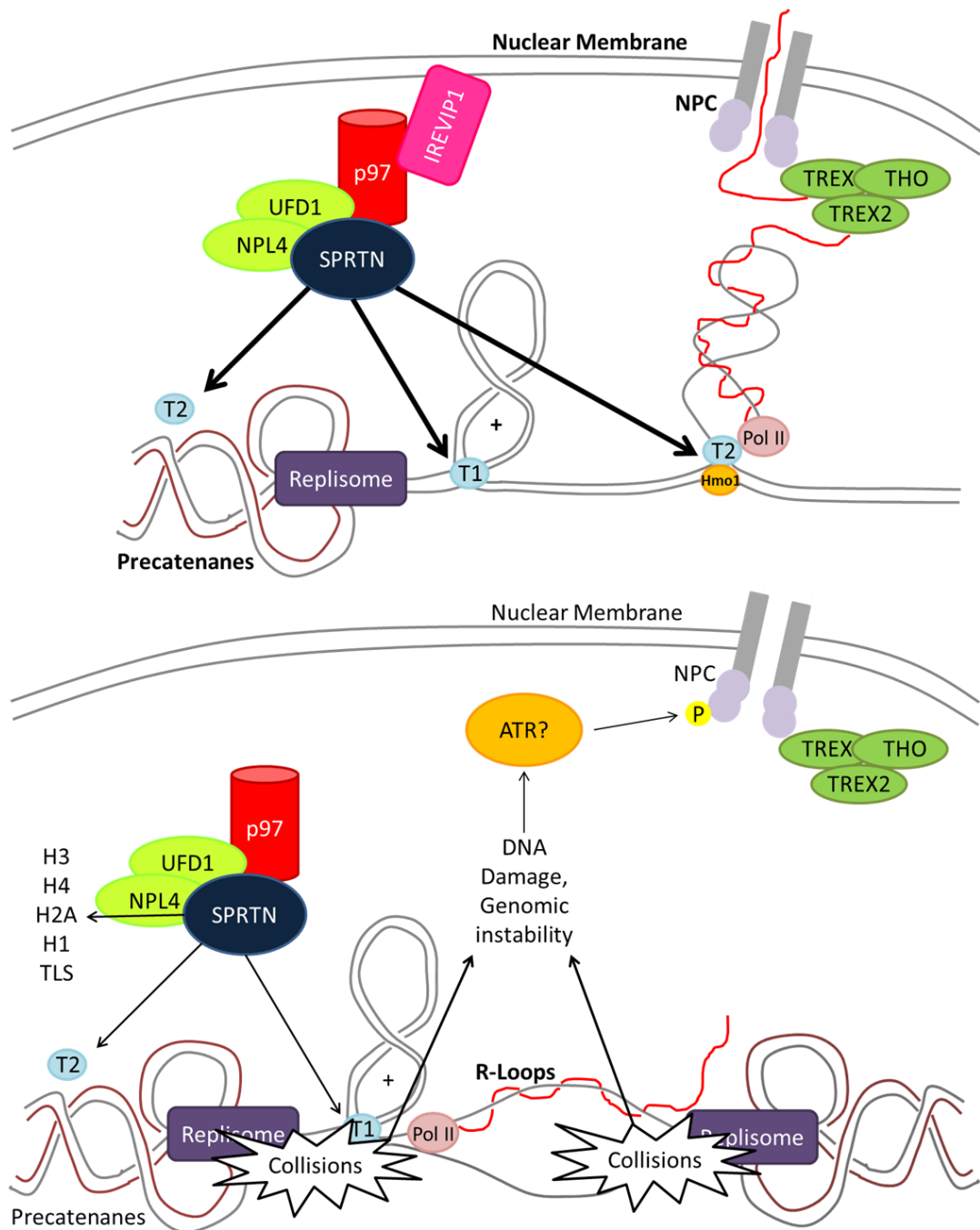


Figure 7-1 IREVIP1 is necessary to prevent genomic instability and interacts with p97 and Topoisomerases
Upper panel: IREVIP1 is a nuclear membrane anchored protein which is able to bind both p97 and Topoisomerase I or II α and bridges the gap between them. SPRTN, a previously described p97 adaptor, is responsible for the proteolytic digestion of Topo-ccs prior to removal by TDP1/2. IREVIP1 may be responsible for directing p97-SPRTN to Topo-ccs over the other roles that SPRTN plays. This would aid to remove Topo-ccs in regions of DNA around the nuclear membrane which can accumulate high torsional stress. This can be generated from processes such as DNA replication and transcription which open up the DNA helix. Collision of these machineries with DNA obstacles can result in DNA damage and genomic instability.
Lower panel: In the absence of IREVIP1, extensive genomic instability is seen in mammalian cells, resulting in cell cycle arrest. An increase in the number of Topo-ccs is observed in the absence of IREVIP1, which corroborates the hypothesis that IREVIP1 is necessary for their timely removal. It could be posited that in the absence of IREVIP1, the p97-SPRTN complex is thinly spread between many different pathways, leaving Topo-ccs neglected, causing DNA damage and genomic instability.

7.4 Future Work

As briefly discussed earlier, it would be prudent to investigate the regions responsible for the interaction between IREVIP1 and Topoisomerases, and whether this can be applied to a wider range of Topoisomerases. Furthermore, as IREVIP1 localises to the nuclear membrane, a common area for replication-transcription collisions and nuclear pore tethering, which can cause increased DNA torsion, it would be intriguing to test whether an N-terminal deletion mutant would be able to rescue genomic instability and the increased number of Topo-ccs seen. IREVIP1 could have functions away from the nuclear membrane, yet it is exciting to think that this localisation mechanism has evolved to deal with this specific region of high torsional stress. Elucidating the mechanism of this interaction could provide means by which to target the interaction, thereby using it to increase the toxicity of current Topoisomerase poisons. The rationale behind this is to facilitate the use of lower doses of drugs currently in use, thereby reducing off-target toxicity and circumventing potential resistance mechanisms. This would be more beneficial than broad range p97 inhibitors which target the multiple pathways in which p97 is involved, causing extensive toxicity to normal tissue as well as tumours. Being able to target a specific p97 function via one of its adaptors would be an intelligent way of avoiding this problem.

Much more presumably remains to be discovered regarding IREVIP1 which could lead to further understanding of genome stability within the cell, offering potentially more elegant mechanisms to remain one step ahead of the evolution of the ever growing issue of cancers.

8. REFERENCES

- Aboussekhra, A., Biggerstaff, M., Shivji, M. K. K., Vilpo, J. A., *et al.* (1995) 'Mammalian DNA nucleotide excision repair reconstituted with purified protein components', *Cell*, 80(6), pp. 859–868. doi: [https://doi.org/10.1016/0092-8674\(95\)90289-9](https://doi.org/10.1016/0092-8674(95)90289-9).
- Acs, K., Luijsterburg, M. S., Ackermann, L., Salomons, F. a, *et al.* (2011) 'The AAA-ATPase VCP/p97 promotes 53BP1 recruitment by removing L3MBTL1 from DNA double-strand breaks', *Nature Structural & Molecular Biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 18(12), pp. 1345–1350. doi: 10.1038/nsmb.2188.
- Adams, J. (2004) 'The development of proteasome inhibitors as anticancer drugs', *Cancer Cell*, 5(5), pp. 417–421. doi: [https://doi.org/10.1016/S1535-6108\(04\)00120-5](https://doi.org/10.1016/S1535-6108(04)00120-5).
- Adams, J., Palombella, V. J., Sausville, E. A., Johnson, J., *et al.* (1999) 'Proteasome Inhibitors: A Novel Class of Potent and Effective Antitumor Agents', *Cancer Research*, 59(11), p. 2615 LP-2622. Available at: <http://cancerres.aacrjournals.org/content/59/11/2615.abstract>.
- Aguilera, A. and Gómez-González, B. (2008) 'Genome instability: a mechanistic view of its causes and consequences', *Nature Reviews Genetics*. Nature Publishing Group, 9(3), pp. 204–217. doi: 10.1038/nrg2268.
- Akutsu, M., Dikic, I. and Bremm, A. (2016) 'Ubiquitin chain diversity at a glance', *Journal of Cell Science*, 129(5), p. 875 LP-880. Available at: <http://jcs.biologists.org/content/129/5/875.abstract>.
- Anderson, D. J., Le Moigne, R., Djakovic, S., Kumar, B., *et al.* (2015) 'Targeting the AAA ATPase p97 as an Approach to Treat Cancer through Disruption of Protein Homeostasis', *Cancer Cell*, 28(5), pp. 653–665. doi: 10.1016/j.ccell.2015.10.002.
- Austin, C. A., Patel, S., Ono, K., Nakane, H. and Fisher, L. M. (1992) 'Site-specific DNA cleavage by

- mammalian DNA topoisomerase II induced by novel flavone and catechin derivatives', *Biochemical Journal*, 282(3), p. 883 LP-889. Available at:
<http://www.biochemj.org/content/282/3/883.abstract>.
- Bailey, R., Priego Moreno, S. and Gambus, A. (2015) 'Termination of DNA replication forks: "Breaking up is hard to do".', *Nucleus (Austin, Tex.)*, 6(3), pp. 187–96. doi: 10.1080/19491034.2015.1035843.
- Balakrishnan, L., Gloor, J. W. and Bambara, R. A. (2010) 'Reconstitution of eukaryotic lagging strand DNA replication', *Methods*. Elsevier Inc., 51(3), pp. 347–357. doi: 10.1016/j.ymeth.2010.02.017.
- Ballar, P., Shen, Y., Yang, H. and Fang, S. (2006) 'The role of a novel p97/valosin-containing protein-interacting motif of gp78 in endoplasmic reticulum-associated degradation.', *The Journal of biological chemistry*. United States, 281(46), pp. 35359–35368. doi: 10.1074/jbc.M603355200.
- Bandelet, O. J. and Osherooff, N. (2007) 'Bioflavonoids as poisons of human topoisomerase II alpha and II beta.', *Biochemistry*. United States, 46(20), pp. 6097–6108. doi: 10.1021/bi7000664.
- Bartek, J., Lukas, C. and Lukas, J. (2004) 'Checking on DNA damage in S phase', *Nat Rev Mol Cell Biol*. Nature Publishing Group, 5(10), pp. 792–804. Available at:
<http://dx.doi.org/10.1038/nrm1493>.
- Barthelme, D. and Sauer, R. T. (2016) 'Origin and Functional Evolution of the Cdc48/p97/VCP AAA+ Protein Unfolding and Remodeling Machine', *Journal of Molecular Biology*, 428(9, Part B), pp. 1861–1869. doi: <https://doi.org/10.1016/j.jmb.2015.11.015>.
- Bassermann, F., Frescas, D., Guardavaccaro, D., Busino, L., *et al.* (2008) 'The Cdc14B-Cdh1-Plk1 Axis Controls the G2 DNA-Damage-Response Checkpoint', *Cell*, 134(2), pp. 256–267.

doi: 10.1016/j.cell.2008.05.043.

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., *et al.* (2017) 'UniProt: The universal protein knowledgebase', *Nucleic Acids Research*, 45(D1), pp. D158–D169. doi: 10.1093/nar/gkw1099.

Baxter, J. and Diffley, J. F. X. (2008) 'Topoisomerase II Inactivation Prevents the Completion of DNA Replication in Budding Yeast', *Molecular Cell*, 30(6), pp. 790–802. doi: 10.1016/j.molcel.2008.04.019.

Bayona-Feliu, A., Casas-Lamesa, A., Reina, O., Bernues, J. and Azorin, F. (2017) 'Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin.', *Nature communications*. England, 8(1), p. 283. doi: 10.1038/s41467-017-00338-5.

Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A. and Hampton, R. Y. (2001) 'Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation.', *Nature cell biology*. England, 3(1), pp. 24–29. doi: 10.1038/35050524.

Bebeacua, C., Förster, A., Mckeown, C., Meyer, H. H., *et al.* (2016) 'Distinct conformations of the protein complex p97- Ufd1-Npl4 revealed by electron cryomicroscopy', *Pnas*, 2015. doi: 10.1073/pnas.1114341109/-
/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1114341109.

Bedford, L., Lowe, J., Dick, L. R., Mayer, R. J. and Brownell, J. E. (2011) 'Ubiquitin-like protein conjugation and the ubiquitin–proteasome system as drug targets', *Nat Rev Drug Discov*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 10(1), pp. 29–46. Available at: <http://dx.doi.org/10.1038/nrd3321>.

Bell, S. P. (2017) 'Rethinking origin licensing', *eLife*. eLife Sciences Publications, Ltd, 6, p. e24052. doi: 10.7554/eLife.24052.

Bentley, N. J., Holtzman, D. A., Flaggs, G., Keegan, K. S., *et al.* (1996) 'The Schizosaccharomyces

- pombe rad3 checkpoint gene.', *The EMBO journal*, 15(23), pp. 6641–51. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=452488&tool=pmcentrez&rendertype=abstract>.
- Bergink, S., Ammon, T., Kern, M., Schermelleh, L., *et al.* (2013) 'Role of Cdc48/p97 as a SUMO-targeted segregase curbing Rad51–Rad52 interaction', *Nat Cell Biol.* Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 15(5), pp. 526–532. Available at: <http://dx.doi.org/10.1038/ncb2729>.
- Berk, A. J. (2005) 'Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus', *Oncogene*, 24(52), pp. 7673–7685. doi: 10.1038/sj.onc.1209040.
- Bermejo, R., Capra, T., Gonzalez-Huici, V., Fachinetti, D., *et al.* (2009) 'Genome-Organizing Factors Top2 and Hmo1 Prevent Chromosome Fragility at Sites of S phase Transcription', *Cell*. Elsevier Ltd, 138(5), pp. 870–884. doi: 10.1016/j.cell.2009.06.022.
- Bermejo, R., Capra, T., Jossen, R., Colosio, A., *et al.* (2011) 'The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores', *Cell*, 146(2), pp. 233–246. doi: 10.1016/j.cell.2011.06.033.
- Bermejo, R., Lai, M. S. and Foiani, M. (2012) 'Preventing Replication Stress to Maintain Genome Stability: Resolving Conflicts between Replication and Transcription', *Molecular Cell*. Elsevier Inc., 45(6), pp. 710–718. doi: 10.1016/j.molcel.2012.03.001.
- Beuron, F., Dreveny, I., Yuan, X., Pye, V. E., *et al.* (2006) 'Conformational changes in the AAA ATPase p97–p47 adaptor complex', *The EMBO Journal*, 25(9), pp. 1967–1976. doi: 10.1038/sj.emboj.7601055.
- de Bie, P. and Ciechanover, A. (2011) 'Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms', *Cell Death and*

- Differentiation*. Nature Publishing Group, 18(9), pp. 1393–1402. doi: 10.1038/cdd.2011.16.
- Blackford, A. N. and Jackson, S. P. (2017) 'ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response', *Molecular Cell*, 66(6), pp. 801–817. doi: <http://dx.doi.org/10.1016/j.molcel.2017.05.015>.
- Blagosklonny, M. V. and Pardee, A. B. (2002) 'The restriction point of the cell cycle.', *Cell Cycle*, 1(2), pp. 103–110. doi: 10.4161/cc.1.2.108.
- Bochman, M. L., Paeschke, K. and Zakian, V. A. (2012) 'DNA secondary structures: stability and function of G-quadruplex structures', *Nat Rev Genet*, 13(11), pp. 770–780. doi: 10.1038/nrg3296.
- Bodnar, N. O. and Rapoport, T. A. (2017) 'Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex', *Cell*. Elsevier Inc., 169(4), p. 722–735.e9. doi: 10.1016/j.cell.2017.04.020.
- Bonifacino, J. S., Cosson, P. and Klausner, R. D. (1990) 'Colocalized transmembrane determinants for ER degradation and subunit assembly explain the intracellular fate of TCR chains', *Cell*, 63(3), pp. 503–513. doi: 10.1016/0092-8674(90)90447-M.
- Booher, R. N., Holman, P. S. and Fattaey, A. (1997) 'Human Myt1 is a cell cycle-regulated kinase that inhibits Cdc2 but not Cdk2 activity', *Journal of Biological Chemistry*, 272(35), pp. 22300–22306. doi: 10.1074/jbc.272.35.22300.
- van den Boom, J., Wolf, M., Weimann, L., Schulze, N., *et al.* (2016) 'VCP/p97 Extracts Sterically Trapped Ku70/80 Rings from DNA in Double-Strand Break Repair', *Molecular Cell*. Elsevier Inc., 64(1), pp. 189–198. doi: 10.1016/j.molcel.2016.08.037.
- Bornhorst, J. A. and Falke, J. J. (2000) '[16] Purification of Proteins Using Polyhistidine Affinity Tags. Methods in enzymology', *Methods in enzymology*, 326, pp. 245–254. Available at:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2909483/>.

- Brambati, A., Colosio, A., Zardoni, L., Galanti, L. and Liberi, G. (2015) 'Replication and transcription on a collision course: eukaryotic regulation mechanisms and implications for DNA stability', *Frontiers in Genetics*, 6(April), pp. 1–8. doi: 10.3389/fgene.2015.00166.
- Buchberger, A., Schindelin, H. and Hänzelmann, P. (2015) 'Control of p97 function by cofactor binding', *FEBS Letters*, pp. 2578–2589. doi: 10.1016/j.febslet.2015.08.028.
- Burden, D. A. and Osheroff, N. (1998) 'Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme', *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1400(1–3), pp. 139–154. doi: 10.1016/S0167-4781(98)00132-8.
- Carroll, T. M. and Setlow, P. (2005) 'Site-directed mutagenesis and structural studies suggest that the germination protease, GPR, in spores of *Bacillus* species is an atypical aspartic acid protease', *Journal of Bacteriology*, 187(20), pp. 7119–7125. doi: 10.1128/JB.187.20.7119-7125.2005.
- Carter, T., Vancurova, I., Sun, I., Lou, W. and DeLeon, S. (1990) 'A DNA-activated protein kinase from HeLa cell nuclei.', *Mol. Cell. Biol.*, 10(12), pp. 6460–6471. doi: 10.1128/MCB.10.12.6460.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., *et al.* (2012) 'The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.', *Cancer discovery*. United States, 2(5), pp. 401–404. doi: 10.1158/2159-8290.CD-12-0095.
- Chaganti, R. S., Schonberg, S. and German, J. (1974) 'A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 71(11), pp. 4508–4512.
- Chang, D. C., Xu, N. and Luo, K. Q. (2003) 'Degradation of cyclin B is required for the onset of

- anaphase in mammalian cells', *Journal of Biological Chemistry*, 278(39), pp. 37865–37873. doi: 10.1074/jbc.M306376200.
- Chapman, J. R., Barral, P., Vannier, J.-B., Borel, V., *et al.* (2013) 'RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection.', *Molecular cell*. United States, 49(5), pp. 858–871. doi: 10.1016/j.molcel.2013.01.002.
- Chapman, J. R., Sossick, A. J., Boulton, S. J. and Jackson, S. P. (2012) 'BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair', *Journal of Cell Science*, 125(15), p. 3529 LP-3534. Available at: <http://jcs.biologists.org/content/125/15/3529.abstract>.
- Chatterji, M. and Nagaraja, V. (2002) 'Gyrl: A counter-defensive strategy against proteinaceous inhibitors of DNA gyrase', *EMBO Reports*, 3(3), pp. 261–267. doi: 10.1093/embo-reports/kvf038.
- Chatterji, M., Sengupta, S. and Nagaraja, V. (2003) 'Chromosomally encoded gyrase inhibitor Gyrl protects Escherichia coli against DNA-damaging agents', *Archives of Microbiology*, 180(5), pp. 339–346. doi: 10.1007/s00203-003-0598-4.
- Chaudhuri, A., Callen, E., Ding, X., Gogola, E., *et al.* (2016) 'Replication fork stability confers chemoresistance in BRCA-deficient cells', *Nature*. Nature Publishing Group, 535(7612), pp. 382–387. doi: 10.1038/nature18325.
- Chia, W. S., Chia, D. X., Rao, F., Bar Nun, S. and Geifman Shochat, S. (2012) 'ATP Binding to p97/VCP D1 Domain Regulates Selective Recruitment of Adaptors to Its Proximal N-Domain', *PLoS ONE*, 7(12). doi: 10.1371/journal.pone.0050490.
- Chibazakura, T., Kamachi, K., Ohara, M., Tane, S., *et al.* (2011) 'Cyclin A Promotes S-Phase Entry via Interaction with the Replication Licensing Factor Mcm7', *Molecular and Cellular*

- Biology*. American Society for Microbiology (ASM), 31(2), pp. 248–255. doi: 10.1128/MCB.00630-10.
- Chou, T.-F., Brown, S. J., Minond, D., Nordin, B. E., *et al.* (2011) 'Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 108(12), pp. 4834–4839. doi: 10.1073/pnas.1015312108.
- Chrencik, J. E., Staker, B. L., Burgin, A. B., Pourquier, P., *et al.* (2004) 'Mechanisms of camptothecin resistance by human topoisomerase I mutations', *Journal of Molecular Biology*, 339(4), pp. 773–784. doi: 10.1016/j.jmb.2004.03.077.
- Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., *et al.* (2012) 'Defining human ERAD networks through an integrative mapping strategy', *Nature cell biology*, 14(1), pp. 93–105. doi: 10.1038/ncb2383.
- Christianson, J. C. and Ye, Y. (2014) 'Cleaning up in the endoplasmic reticulum: ubiquitin in charge', *Nat Struct Mol Biol*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 21(4), pp. 325–335. Available at: <http://dx.doi.org/10.1038/nsmb.2793>.
- Cleary, J. D., Nichol, K., Wang, Y.-H. and Pearson, C. E. (2002) 'Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells.', *Nature genetics*. United States, 31(1), pp. 37–46. doi: 10.1038/ng870.
- Cortes-Ciriano, I., Lee, S., Park, W.-Y., Kim, T.-M. and Park, P. J. (2017) 'A molecular portrait of microsatellite instability across multiple cancers', *Nature Communications*, 8, p. 15180. doi: 10.1038/ncomms15180.
- Costa, S., Almeida, A., Castro, A. and Domingues, L. (2014) 'Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: The novel Fh8 system', *Frontiers in*

- Microbiology*, 5(FEB), pp. 1–20. doi: 10.3389/fmicb.2014.00063.
- Coulonval, K., Kookan, H. and Roger, P. P. (2011) 'Coupling of T161 and T14 phosphorylations protects cyclin B–CDK1 from premature activation', *Molecular Biology of the Cell*. Edited by M. J. Solomon. The American Society for Cell Biology, 22(21), pp. 3971–3985. doi: 10.1091/mbc.E11-02-0136.
- Crespan, E., Hübscher, U. and Maga, G. (2015) 'Expansion of CAG triplet repeats by human DNA polymerases λ and β in vitro, is regulated by flap endonuclease 1 and DNA ligase 1', *DNA Repair*, 29, pp. 101–111. doi: <https://doi.org/10.1016/j.dnarep.2015.01.005>.
- Crowe, J., Masone, B. S. and Ribbe, J. (1995) 'One-step purification of recombinant proteins with the 6xHis tag and Ni-NTA resin', in Harwood, A. J. (ed.) *Molecular Biotechnology*. Totowa, NJ: Humana Press, pp. 247–258. doi: 10.1007/BF02779018.
- Croy, R. G. and Pardee, A. B. (1983) 'Enhanced synthesis and stabilization of Mr 68,000 protein in transformed BALB/c-3T3 cells: candidate for restriction point control of cell growth', *Proceedings of the National Academy of Sciences of the United States of America*, 80(15), pp. 4699–4703. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6308643.
- D'Angelo, M. A., Gomez-Cavazos, J. S., Mei, A., Lackner, D. H. and Hetzer, M. W. (2012) 'A Change in Nuclear Pore Complex Composition Regulates Cell Differentiation', *Developmental Cell*, 22(2), pp. 446–458. doi: <https://doi.org/10.1016/j.devcel.2011.11.021>.
- D'Arcy, P., Wang, X. and Linder, S. (2015) 'Deubiquitinase inhibition as a cancer therapeutic strategy', *Pharmacol Ther*, 147, pp. 32–54. doi: 10.1016/j.pharmthera.2014.11.002.

- Dai, Y. and Grant, S. (2010) 'New Insights into Checkpoint Kinase 1 in the DNA Damage Response Signaling Network', *Clinical Cancer Research*, 16(2), p. 376 LP-383. doi: 10.1158/1078-0432.CCR-09-1029.
- Dalal, S. N., Yaffe, M. B. and DeCaprio, J. A. (2004) '14-3-3 family members act coordinately to regulate mitotic progression.', *Cell cycle (Georgetown, Tex.)*. United States, 3(5), pp. 672–677.
- Dantuma, N. P., Acs, K. and Luijsterburg, M. S. (2014) 'Should I stay or should I go: VCP/p97-mediated chromatin extraction in the DNA damage response', *Experimental Cell Research*. Elsevier, 329(1), pp. 9–17. doi: 10.1016/j.yexcr.2014.08.025.
- Davis, E. J., Lachaud, C., Appleton, P., Macartney, T. J., *et al.* (2012) 'DVC1 (C1orf124) recruits the p97 protein segregase to sites of DNA damage', *Nature Structural & Molecular Biology*. Nature Publishing Group, 19(11), pp. 1093–1100. doi: 10.1038/nsmb.2394.
- Deb  thune, L., Kohlhagen, G., Grandas, A. and Pommier, Y. (2002) 'Processing of nucleopeptides mimicking the topoisomerase I–DNA covalent complex by tyrosyl-DNA phosphodiesterase', *Nucleic Acids Research*, 30(5), pp. 1198–1204. Available at: <http://dx.doi.org/10.1093/nar/30.5.1198>.
- Deshmukh, A. L., Kumar, C., Singh, D. K., Maurya, P. and Banerjee, D. (2016) 'Dynamics of replication proteins during lagging strand synthesis: A crossroads for genomic instability and cancer', *DNA Repair*. Elsevier B.V., 42, pp. 72–81. doi: 10.1016/j.dnarep.2016.04.010.
- Dewar, J. M., Budzowska, M. and Walter, J. C. (2015) 'The mechanism of DNA replication termination in vertebrates.', *Nature*, 525(7569), pp. 345–50. doi: 10.1038/nature14887.
- Deweese, J. E. and Osheroff, N. (2009) 'The DNA cleavage reaction of topoisomerase II: wolf in

- sheep's clothing', *Nucleic Acids Research*. Oxford University Press, 37(3), pp. 738–748. doi: 10.1093/nar/gkn937.
- Dobrynin, G., Popp, O., Romer, T., Bremer, S., *et al.* (2011) 'Cdc48/p97–Ufd1–Npl4 antagonizes Aurora B during chromosome segregation in HeLa cells', *Journal of Cell Science*, 124(9), p. 1571 LP-1580. Available at: <http://jcs.biologists.org/content/124/9/1571.abstract>.
- Dozier, C., Mazzolini, L., Cenac, C., Froment, C., *et al.* (2017) 'CyclinD-CDK4/6 complexes phosphorylate CDC25A and regulate its stability.', *Oncogene*. England, 36(26), pp. 3781–3788. doi: 10.1038/onc.2016.506.
- Dungrawala, H., Bhat, K. P., Le Meur, R., Chazin, W. J., *et al.* (2017) 'RADX Promotes Genome Stability and Modulates Chemosensitivity by Regulating RAD51 at Replication Forks', *Molecular Cell*. Elsevier, 67(3), p. 374–386.e5. doi: 10.1016/j.molcel.2017.06.023.
- Duxin, J. P., Dewar, J. M., Yardimci, H. and Walter, J. C. (2014) 'Repair of a DNA-protein crosslink by replication-coupled proteolysis', *Cell*. Elsevier Inc., 159(2), pp. 349–357. doi: 10.1016/j.cell.2014.09.024.
- Dynan, W. S. and Tjian, R. (1983) 'Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II', *Cell*, 32(3), pp. 669–680. doi: [https://doi.org/10.1016/0092-8674\(83\)90053-3](https://doi.org/10.1016/0092-8674(83)90053-3).
- Eid, W., Steger, M., El-Shemerly, M., Ferretti, L. P., *et al.* (2010) 'DNA end resection by CtIP and exonuclease 1 prevents genomic instability.', *EMBO reports*. England, 11(12), pp. 962–968. doi: 10.1038/embor.2010.157.
- Fachinetti, D., Bermejo, R., Cocito, A., Minardi, S., *et al.* (2010) 'Replication Termination at Eukaryotic Chromosomes Is Mediated by Top2 and Occurs at Genomic Loci Containing Pausing Elements', *Molecular Cell*. Elsevier, 39(4), pp. 595–605. doi: 10.1016/j.molcel.2010.07.024.

- Falck, J., Coates, J. and Jackson, S. P. (2005) 'Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage', *Nature*, 434(7033), pp. 605–611. Available at: <http://dx.doi.org/10.1038/nature03442>.
- Falck, J., Mailand, N., Syljuasen, R. G., Bartek, J. and Lukas, J. (2001) 'The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis', *Nature*. Macmillan Magazines Ltd., 410(6830), pp. 842–847. Available at: <http://dx.doi.org/10.1038/35071124>.
- Fanconi, G. (1927) 'Familiäre infantile perniziosaartige Anämie (perniziöses Blutbild und Konstitution)', *Jahrbuch für Kinderheilkunde und physische Erziehung (Wien)*, 177, pp. 257–280.
- Fesquet, D., Labbé, J. C., Derancourt, J., Capony, J. P., *et al.* (1993) 'The MO15 gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases (CDKs) through phosphorylation of Thr161 and its homologues.', *The EMBO journal*, 12(8), pp. 3111–21. doi: 10.1016/0168-9525(93)90030-L.
- Fiebigler, E., Hirsch, C., Vyas, J. M., Gordon, E., *et al.* (2004) 'Dissection of the Dislocation Pathway for Type I Membrane Proteins with a New Small Molecule Inhibitor, Eeyarestatin', *Molecular Biology of the Cell*. Edited by P. Walter. The American Society for Cell Biology, 15(4), pp. 1635–1646. doi: 10.1091/mbc.E03-07-0506.
- Fisher, R. P. and Morgan, D. O. (1994) 'A novel cyclin associates with M015/CDK7 to form the CDK-activating kinase', *Cell*, 78(4), pp. 713–724. doi: 10.1016/0092-8674(94)90535-5.
- Franz, A., Ackermann, L. and Hoppe, T. (2014) 'Create and preserve: Proteostasis in development and aging is governed by Cdc48/p97/VCP', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., 1843(1), pp. 205–215. doi: 10.1016/j.bbamcr.2013.03.031.

- Franz, A., Ackermann, L. and Hoppe, T. (2016) 'Ring of Change: CDC48/p97 Drives Protein Dynamics at Chromatin.', *Frontiers in genetics*. Switzerland, 7, p. 73. doi: 10.3389/fgene.2016.00073.
- Franz, A., Orth, M., Pirson, P. A., Sonnevile, R., *et al.* (2011) 'CDC-48/p97 Coordinates CDT-1 Degradation with GINS Chromatin Dissociation to Ensure Faithful DNA Replication', *Molecular Cell*, 44(1), pp. 85–96. doi: 10.1016/j.molcel.2011.08.028.
- Frigola, J., He, J., Kinkelin, K., Pye, V. E., *et al.* (2017) 'Cdt1 stabilizes an open MCM ring for helicase loading', *Nature Communications*, 8, p. 15720. doi: 10.1038/ncomms15720.
- Galaktionov, K. and Beach, D. (1991) 'Specific activation of cdc25 tyrosine phosphatases by B-type cyclins: Evidence for multiple roles of mitotic cyclins', *Cell*, 67(6), pp. 1181–1194. doi: [https://doi.org/10.1016/0092-8674\(91\)90294-9](https://doi.org/10.1016/0092-8674(91)90294-9).
- Galanty, Y., Belotserkovskaya, R., Coates, J. and Jackson, S. P. (2012) 'RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair', *Genes and Development*, 26(11), pp. 1179–1195. doi: 10.1101/gad.188284.112.
- Gao, R., Schellenberg, M. J., Huang, S. Y. N., Abdelmalak, M., *et al.* (2014) 'Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2-DNA and Top2-RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2)', *Journal of Biological Chemistry*, 289(26), pp. 17960–17969. doi: 10.1074/jbc.M114.565374.
- García-Muse, T. and Aguilera, A. (2016) 'Transcription-replication conflicts: how they occur and how they are resolved.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 17(9), pp. 553–63. doi: 10.1038/nrm.2016.88.
- Gardino, A. K. and Yaffe, M. B. (2011) '14-3-3 Proteins As Signaling Integration Points for Cell Cycle Control and Apoptosis', *Seminars in cell & developmental biology*, 22(7), pp. 688–

695. doi: 10.1016/j.semcdb.2011.09.008.
- Gavet, O. and Pines, J. (2010) 'Activation of cyclin B1–Cdk1 synchronizes events in the nucleus and the cytoplasm at mitosis', *The Journal of Cell Biology*, 189(2), p. 247 LP-259. Available at: <http://jcb.rupress.org/content/189/2/247.abstract>.
- Gell, D. and Jackson, S. P. (1999) 'Mapping of protein-protein interactions within the DNA-dependent protein kinase complex', *Nucleic Acids Research*, 27(17), pp. 3494–3502. Available at: <http://dx.doi.org/10.1093/nar/27.17.3494>.
- Glover, T. W., Berger, C., Coyle, J. and Echo, B. (1984) 'DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes.', *Human genetics*. Germany, 67(2), pp. 136–142.
- Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., *et al.* (2011) 'Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles.', *The EMBO journal*, 30(15), pp. 3106–19. doi: 10.1038/emboj.2011.206.
- Gómez-Herreros, F., Zagnoli-Vieira, G., Ntai, I., Martínez-Macías, M. I., *et al.* (2017) 'TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription', *Nature Communications*. London: Nature Publishing Group UK, 8, p. 233. doi: 10.1038/s41467-017-00307-y.
- Gongora, C., Vezzio-Vie, N., Tuduri, S., Denis, V., *et al.* (2011) 'New Topoisomerase I mutations are associated with resistance to camptothecin.', *Molecular cancer*, 10(1), p. 64. doi: 10.1186/1476-4598-10-64.
- Graham, T. G. W., Walter, J. C. and Loparo, J. J. (2017) 'Two-Stage Synapsis of DNA Ends during Non-homologous End Joining', *Molecular Cell*. Elsevier, 61(6), pp. 850–858. doi: 10.1016/j.molcel.2016.02.010.
- Guervilly, J.-H., Takedachi, A., Naim, V., Scaglione, S., *et al.* (2015) 'The SLX4 complex is a SUMO

- E3 ligase that impacts on replication stress outcome and genome stability.’, *Molecular cell*. United States, 57(1), pp. 123–137. doi: 10.1016/j.molcel.2014.11.014.
- Guturi, K. K. N., Bohgaki, M., Bohgaki, T., Srikumar, T., *et al.* (2016) ‘RNF168 and USP10 regulate topoisomerase II α function via opposing effects on its ubiquitylation.’, *Nature communications*, 7, p. 12638. doi: 10.1038/ncomms12638.
- Hagting, A., Karlsson, C., Clute, P., Jackman, M. and Pines, J. (1998) ‘MPF localization is controlled by nuclear export.’, *The EMBO Journal*, 17(14), pp. 4127–4138. doi: 10.1093/emboj/17.14.4127.
- Hanahan, D. and Weinberg, R. A. (2011) ‘Hallmarks of cancer: The next generation’, *Cell*. Elsevier Inc., 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.
- Hänzelmann, P., Buchberger, A. and Schindelin, H. (2011) ‘Hierarchical binding of cofactors to the AAA ATPase p97’, *Structure*, 19(6), pp. 833–843. doi: 10.1016/j.str.2011.03.018.
- Hänzelmann, P. and Schindelin, H. (2011) ‘The structural and functional basis of the p97/valosin-containing protein (VCP)-interacting motif (VIM): Mutually exclusive binding of cofactors to the n-terminal domain of p97’, *Journal of Biological Chemistry*, 286(44), pp. 38679–38690. doi: 10.1074/jbc.M111.274506.
- Hänzelmann, P. and Schindelin, H. (2016) ‘Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions’, *Structure*, 24(1), pp. 140–147. doi: 10.1016/j.str.2015.10.027.
- Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N. and Nakayama, K. I. (2001) ‘U Box Proteins as a New Family of Ubiquitin-Protein Ligases’, *Journal of Biological Chemistry*, 276(35), pp. 33111–33120. doi: 10.1074/jbc.M102755200.
- Helleday, T., Eshtad, S. and Nik-Zainal, S. (2014) ‘Mechanisms underlying mutational signatures in human cancers’, *Nature Reviews Genetics*. Nature Publishing Group, 15(9), pp. 585–

598. doi: 10.1038/nrg3729.

Heller, R. C., Kang, S., Lam, W. M., Chen, S., *et al.* (2011) 'Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases', *Cell*. Elsevier, 146(1), pp. 80–91. doi: 10.1016/j.cell.2011.06.012.

Helmrich, A., Ballarino, M., Nudler, E. and Tora, L. (2013) 'Transcription-replication encounters, consequences and genomic instability.', *Nature structural & molecular biology*, 20(4), pp. 412–8. doi: 10.1038/nsmb.2543.

Helmrich, A., Ballarino, M. and Tora, L. (2011) 'Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes', *Molecular Cell*, 44(6), pp. 966–977. doi: <http://dx.doi.org/10.1016/j.molcel.2011.10.013>.

Henneke, G., Koundrioukoff, S. and Hubscher, U. (2003) 'Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation', *Oncogene*. Nature Publishing Group, 22(28), pp. 4301–4313. Available at: <http://dx.doi.org/10.1038/sj.onc.1206606>.

Hiller, M. M., Finger, A., Schweiger, M. and Wolf, D. H. (1996) 'ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway.', *Science (New York, N.Y.)*. United States, 273(5282), pp. 1725–1728.

Hoffmann, I. (2000) 'The role of Cdc25 phosphatases in cell cycle checkpoints', *Protoplasma*, 211(1), pp. 8–11. doi: 10.1007/BF01279894.

Hofmann, K. and Stoffel, W. (1993) 'TMbase: A Database of Membrane Spanning Protein Segments.', *Biological Chemistry*, 374, p. 166. doi: 10.1515/bchm3.1993.374.1-6.143.

Hunter, L., Bromberg, K. D., Felix, C. A. and Osheroff, N. (2004) '1,4-Benzoquinone Is a Topoisomerase II Poison', *Biochemistry*. American Chemical Society, 43(23), pp.

7563–7574. doi: 10.1021/bi049756r.

Hustedt, N. and Durocher, D. (2017) 'The control of DNA repair by the cell cycle', *Nat Cell Biol.*

Nature Publishing Group, 19(1), pp. 1–9. doi: 10.1038/ncb3452.

Ibarra, A. and Hetzer, M. W. (2015) 'Nuclear pore proteins and the control of genome functions',

Genes and Development, 29(4), pp. 337–349. doi: 10.1101/gad.256495.114.

Interthal, H. and Champoux, J. J. (2011) 'Effects of DNA and protein size on substrate cleavage by

human tyrosyl-DNA phosphodiesterase 1', *Biochemical Journal*, 436(3), p. 559 LP-566.

Available at: <http://www.biochemj.org/content/436/3/559.abstract>.

Ishida, T. and Kinoshita, K. (2007) 'PrDOS: Prediction of disordered protein regions from amino

acid sequence', *Nucleic Acids Research*, 35(SUPPL.2). doi: 10.1093/nar/gkm363.

Itaya, M., Brett, I. C. and Smith, S. O. (2012) 'Synthesis, purification, and characterization of

single helix membrane peptides and proteins for NMR spectroscopy', *Methods in*

Molecular Biology, 831, pp. 333–357. doi: 10.1007/978-1-61779-480-3_18.

Jackson, P. K., Chevalier, S., Philippe, M. and Kirschner, M. W. (1995) 'Early events in DNA

replication require cyclin E and are blocked by p21CIP1.', *The Journal of Cell Biology*,

130(4), p. 755 LP-769. Available at: <http://jcb.rupress.org/content/130/4/755.abstract>.

Jackson, S. P. and Bartek, J. (2009) 'The DNA-damage response in human biology and disease.',

Nature. Nature Publishing Group, 461(7267), pp. 1071–1078. doi:

10.1038/nature08467.

Jackson, S. P., MacDonald, J. J., Lees-Miller, S. and Tjian, R. (1990) 'GC box binding induces

phosphorylation of Sp1 by a DNA-dependent protein kinase', *Cell*, 63(1), pp. 155–

165. doi: [http://dx.doi.org/10.1016/0092-8674\(90\)90296-Q](http://dx.doi.org/10.1016/0092-8674(90)90296-Q).

Jiricny, J. (2006) 'The multifaceted mismatch-repair system', *Nature Reviews Molecular Cell*

- Biology*. Nature Publishing Group, 7, p. 335. Available at:
<http://dx.doi.org/10.1038/nrm1907>.
- Kakarougkas, A., Ismail, A., Katsuki, Y., Freire, R., *et al.* (2013) 'Co-operation of BRCA1 and POH1 relieves the barriers posed by 53BP1 and RAP80 to resection', *Nucleic Acids Research*, 41(22), pp. 10298–10311. Available at: <http://dx.doi.org/10.1093/nar/gkt802>.
- Kakarougkas, A., Ismail, A., Klement, K., Goodarzi, A. A., *et al.* (2013) 'Opposing roles for 53BP1 during homologous recombination', *Nucleic Acids Research*, 41(21), pp. 9719–9731. doi: 10.1093/nar/gkt729.
- Kannouche, P. L., Wing, J. and Lehmann, A. R. (2004) 'Interaction of Human DNA Polymerase η with Monoubiquitinated PCNA: A Possible Mechanism for the Polymerase Switch in Response to DNA Damage', *Molecular Cell*, 14(4), pp. 491–500. doi: [https://doi.org/10.1016/S1097-2765\(04\)00259-X](https://doi.org/10.1016/S1097-2765(04)00259-X).
- Karimian, A., Ahmadi, Y. and Yousefi, B. (2016) 'Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage', *DNA Repair*. Elsevier B.V., 42, pp. 63–71. doi: 10.1016/j.dnarep.2016.04.008.
- Kastan, M. B. and Bartek, J. (2004) 'Cell cycle checkpoints and cancer', *Nature*, 432(7015), pp. 316–323. doi: 10.1038/nature03097.
- Kiiianitsa, K. and Maizels, N. (2013) 'A rapid and sensitive assay for DNA-protein covalent complexes in living cells.', *Nucleic acids research*. England, 41(9), p. e104. doi: 10.1093/nar/gkt171.
- Kim, J.-S. and Raines, R. T. (2008) 'Ribonuclease S-peptide as a carrier in fusion proteins', *Protein Science*, 2(3), pp. 348–356. doi: 10.1002/pro.5560020307.
- Kimple, M. E., Brill, A. L. and Pasker, R. L. (2013) 'Overview of affinity tags for protein purification', *Current Protocols in Protein Science*, (SUPPL.73). doi:

10.1002/0471140864.ps0909s73.

Kloppsteck, P., Ewens, C. A., Förster, A., Zhang, X. and Freemont, P. S. (2012) 'Regulation of p97 in the ubiquitin-proteasome system by the UBX protein-family', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1823(1), pp. 125–129. doi: 10.1016/j.bbamcr.2011.09.006.

Kneusel, R. E., Crowe, J., Wulbeck, M. and Ribbe, J. (1998) 'Procedures for the analysis and purification of his-tagged proteins.', in Reischl, U. (ed.) *Methods in molecular medicine*. Totowa, NJ: Humana Press, pp. 293–308. doi: 10.1385/0-89603-485-2:293.

Komander, D. (2009) 'The emerging complexity of protein ubiquitination', *Biochemical Society Transactions*, 37(5), p. 937 LP-953. Available at: <http://www.biochemsoctrans.org/content/37/5/937.abstract>.

Komander, D., Clague, M. J. and Urbe, S. (2009) 'Breaking the chains: structure and function of the deubiquitinases', *Nat Rev Mol Cell Biol*. Nature Publishing Group, 10(8), pp. 550–563. Available at: <http://dx.doi.org/10.1038/nrm2731>.

Komander, D. and Rape, M. (2012) 'The Ubiquitin Code', *Annual Review of Biochemistry*. Annual Reviews, 81(1), pp. 203–229. doi: 10.1146/annurev-biochem-060310-170328.

Koseoglu, M. M., Graves, L. M. and Marzluff, W. F. (2008) 'Phosphorylation of threonine 61 by cyclin a/Cdk1 triggers degradation of stem-loop binding protein at the end of S phase.', *Molecular and cellular biology*, 28(14), pp. 4469–79. doi: 10.1128/MCB.01416-07.

Koster, D. A., Palle, K., Bot, E. S. M., Bjornsti, M.-A. and Dekker, N. H. (2007) 'Antitumour drugs impede DNA uncoiling by topoisomerase I', *Nature*, 448(7150), pp. 213–217. Available at: <http://dx.doi.org/10.1038/nature05938>.

Kravtsova-Ivantsiv, Y. and Ciechanover, A. (2012) 'Non-canonical ubiquitin-based signals for proteasomal degradation', *Journal of Cell Science*, 125(3), pp. 539–548. doi:

10.1242/jcs.093567.

- Kress, E., Schwager, F., Holtackers, R., Seiler, J., *et al.* (2013) 'The UBXN-2/p37/p47 adaptors of CDC-48/p97 regulate mitosis by limiting the centrosomal recruitment of Aurora A', *Journal of Cell Biology*, 201(4), pp. 559–575. doi: 10.1083/jcb.201209107.
- Kubota, Y., Nash, R. A., Klungland, A., Schär, P., *et al.* (1996) 'Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein.', *The EMBO Journal*, 15(23), pp. 6662–6670. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC452490/>.
- Lai, J. S. and Herr, W. (1992) 'Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations.', *Proceedings of the National Academy of Sciences*, 89(15), pp. 6958–6962. doi: 10.1073/pnas.89.15.6958.
- Lambert, S., Watson, A., Sheedy, D. M., Martin, B. and Carr, A. M. (2017) 'Gross Chromosomal Rearrangements and Elevated Recombination at an Inducible Site-Specific Replication Fork Barrier', *Cell. Elsevier*, 121(5), pp. 689–702. doi: 10.1016/j.cell.2005.03.022.
- Le, L. T. M., Kang, W., Kim, J. Y., Le, O. T. T., *et al.* (2016) 'Structural details of Ufd1 binding to p97 and their functional implications in ER-associated degradation', *PLoS ONE*, 11(9), pp. 1–15. doi: 10.1371/journal.pone.0163394.
- Ledesma, F. C., El Khamisy, S. F., Zuma, M. C., Osborn, K. and Caldecott, K. W. (2009) 'A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage', *Nature*. Macmillan Publishers Limited. All rights reserved, 461(7264), pp. 674–678. Available at: <http://dx.doi.org/10.1038/nature08444>.
- Lee, C. H. and Chung, J. H. (2001) 'The hCds1 (Chk2)-FHA Domain Is Essential for a Chain of Phosphorylation Events on hCds1 That Is Induced by Ionizing Radiation', *Journal of Biological Chemistry*, 276(32), pp. 30537–30541. doi: 10.1074/jbc.M104414200.

- Lee, K. C., Bramley, R. L., Cowell, I. G., Jackson, G. H. and Austin, C. A. (2016) 'Proteasomal inhibition potentiates drugs targeting DNA topoisomerase II', *Biochemical Pharmacology*. Elsevier Inc., 103, pp. 29–39. doi: 10.1016/j.bcp.2015.12.015.
- Lee, K. M., Choi, K. H. and Ouellette, M. M. (2004) 'Use of exogenous hTERT to immortalize primary human cells', *Cytotechnology*, 45(1–2), pp. 33–38. doi: 10.1007/s10616-004-5123-3.
- Lees-Miller, S. P., Chen, Y. R. and Anderson, C. W. (1990) 'Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen.', *Molecular and cellular biology*, 10(12), pp. 6472–81. doi: 10.1128/mcb.10.12.6472.
- Lessel, D., Vaz, B., Halder, S., Lockhart, P. J., *et al.* (2014) 'Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features', *Nat Genet*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 46(11), pp. 1239–1244. Available at: <http://dx.doi.org/10.1038/ng.3103>.
- Li, C., Vassilev, A. and DePamphilis, M. L. (2004) 'Role for Cdk1 (Cdc2)/Cyclin A in Preventing the Mammalian Origin Recognition Complex's Largest Subunit (Orc1) from Binding to Chromatin during Mitosis', *Molecular and Cellular Biology*. American Society for Microbiology, 24(13), pp. 5875–5886. doi: 10.1128/MCB.24.13.5875-5886.2004.
- Lin, C. P., Ban, Y., Lyu, Y. L. and Liu, L. F. (2009) 'Proteasome-dependent processing of topoisomerase I-DNA adducts into DNA double strand breaks at arrested replication forks', *Journal of Biological Chemistry*, 284(41), pp. 28084–28092. doi: 10.1074/jbc.M109.030601.
- Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., *et al.* (2014) 'Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations.', *Nature*

- communications*, 5(11), p. 4767. doi: 10.1038/ncomms5767.
- Lindahl, T. (1993) 'Instability and decay of the primary structure of DNA', *Nature*, 362(6422), pp. 709–715. doi: 10.1038/362709a0.
- Lindahl, T. and Barnes, D. E. (2000) 'Repair of endogenous DNA damage', *Cold Spring Harbor Symposia on Quantitative Biology*, 65, pp. 127–133. doi: 10.1101/sqb.2000.65.127.
- Lippincott-Schwartz, J., Bonifacino, J. S., Yuan, L. C. and Klausner, R. D. (1988) 'Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins.', *Cell*. United States, 54(2), pp. 209–220.
- Liu, F., Stanton, J. J., Wu, Z. and Piwnica-Worms, H. (1997) 'The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine 14 and localizes to the endoplasmic reticulum and Golgi complex.', *Molecular and cellular biology*, 17(2), pp. 571–83. doi: 10.1128/MCB.17.2.571.
- van Loon, B., Hübscher, U. and Maga, G. (2017) 'Living on the Edge: DNA Polymerase Lambda between Genome Stability and Mutagenesis', *Chemical Research in Toxicology*. American Chemical Society. doi: 10.1021/acs.chemrestox.7b00152.
- Lundberg, A. S. and Weinberg, R. A. (1998) 'Functional Inactivation of the Retinoblastoma Protein Requires Sequential Modification by at Least Two Distinct Cyclin-cdk Complexes', *Molecular and Cellular Biology*. American Society for Microbiology, 18(2), pp. 753–761. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC108786/>.
- Maede, Y., Shimizu, H., Fukushima, T., Kogame, T., *et al.* (2014) 'Differential and common DNA repair pathways for topoisomerase I- and II-targeted drugs in a genetic DT40 repair cell screen panel', *Molecular cancer therapeutics*, 13(1), pp. 214–220. doi: 10.1158/1535-7163.MCT-13-0551.

- Maere, S., Heymans, K. and Kuiper, M. (2005) 'BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks', *Bioinformatics*, 21(16), pp. 3448–3449. doi: 10.1093/bioinformatics/bti551.
- Magnaghi, P., D'Alessio, R., Valsasina, B., Avanzi, N., *et al.* (2013) 'Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death', *Nature Chemical Biology*, 9(9), pp. 548–559. doi: 10.1038/nchembio.1313.
- Makela, T. P., Tassan, J. P., Nigg, E. A., Frutiger, S., *et al.* (1994) 'A cyclin associated with the CDK-activating kinase MO15', *Nature*, 371(6494), pp. 254–257. doi: 10.1038/371254a0.
- Mari, P.-O., Florea, B. I., Persengiev, S. P., Verkaik, N. S., *et al.* (2006) 'Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 103(49), pp. 18597–18602. doi: 10.1073/pnas.0609061103.
- Maric, M., Maculins, T., De Piccoli, G. and Labib, K. (2014) 'Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication.', *Science (New York, N.Y.)*, 346(6208), p. 1253596. doi: 10.1126/science.1253596.
- Maric, M., Mukherjee, P., Tatham, M. H., Hay, R. and Labib, K. (2017) 'Ufd1-Npl4 Recruit Cdc48 for Disassembly of Ubiquitylated CMG Helicase at the End of Chromosome Replication', *Cell Reports*. ElsevierCompany., 18(13), pp. 3033–3042. doi: 10.1016/j.celrep.2017.03.020.
- Marnef, A., Cohen, S. and Legube, G. (2017) 'Transcription-Coupled DNA Double-Strand Break Repair: Active Genes Need Special Care', *Journal of Molecular Biology*, 429(9), pp. 1277–1288. doi: 10.1016/j.jmb.2017.03.024.
- Maskey, R. S., Flatten, K. S., Sieben, C. J., Peterson, K. L., *et al.* (2017) 'Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis', *Nucleic Acids*

- Research*, 45(8), pp. 4564–4576. doi: 10.1093/nar/gkx107.
- Massé, E. and Drolet, M. (1999) 'Escherichia coli DNA topoisomerase I inhibits R-loop formation by relaxing transcription-induced negative supercoiling', *Journal of Biological Chemistry*, 274(23), pp. 16659–16664. doi: 10.1074/jbc.274.23.16659.
- Matos, J. and West, S. C. (2014) 'Holliday junction resolution: Regulation in space and time', *DNA Repair*, 19, pp. 176–181. doi: <https://doi.org/10.1016/j.dnarep.2014.03.013>.
- Matsuoka, S., Rotman, G., Ogawa, A., Shiloh, Y., *et al.* (2000) 'Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro.', *Proceedings of the National Academy of Sciences of the United States of America*, 97(19), pp. 10389–94. doi: 10.1073/pnas.190030497.
- Mattaj, I. W. (2004) 'Sorting out the nuclear envelope from the endoplasmic reticulum', *Nat Rev Mol Cell Biol.* Nature Publishing Group, 5(1), pp. 65–69. Available at: <http://dx.doi.org/10.1038/nrm1263>.
- Mazouzi, A., Velimezi, G. and Loizou, J. I. (2014) 'DNA replication stress: Causes, resolution and disease', *Experimental Cell Research*, 329(1), pp. 85–93. doi: 10.1016/j.yexcr.2014.09.030.
- McCracken, A. A. and Brodsky, J. L. (1996) 'Assembly of ER-associated protein degradation in vitro: dependence on cytosol, calnexin, and ATP.', *The Journal of Cell Biology*, 132(3), p. 291 LP-298. Available at: <http://jcb.rupress.org/content/132/3/291.abstract>.
- Medema, R. H., Klompaker, R., Smits, V. A. and Rijksen, G. (1998) 'p21waf1 can block cells at two points in the cell cycle, but does not interfere with processive DNA-replication or stress-activated kinases.', *Oncogene*, 16(4), pp. 431–41. doi: 10.1038/sj.onc.1201558.
- Meerang, M., Ritz, D., Paliwal, S., Garajova, Z., *et al.* (2011) 'The ubiquitin-selective segregase

- VCP/p97 orchestrates the response to DNA double-strand breaks', *Nature Cell Biology*.
Nature Publishing Group, 13(11), pp. 1376–1382. doi: 10.1038/ncb2367.
- Mehnert, M., Sommer, T. and Jarosch, E. (2010) 'ERAD ubiquitin ligases: Multifunctional tools for protein quality control and waste disposal in the endoplasmic reticulum', *BioEssays*, pp. 905–913. doi: 10.1002/bies.201000046.
- Meisenberg, C., Ward, S. E., Schmid, P. and El-Khamisy, S. F. (2014) 'TDP1/TOP1 Ratio as a Promising Indicator for the Response of Small Cell Lung Cancer to Topotecan.', *Journal of cancer science & therapy*. United States, 6(7), pp. 258–267. doi: 10.4172/1948-5956.1000280.
- Mekhail, K. and Moazed, D. (2011) 'The nuclear envelope in genome organization, expression and stability', *Nat. Rev. Mol. Cell Biol.*, 11(5), pp. 317–328. doi: 10.1038/nrm2894.The.
- Méndez, J. and Stillman, B. (2000) 'Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis', *Molecular and Cellular Biology*. American Society for Microbiology, 20(22), pp. 8602–8612. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC102165/>.
- Mentegari, E., Kissova, M., Bavagnoli, L., Maga, G. and Crespan, E. (2016) 'DNA Polymerases λ and β : The Double-Edged Swords of DNA Repair', *Genes*. Edited by R. T. Pomerantz. MDPI, 7(9), p. 57. doi: 10.3390/genes7090057.
- Meyer, H. H., Bug, M. and Bremer, S. (2012) 'Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system', *Nat Cell Biol.* Nature Publishing Group, 14(2), pp. 117–123. doi: 10.1038/ncb2407.
- Meyer, H. H., Shorter, J. G., Seemann, J., Pappin, D. and Warren, G. (2000) 'A complex of mammalian Ufd1 and Npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport

- pathways', *The EMBO Journal*, 19(10), pp. 2181–2192. doi: 10.1093/emboj/19.10.2181.
- Meyer, H. H. and Weihl, C. C. (2014) 'The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis.', *Journal of cell science*, 127(Pt 18), pp. 3877–83. doi: 10.1242/jcs.093831.
- Mirkin, S. M. (2007) 'Expandable DNA repeats and human disease', *Nature*. Nature Publishing Group, 447, p. 932. Available at: <http://dx.doi.org/10.1038/nature05977>.
- Momand, J., Zambetti, G. P., Olson, D. C., George, D. and Levine, A. J. (1992) 'The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation.', *Cell*. United States, 69(7), pp. 1237–1245.
- Moreno, P. S. and Gambus, A. (2015) 'Regulation of Unperturbed DNA Replication by Ubiquitylation', *Genes*. doi: 10.3390/genes6030451.
- Mosbech, A., Gibbs-Seymour, I., Kagias, K., Thorslund, T., *et al.* (2012) 'DVC1 (C1orf124) is a DNA damage–targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks', *Nature Structural & Molecular Biology*. Nature Publishing Group, 1(October), pp. 1–11. doi: 10.1038/nsmb.2395.
- Mueller, P. R., Coleman, T. R., Kumagai, A. and Dunphy, W. G. (1995) 'Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15', *Science*, 270(5233), pp. 86–90. doi: 10.1126/science.270.5233.86.
- Nagata, A., Igarashi, M., Jinno, S., Suto, K. and Okayama, H. (1991) 'An additional homolog of the fission yeast cdc25+ gene occurs in humans and is highly expressed in some cancer cells.', *The New biologist*. United States, 3(10), pp. 959–968.
- Nakanishi, A., Oshida, T., Matsushita, T., Imajoh-Ohmi, S. and Ohnuki, T. (1998) 'Identification of DNA gyrase inhibitor (GyrI) in Escherichia coli', *Journal of Biological Chemistry*, 273(4),

- pp. 1933–1938. doi: 10.1074/jbc.273.4.1933.
- Nalepa, G., Rolfe, M. and Harper, J. W. (2006) 'Drug discovery in the ubiquitin-proteasome system', *Nat Rev Drug Discov*, 5(7), pp. 596–613. Available at: <http://dx.doi.org/10.1038/nrd2056>.
- Nguyen, H. D., Yadav, T., Giri, S., Saez, B., *et al.* (2017) 'Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1', *Molecular Cell*. Elsevier, 65(5), p. 832–847.e4. doi: 10.1016/j.molcel.2017.01.029.
- Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., *et al.* (2011) 'BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair.', *Genes & development*. United States, 25(4), pp. 350–362. doi: 10.1101/gad.2003811.
- Nitiss, J. L. (2009) 'DNA topoisomerase II and its growing repertoire of biological functions.', *Nature reviews. Cancer*, 9(5), pp. 327–337. doi: 10.1038/nrc2608.
- Nordlund, P., Weigelt, J., Bray, J., Gileadi, O., *et al.* (2008) 'Protein production and purification', *Nature Methods*, 5(2), pp. 135–146. doi: 10.1038/nmeth.f.202 [doi].
- Ochi, T., Blackford, A. N., Coates, J., Jhujh, S., *et al.* (2015) 'PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair', *Science*, 347(6218), pp. 185–188. doi: 10.1126/science.1261971.
- Olzmann, J. A. and Kopito, R. R. (2011) 'Lipid droplet formation is dispensable for endoplasmic reticulum-associated degradation.', *The Journal of biological chemistry*, 286(32), pp. 27872–4. doi: 10.1074/jbc.C111.266452.
- Olzmann, J. A., Kopito, R. R. and Christianson, J. C. (2013) 'The mammalian endoplasmic reticulum-associated degradation system', *Cold Spring Harbor Perspectives in Biology*, 5(9). doi: 10.1101/cshperspect.a013185.

- Olzmann, J. A., Richter, C. M. and Kopito, R. R. (2013) 'Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 110(4), pp. 1345–1350. doi: 10.1073/pnas.1213738110.
- Osheroff, N., Shelton, E. R. and Brutlag, D. L. (1983) 'DNA topoisomerase II from *Drosophila melanogaster*. Relaxation of supercoiled DNA.', *The Journal of biological chemistry*, 258(15), pp. 9536–9543.
- Panier, S. and Boulton, S. J. (2014) 'Double-strand break repair: 53BP1 comes into focus.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 15(1), pp. 7–18. doi: 10.1038/nrm3719.
- Parker, L. L., Atherton-Fessler, S., Lee, M. S., Ogg, S., *et al.* (1991) 'Cyclin promotes the tyrosine phosphorylation of p34cdc2 in a wee1+ dependent manner.', *The EMBO Journal*, 10(5), pp. 1255–1263. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC452780/>.
- Pfeiffer, A., Luijsterburg, M. S., Acs, K., Wiegant, W. W., *et al.* (2017) 'Ataxin-3 consolidates the MDC1-dependent DNA double-strand break response by counteracting the SUMO-targeted ubiquitin ligase RNF4', *The EMBO Journal*. Available at: <http://emboj.embopress.org/content/early/2017/03/08/emboj.201695151.abstract>.
- Pommier, Y., Leo, E., Zhang, H. and Marchand, C. (2010) 'DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs', *Chemistry & Biology*, 17(5), pp. 421–433. doi: <https://doi.org/10.1016/j.chembiol.2010.04.012>.
- Pommier, Y., Pourquier, P., Fan, Y. and Strumberg, D. (1998) 'Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme', *Biochimica et Biophysica Acta - Gene Structure and Expression*, pp. 83–106. doi: 10.1016/S0167-4781(98)00129-8.
- Pommier, Y., Sun, Y., Huang, S.-Y. N. and Nitiss, J. L. (2016) 'Roles of eukaryotic topoisomerases

- in transcription, replication and genomic stability.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 17(11), pp. 703–721. doi: 10.1038/nrm.2016.111.
- Poon, R. Y., Yamashita, K., Adamczewski, J. P., Hunt, T. and Shuttleworth, J. (1993) 'The cdc2-related protein p40MO15 is the catalytic subunit of a protein kinase that can activate p33cdk2 and p34cdc2.', *The EMBO journal*, 12(8), pp. 3123–32. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=413578&tool=pmcentrez&rendertype=abstract>.
- Priego Moreno, S., Bailey, R., Campion, N., Herron, S., *et al.* (2014) 'Polyubiquitylation drives replisome disassembly at the termination of DNA replication', *Science*, 346(6208), p. 477 LP-481. doi: 10.1126/science.1253585.
- Puumalainen, M.-R., Lessel, D., Rüthemann, P., Kaczmarek, N., *et al.* (2014) 'Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity', *Nature Communications*, 5, p. 3695. doi: 10.1038/ncomms4695.
- Raines, R. T., McCormick, M., Van Oosbree, T. R. and Mierendorf, R. C. (2000) 'The S.Tag fusion system for protein purification.', *Methods in enzymology*, 326, pp. 362–376. doi: 10.1016/S0076-6879(00)26065-5.
- Ramadan, K., Bruderer, R., Spiga, F. M., Popp, O., *et al.* (2007) 'Cdc48/p97 promotes reformation of the nucleus by extracting the kinase Aurora B from chromatin.', *Nature*, 450(7173), pp. 1258–1262. doi: 10.1038/nature06388.
- Ramadan, K., Halder, S., Wiseman, K. and Vaz, B. (2017) 'Strategic role of the ubiquitin-dependent segregase p97 (VCP or Cdc48) in DNA replication', *Chromosoma*. Chromosoma, 126(1), pp. 17–32. doi: 10.1007/s00412-016-0587-4.
- Ramadan, K., Shevelev, I. and Hübscher, U. (2004) 'The DNA-polymerase-X family: controllers of

- DNA quality?', *Nature reviews. Molecular cell biology*, 5(12), pp. 1038–1043. doi: 10.1038/nrm1530.
- Raman, M., Havens, C. G., Walter, J. C. and Harper, J. W. (2011) 'A Genome-wide Screen Identifies p97 as an Essential Regulator of DNA Damage-Dependent CDT1 Destruction', *Molecular Cell*, 44(1), pp. 72–84. doi: 10.1016/j.molcel.2011.06.036.
- Riemer, A., Dobrynin, G., Dressler, A., Bremer, S., *et al.* (2014) 'The p97-Ufd1-Npl4 ATPase complex ensures robustness of the G(2)/M checkpoint by facilitating CDC25A degradation', *Cell Cycle*. Landes Bioscience, 13(6), pp. 919–927. doi: 10.4161/cc.27779.
- Sadhu, K., Reed, S. I., Richardson, H. and Russell, P. (1990) 'Human homolog of fission yeast cdc25 mitotic inducer is predominantly expressed in G2.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 87(13), pp. 5139–5143.
- Saeki, Y., Kudo, T., Sone, T., Kikuchi, Y., *et al.* (2009) 'Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome', *The EMBO Journal*, 28(4), p. 359 LP-371. Available at: <http://emboj.embopress.org/content/28/4/359.abstract>.
- Saugar, I., Ortiz-Bazán, M. Á. and Tercero, J. A. (2014) 'Tolerating DNA damage during eukaryotic chromosome replication', *Experimental Cell Research*. Elsevier, 329(1), pp. 170–177. doi: 10.1016/j.yexcr.2014.07.009.
- Savitsky, K., Bar-Shira, A., Gilad, S., Rotman, G., *et al.* (1995) 'A single ataxia telangiectasia gene with a product similar to PI-3 kinase', *Science*, 268(5218), pp. 1749–1753. doi: 10.1126/science.7792600.
- van Schendel, R., Roerink, S. F., Portegijs, V., van den Heuvel, S. and Tijsterman, M. (2015) 'Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated

- mutagenesis', *Nature Communications*, 6, p. 7394. doi: 10.1038/ncomms8394.
- Schulz, J., Avci, D., Queisser, M. A., Gutschmidt, A., *et al.* (2017) 'Conserved cytoplasmic domains promote Hrd1 ubiquitin ligase complex formation for ER-associated degradation (ERAD)', *Journal of Cell Science*. Available at:
<http://jcs.biologists.org/content/early/2017/08/18/jcs.206847.abstract>.
- Schwertman, P., Bekker-Jensen, S. and Mailand, N. (2016) 'Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers', *Nat Rev Mol Cell Biol.* Nature Publishing Group, 17(6), pp. 379–394. doi: 10.1038/nrm.2016.58.
- Shi, W., Ma, Z., Willers, H., Akhtar, K., *et al.* (2008) 'Disassembly of MDC1 foci is controlled by ubiquitin-proteasome-dependent degradation', *Journal of Biological Chemistry*, 283(46), pp. 31608–31616. doi: 10.1074/jbc.M801082200.
- Shibata, A. (2017) 'Regulation of repair pathway choice at two-ended DNA double-strand breaks', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. doi: 10.1016/j.mrfmmm.2017.07.011.
- Shima, N. and Pederson, K. D. (2017) 'Dormant origins as a built-in safeguard in eukaryotic DNA replication against genome instability and disease development', *DNA Repair*, 56, pp. 166–173. doi: <https://doi.org/10.1016/j.dnarep.2017.06.019>.
- Singleton, B. K., Torres-Arzayus, M. I., Rottinghaus, S. T., Taccioli, G. E. and Jeggo, P. A. (1999) 'The C Terminus of Ku80 Activates the DNA-Dependent Protein Kinase Catalytic Subunit', *Molecular and cellular biology*, 19(5), pp. 3267–3277. doi: 10.1128/MCB.19.5.3267.
- Sinha, B. K., Kumar, A. and Mason, R. P. (2017) 'Nitric oxide inhibits ATPase activity and induces resistance to topoisomerase II-poisons in human MCF-7 breast tumor cells', *Biochemistry and Biophysics Reports*, 10(Supplement C), pp. 252–259. doi:
<https://doi.org/10.1016/j.bbrep.2017.04.011>.

- Sirbu, B. M., Couch, F. B. and Cortez, D. (2012) 'Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA.', *Nature protocols*. England, 7(3), pp. 594–605. doi: 10.1038/nprot.2012.010.
- Solomon, M. J., Harper, J. W. and Shuttleworth, J. (1993) 'CAK, the p34cdc2 activating kinase, contains a protein identical or closely related to p40MO15.', *The EMBO journal*, 12(8), pp. 3133–3142. doi: 10.1016/0168-9525(93)90032-D.
- Song, C., Wang, Q., Song, C., Lockett, S. J., *et al.* (2015) 'Nucleocytoplasmic shuttling of valosin-containing protein (VCP/p97) regulated by its N domain and C-terminal region', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., 1853(1), pp. 222–232. doi: 10.1016/j.bbamcr.2014.10.019.
- Sørensen, H. P., Sperling-Petersen, H. U. and Mortensen, K. K. (2003) 'Production of recombinant thermostable proteins expressed in *Escherichia coli*: Completion of protein synthesis is the bottleneck', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 786(1–2), pp. 207–214. doi: 10.1016/S1570-0232(02)00689-X.
- Soto, M., Raaijmakers, J. A., Bakker, B., Spierings, D. C. J., *et al.* (2017) 'p53 Prohibits Propagation of Chromosome Segregation Errors that Produce Structural Aneuploidies', *Cell Reports*. Elsevier, 19(12), pp. 2423–2431. doi: 10.1016/j.celrep.2017.05.055.
- Stillman, B. (2015) 'Reconsidering DNA Polymerases at the Replication Fork in Eukaryotes', *Molecular Cell*. Elsevier Inc., 59(2), pp. 139–141. doi: 10.1016/j.molcel.2015.07.004.
- Stingele, J., Bellelli, R., Alte, F., Hewitt, G., *et al.* (2016) 'Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN', *Molecular Cell*. Elsevier Inc., 64(4), pp. 688–703. doi: 10.1016/j.molcel.2016.09.031.
- Strick, R., Strissel, P. L., Borgers, S., Smith, S. L. and Rowley, J. D. (2000) 'Dietary bioflavonoids

- induce cleavage in the MLL gene and may contribute to infant leukemia.’, *Proceedings of the National Academy of Sciences of the United States of America*. United States, 97(9), pp. 4790–4795. doi: 10.1073/pnas.070061297.
- Sutherland, G. R., Jacky, P. B., Baker, E. and Manuel, A. (1983) ‘Heritable fragile sites on human chromosomes. X. New folate-sensitive fragile sites: 6p23, 9p21, 9q32, and 11q23.’, *American Journal of Human Genetics*, 35(3), pp. 432–437. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1685660/>.
- Swanson, R., Locher, M. and Hochstrasser, M. (2001) ‘A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and MafK repressor degradation’, *Genes & Development*. Cold Spring Harbor Laboratory Press, 15(20), pp. 2660–2674. doi: 10.1101/gad.933301.
- Tang, X., Yuan, F. and Guo, K. (2014) ‘Repair of radiation damage of U2OS osteosarcoma cells is related to DNA-dependent protein kinase catalytic subunit (DNA-PKcs) activity’, *Molecular and Cellular Biochemistry*, 390(1–2), pp. 51–59. doi: 10.1007/s11010-013-1955-5.
- Tassan, J. P., Jaquenoud, M., Fry, a M., Frutiger, S., *et al.* (1995) ‘In vitro assembly of a functional human CDK7-cyclin H complex requires MAT1, a novel 36 kDa RING finger protein.’, *The EMBO journal*, 14(22), pp. 5608–17. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394676&tool=pmcentrez&rendertype=abstract>.
- Taylor, W. R. and Stark, G. R. (2001) ‘Regulation of the G2/M transition by p53.’, *Oncogene*, 20(15), pp. 1803–1815. doi: 10.1038/sj.onc.1204252.
- Toledo, L., Neelsen, K. J. and Lukas, J. (2017) ‘Replication Catastrophe: When a Checkpoint Fails because of Exhaustion’, *Molecular Cell*. Elsevier Inc., 66(6), pp. 735–749. doi:

10.1016/j.molcel.2017.05.001.

- Torrecilla, I., Oehler, J. and Ramadan, K. (2017) 'The role of ubiquitin-dependent segregase p97 (VCP or Cdc48) in chromatin dynamics after DNA double strand breaks', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1731). Available at: <http://rstb.royalsocietypublishing.org/content/372/1731/20160282.abstract>.
- Tsutakawa, S. E., Thompson, M. J., Arvai, A. S., Neil, A. J., *et al.* (2017) 'Phosphate steering by Flap Endonuclease 1 promotes 5'-flap specificity and incision to prevent genome instability', *Nature Communications*, 8, p. 15855. doi: 10.1038/ncomms15855.
- Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., *et al.* (2009) 'Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription', *Nature cell biology*. Nature Publishing Group, 11(11), pp. 1315–1324. doi: 10.1038/ncb1984.
- Vaz, B., Halder, S. and Ramadan, K. (2013) 'Role of p97/VCP (Cdc48) in genome stability.', *Frontiers in genetics*, 4(April), p. 60. doi: 10.3389/fgene.2013.00060.
- Vaz, B., Popovic, M., Newman, J. A., Fielden, J., *et al.* (2016) 'Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair', *Molecular Cell*. Elsevier Inc., 64(4), pp. 704–719. doi: 10.1016/j.molcel.2016.09.032.
- Vijayaraghavan, S., Karakas, C., Doostan, I., Chen, X., *et al.* (2017) 'CDK4/6 and autophagy inhibitors synergistically induce senescence in Rb positive cytoplasmic cyclin E negative cancers.', *Nature communications*. England, 8, p. 15916. doi: 10.1038/ncomms15916.
- Villani, G. and Tanguy Le Gac, N. (2014) 'DNA Repair Polymerases', in Bell, E. (ed.) *Molecular Life Sciences: An Encyclopedic Reference*. New York, NY: Springer New York, pp. 1–13. doi: 10.1007/978-1-4614-6436-5_61-1.
- Vos, S. M., Tretter, E. M., Schmidt, B. H. and Berger, J. M. (2011) 'All tangled up: how cells direct,

- manage and exploit topoisomerase function.', *Nature reviews. Molecular cell biology*.
Nature Publishing Group, 12(12), pp. 827–41. doi: 10.1038/nrm3228.
- Wang, J. C. (2002) 'Cellular roles of DNA topoisomerases: a molecular perspective.', *Nature reviews. Molecular cell biology*, 3(6), pp. 430–440. doi: 10.1038/nrm831.
- Wang, Q., Li, A. and Ye, Y. (2008) 'Inhibition of p97-dependent protein degradation by Eeyarestatin I', *Journal of Biological Chemistry*, 283(12), pp. 7445–7454. doi: 10.1074/jbc.M708347200.
- Wang, Q., Mora-Jensen, H., Weniger, M. A., Perez-Galan, P., *et al.* (2009) 'ERAD inhibitors integrate ER stress with an epigenetic mechanism to activate BH3-only protein NOXA in cancer cells', *Proceedings of the National Academy of Sciences*, 106(7), pp. 2200–2205. doi: 10.1073/pnas.0807611106.
- Wang, Q., Shinkre, B. A., Lee, J. gu, Weniger, M. A., *et al.* (2010) 'The ERAD inhibitor eeyarestatin I is a bifunctional compound with a membrane-binding domain and a p97/VCP inhibitory group', *PLoS ONE*, 5(11). doi: 10.1371/journal.pone.0015479.
- Wang, Y., Knudsen, B. R., Bjergbæk, L., Westergaard, O. and Andersen, A. H. (1999) 'Stimulated Activity of Human Topoisomerases II α and II β on RNA-containing Substrates', *Journal of Biological Chemistry*, 274(32), pp. 22839–22846. doi: 10.1074/jbc.274.32.22839.
- Waugh, D. S. (2005) 'Making the most of affinity tags', *Trends in Biotechnology*, 23(6), pp. 316–320. doi: 10.1016/j.tibtech.2005.03.012.
- Weinberg, R. A. (1995) 'The retinoblastoma protein and cell cycle control', *Cell*, 81(3), pp. 323–330. doi: [http://dx.doi.org/10.1016/0092-8674\(95\)90385-2](http://dx.doi.org/10.1016/0092-8674(95)90385-2).
- de Winter, J. P. and Joenje, H. (2009) 'The genetic and molecular basis of Fanconi anemia', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 668(1), pp. 11–19. doi: <https://doi.org/10.1016/j.mrfmmm.2008.11.004>.

- Woo, R. A. and Poon, R. Y. C. (2003) 'Cyclin-dependent kinases and S phase control in mammalian cells.', *Cell cycle (Georgetown, Tex.)*. United States, 2(4), pp. 316–324.
- Wu, J. and Liu, L. F. (1997) 'Processing of topoisomerase I cleavable complexes into DNA damage by transcription', *Nucleic Acids Research*, 25(21), pp. 4181–4186. doi: 10.1093/nar/25.21.4181.
- Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., *et al.* (2009) 'Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation', *Cell*, 137(1), pp. 133–145. doi: 10.1016/j.cell.2009.01.041.
- Yadav, P., Harcy, V., Argueso, J. L., Dominska, M., *et al.* (2014) 'Topoisomerase I Plays a Critical Role in Suppressing Genome Instability at a Highly Transcribed G-Quadruplex-Forming Sequence', *PLoS Genetics*, 10(12). doi: 10.1371/journal.pgen.1004839.
- Yang, S. W., Burgin, A. B., Huizenga, B. N., Robertson, C. A., *et al.* (1996) 'A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases.', *Proceedings of the National Academy of Sciences of the United States of America*, 93(21), pp. 11534–11539. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC38092/>.
- Yau, R. and Rape, M. (2016) 'The increasing complexity of the ubiquitin code', *Nature Cell Biology*. Nature Publishing Group, 18(6), pp. 579–586. doi: 10.1038/ncb3358.
- Ye, Y., Meyer, H. H. and Rapoport, T. A. (2001) 'The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol', *Nature*. Macmillan Magazines Ltd., 414, p. 652. Available at: <http://dx.doi.org/10.1038/414652a>.
- Ye, Y., Meyer, H. H. and Rapoport, T. A. (2003) 'Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: Dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains', *Journal of Cell Biology*, 162(1), pp. 71–84. doi: 10.1083/jcb.200302169.

- Ye, Y., Shibata, Y., Kikkert, M., van Voorden, S., *et al.* (2005) 'Recruitment of the p97 ATPase and ubiquitin ligases to the site of retrotranslocation at the endoplasmic reticulum membrane', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 102(40), pp. 14132–14138. doi: 10.1073/pnas.0505006102.
- Yekezare, M., Gómez-González, B. and Diffley, J. F. X. (2013) 'Controlling DNA replication origins in response to DNA damage – inhibit globally, activate locally', *Journal of Cell Science*, 126(6), p. 1297 LP-1306. Available at: <http://jcs.biologists.org/content/126/6/1297.abstract>.
- Yeo, C. Q. X., Alexander, I., Lin, Z., Lim, S., *et al.* (2016) 'p53 Maintains Genomic Stability by Preventing Interference between Transcription and Replication', *Cell Reports*. The Authors, 15(1), pp. 132–146. doi: 10.1016/j.celrep.2016.03.011.
- Yeung, H. O., Kloppsteck, P., Niwa, H., Isaacson, R. L., *et al.* (2008) 'Insights into adaptor binding to the AAA protein p97.', *Biochemical Society transactions*, 36(Pt 1), pp. 62–67. doi: 10.1042/BST0360062.
- Zamiri, B., Mirceta, M., Bomsztyk, K., Macgregor, R. B. and Pearson, C. E. (2015) 'Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)₈•(GGCCCC)₈ repeat: Effect of CpG methylation', *Nucleic Acids Research*, 43(20), pp. 10055–10064. doi: 10.1093/nar/gkv1008.
- Zeng, Z., Cortes-Ledesma, F., El Khamisy, S. F. and 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.', *The Journal of biological chemistry*. United States, 286(1), pp. 403–409. doi: 10.1074/jbc.M110.181016.
- Zeng, Z., Sharma, A., Ju, L., Murai, J., *et al.* (2012) 'TDP2 promotes repair of topoisomerase I-

- mediated DNA damage in the absence of TDP1.', *Nucleic acids research*. England, 40(17), pp. 8371–8380. doi: 10.1093/nar/gks622.
- Zhai, Y., Li, N., Jiang, H., Huang, X., *et al.* (2017) 'Unique Roles of the Non-identical MCM Subunits in DNA Replication Licensing', *Molecular Cell*, 67(2), pp. 168–179. doi: <http://dx.doi.org/10.1016/j.molcel.2017.06.016>.
- Zhang, F., Fan, Q., Ren, K. and Andreassen, P. R. (2009) 'PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2.', *Molecular cancer research : MCR*. United States, 7(7), pp. 1110–1118. doi: 10.1158/1541-7786.MCR-09-0123.
- Zhang, Q., Karnak, D., Tan, M., Lawrence, T. S., *et al.* (2016) 'FBXW7 Facilitates Nonhomologous End-Joining via K63-Linked Polyubiquitylation of XRCC4', *Molecular Cell*. Elsevier Inc., 61(3), pp. 419–433. doi: 10.1016/j.molcel.2015.12.010.
- Zhou, H. J., Wang, J., Yao, B., Wong, S., *et al.* (2015) 'Discovery of a First-in-Class, Potent, Selective, and Orally Bioavailable Inhibitor of the p97 AAA ATPase (CB-5083)', *Journal of Medicinal Chemistry*, 58(24), pp. 9480–9497. doi: 10.1021/acs.jmedchem.5b01346.
- Zlotorynski, E., Rahat, A., Skaug, J., Ben-Porat, N., *et al.* (2003) 'Molecular basis for expression of common and rare fragile sites.', *Molecular and cellular biology*. United States, 23(20), pp. 7143–7151.
- Zou, L. and Elledge, S. J. (2003) 'Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes', *Science*, 300(5625), pp. 1542–1548. doi: 10.1126/science.1083430.
- Zucca, E., Bertoletti, F., Wimmer, U., Ferrari, E., *et al.* (2013) 'Silencing of human DNA polymerase λ causes replication stress and is synthetically lethal with an impaired S phase checkpoint', *Nucleic Acids Research*. Oxford University Press, 41(1), pp. 229–241. doi: 10.1093/nar/gks1016.

9. APPENDIX

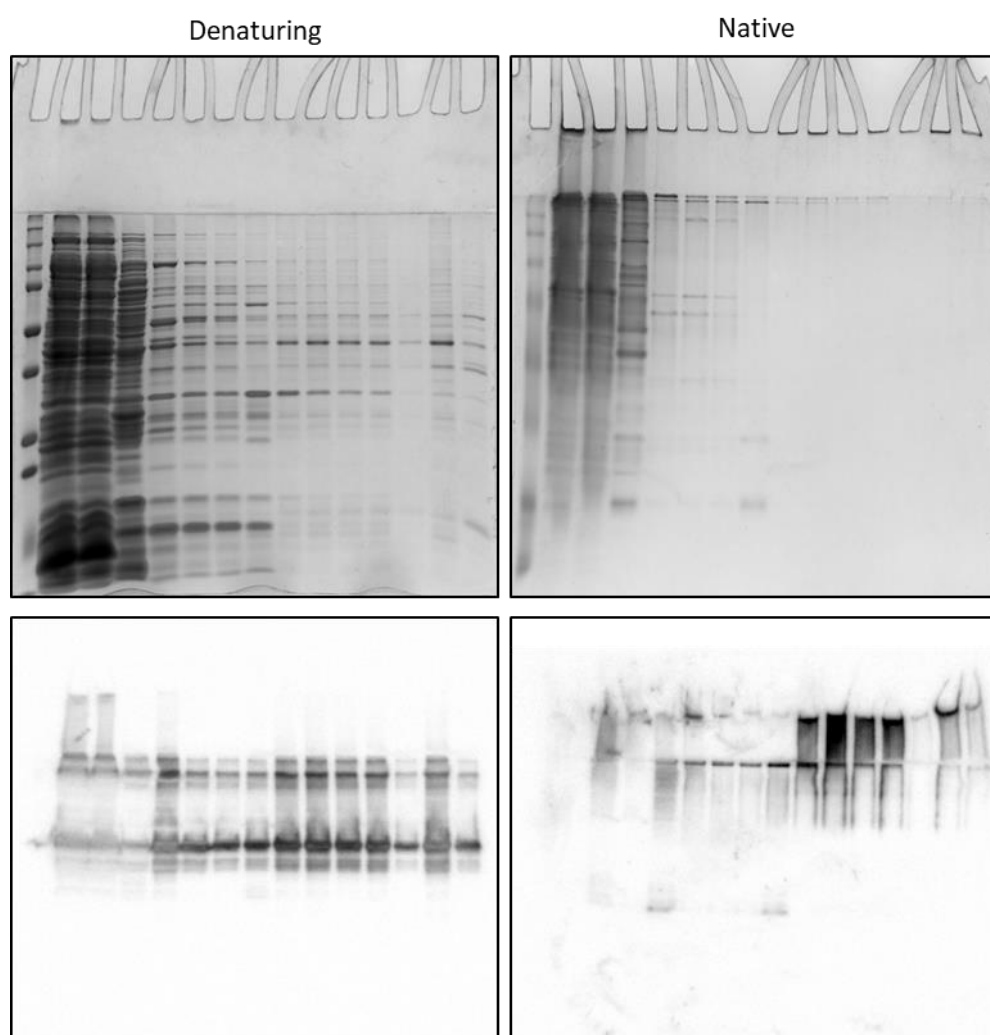


Figure 9-1 IREVIP1 cHis enters denaturing but not native gels
IREVIP1 cHis purification fractions were run on denaturing (left) or native (right) polyacrylamide gels. Samples were analysed by Coomassie (top) and WB (bottom). IREVIP1 cHis was able to enter denaturing gels yet unable to enter native gels.

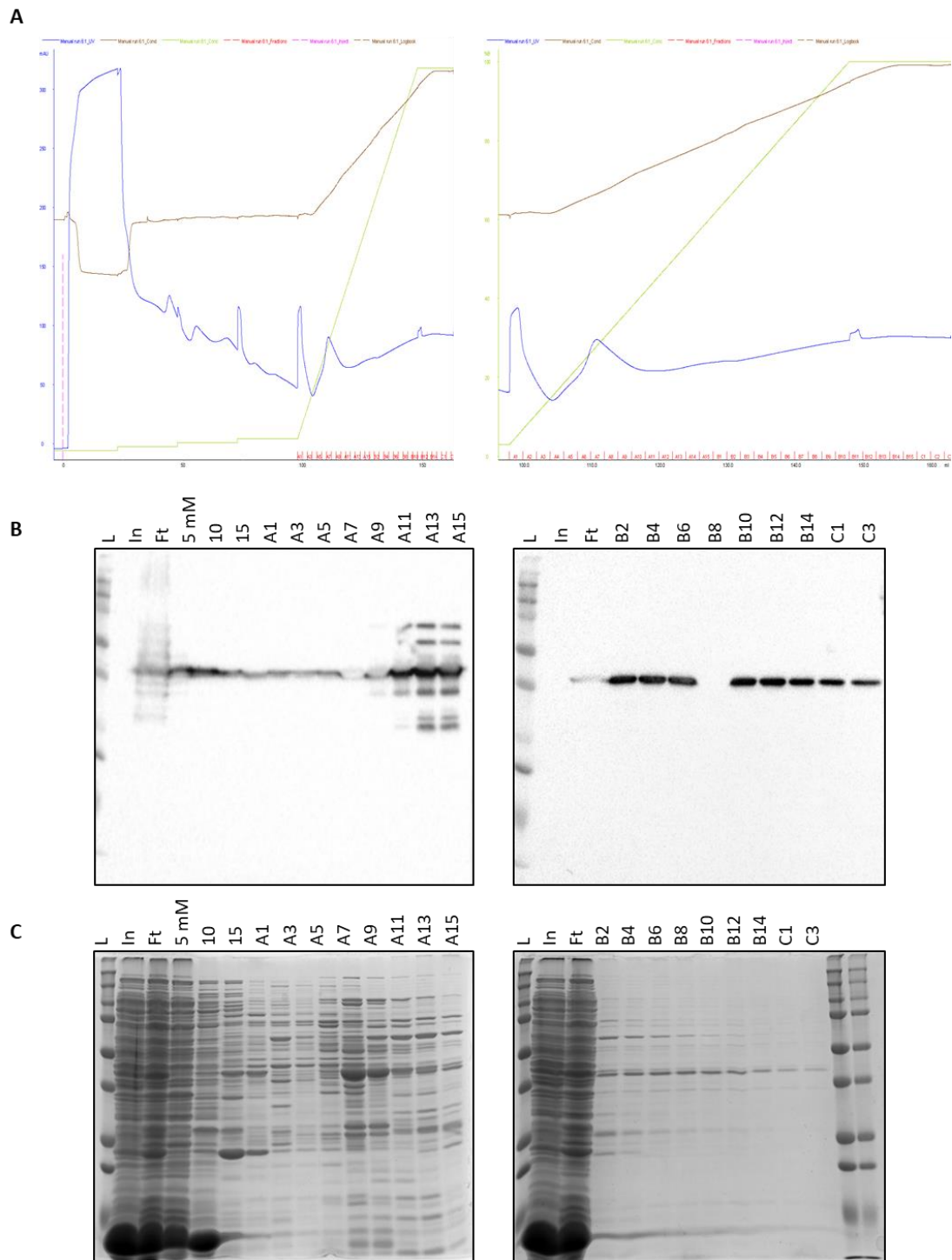


Figure 9-2 IREVIP1 NTD cHis binds poorly, giving low elutions
 IREVIP1 NTD cHis was purified using the 5 mL FF crude columns on the ÄKTA FPLC system. A) ÄKTA FPLC read out. Blue – UV, Green – Concentration of solution B (1 M Imidazole), Brown – conductivity. Left panel – full view, right panel – zoomed view of elution fraction. B) WB analysis of purification fraction. IREVIP1 is lost throughout the purification, but the main bulk appears from fraction A11 onwards. This does not correspond to the blue elution peak in A. C) Coomassie analysis of purification fractions. IREVIP1 is highly contaminated.

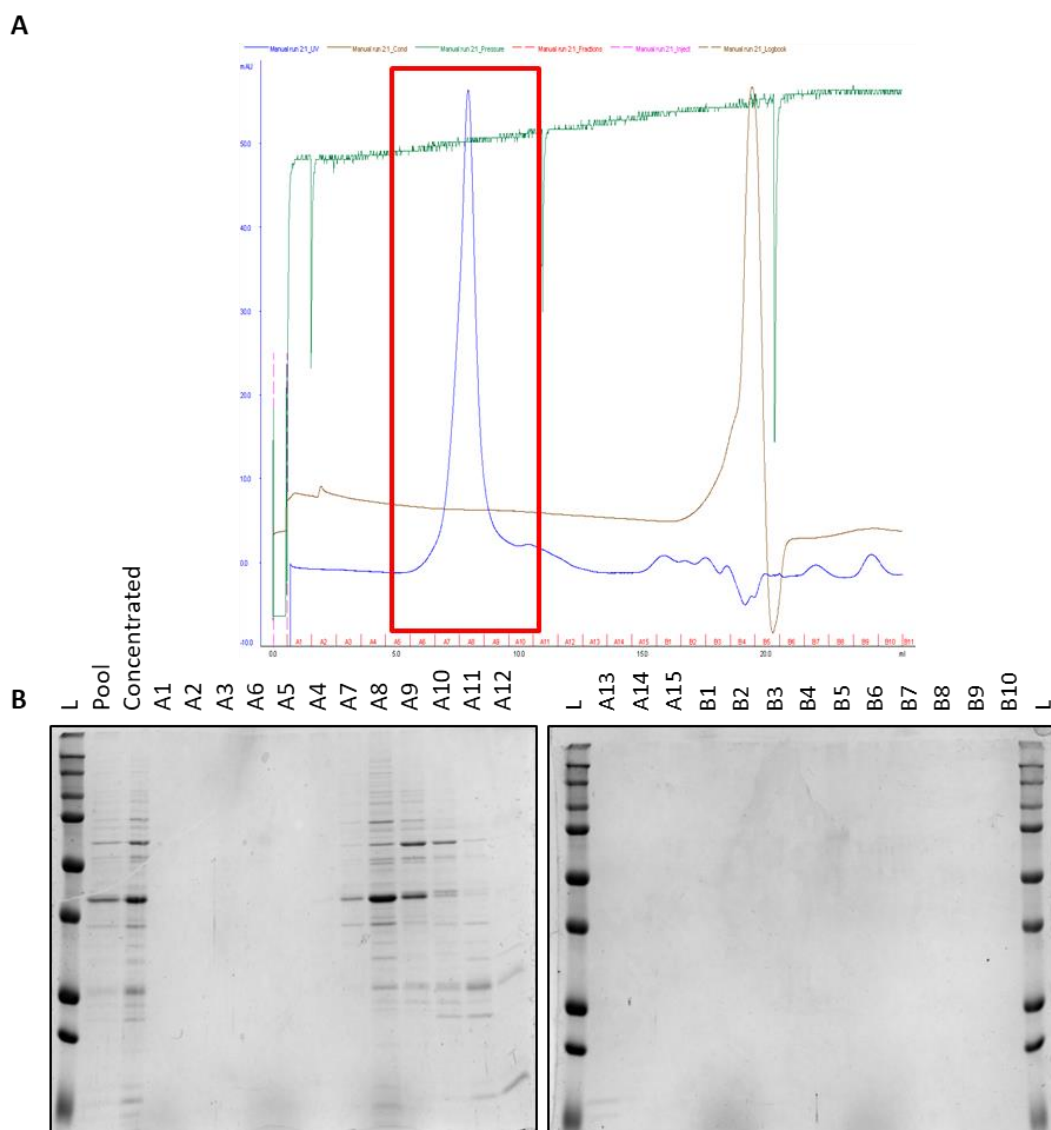


Figure 9-3 IREVIP1 NTA cHis is found in the deal volume of Superdex 75 gel filtration column
 IREVIP1 NTA cHis was run through Superdex 75 gel filtration columns as an additional purification step. A) ÄKTA FPLC read out. Blue – UV, Green – Concentration of solution B (none), Brown – conductivity. B) IREVIP1 fraction analysis by Coomassie. IREVIP1 appears in fraction A8, which corresponds to the blue peak in A. However, this is the 'dead volume' of the column where a protein the size of IREVIP1, 34 kDa, should not be appearing.

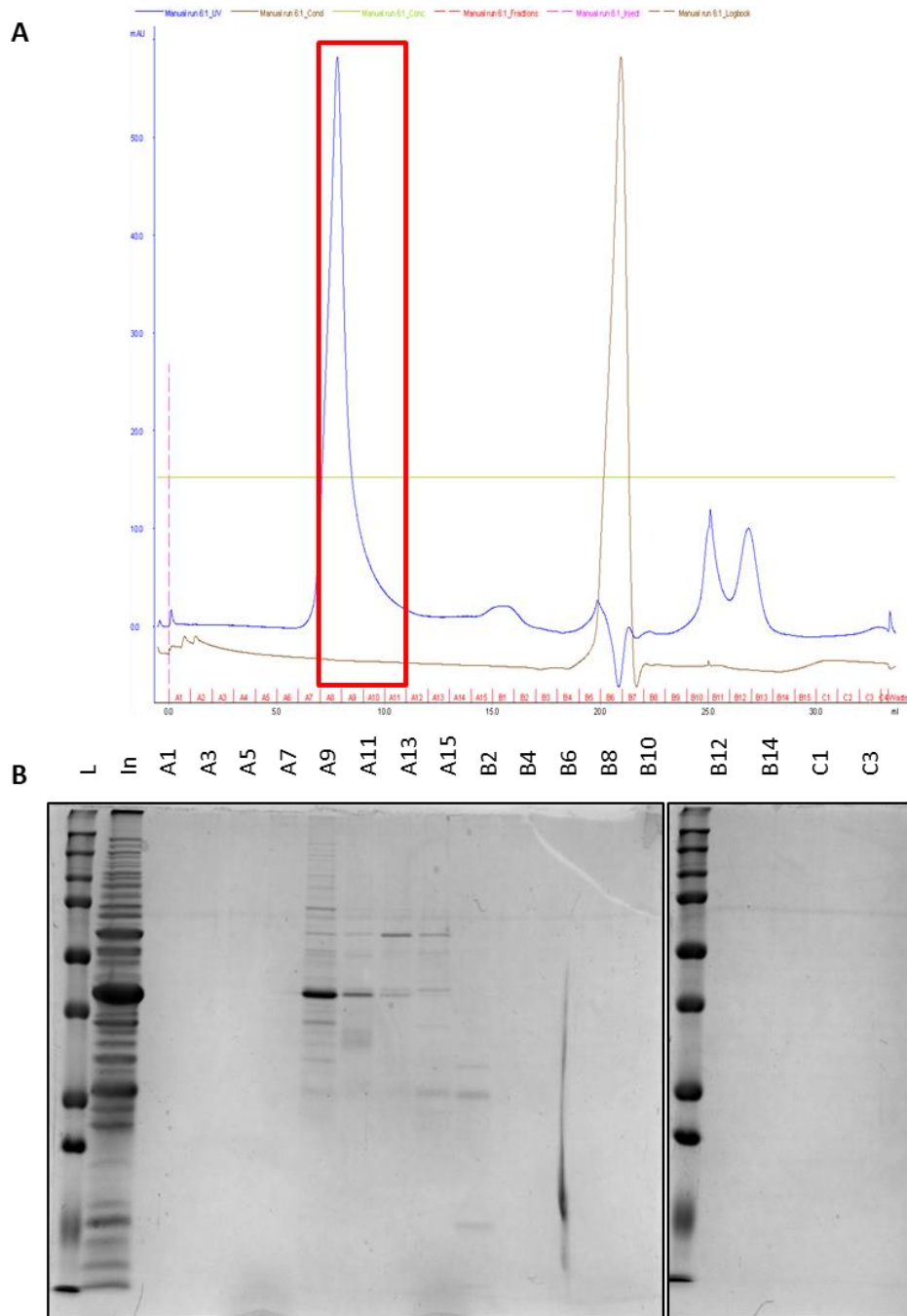


Figure 9-4 IREVIP1 NTA cHis is found in the dead volume of Superdex 200 gel filtration column
 IREVIP1 NTA cHis was run through Superdex 200 gel filtration columns as an additional purification step. A) ÄKTA FPLC read out. Blue – UV, Green – Concentration of solution B (none), Brown – conductivity. B) IREVIP1 fraction analysis by Coomassie. IREVIP1 appears in fraction A9, which corresponds to the blue peak in A. However, this is the ‘dead volume’ of the column where a protein the size of IREVIP1, 34 kDa, should not be appearing.

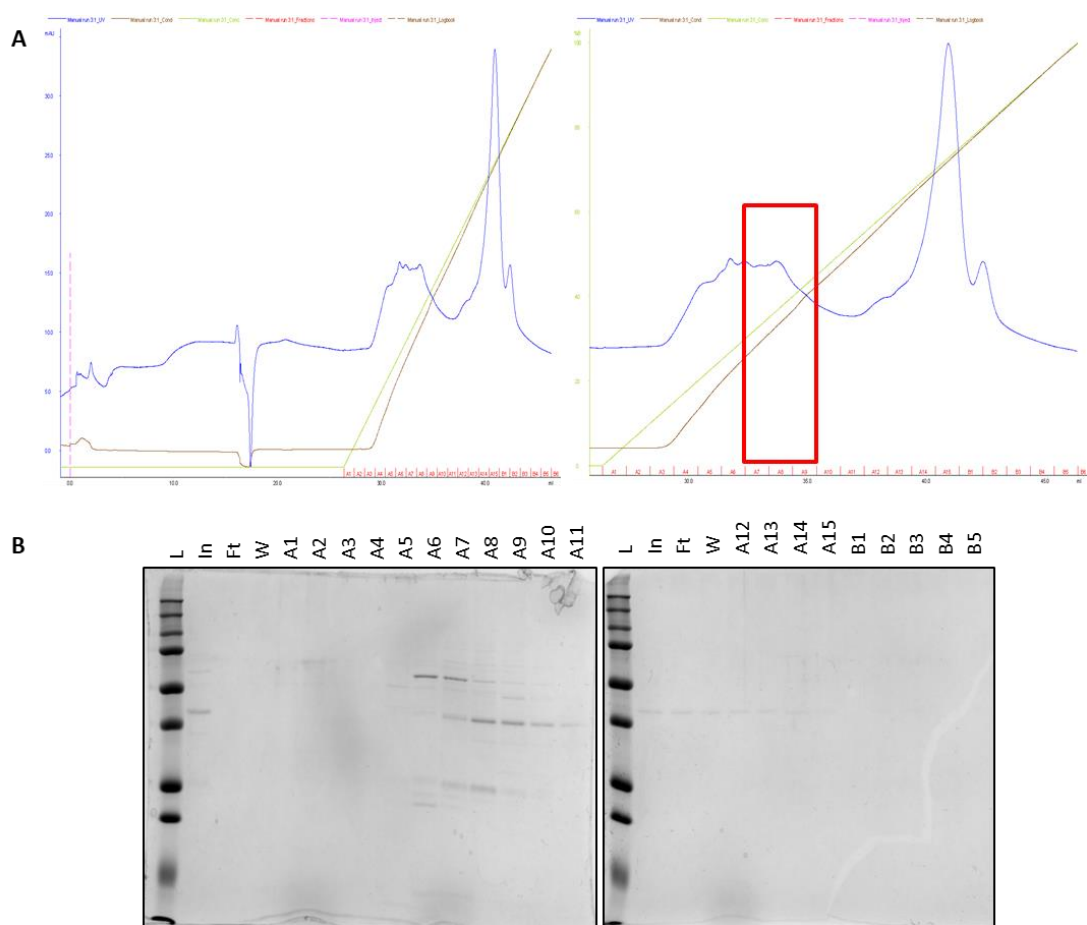


Figure 9-5 IREVIP1 NTA cHis is lost during MonoQ Ion Exchange
IREVIP1 NTA cHis was run through a MonoQ Ion Exchange column as an additional purification step. A) ÄKTA FPLC read out. Blue – UV, Green – Concentration of solution B (1 M NaCl), Brown – conductivity. B) IREVIP1 fraction analysis by Coomassie. IREVIP1 appears in fractions A7-A11, which corresponds to the blue peak in A (right panel). However, the remaining amount of protein is very low.

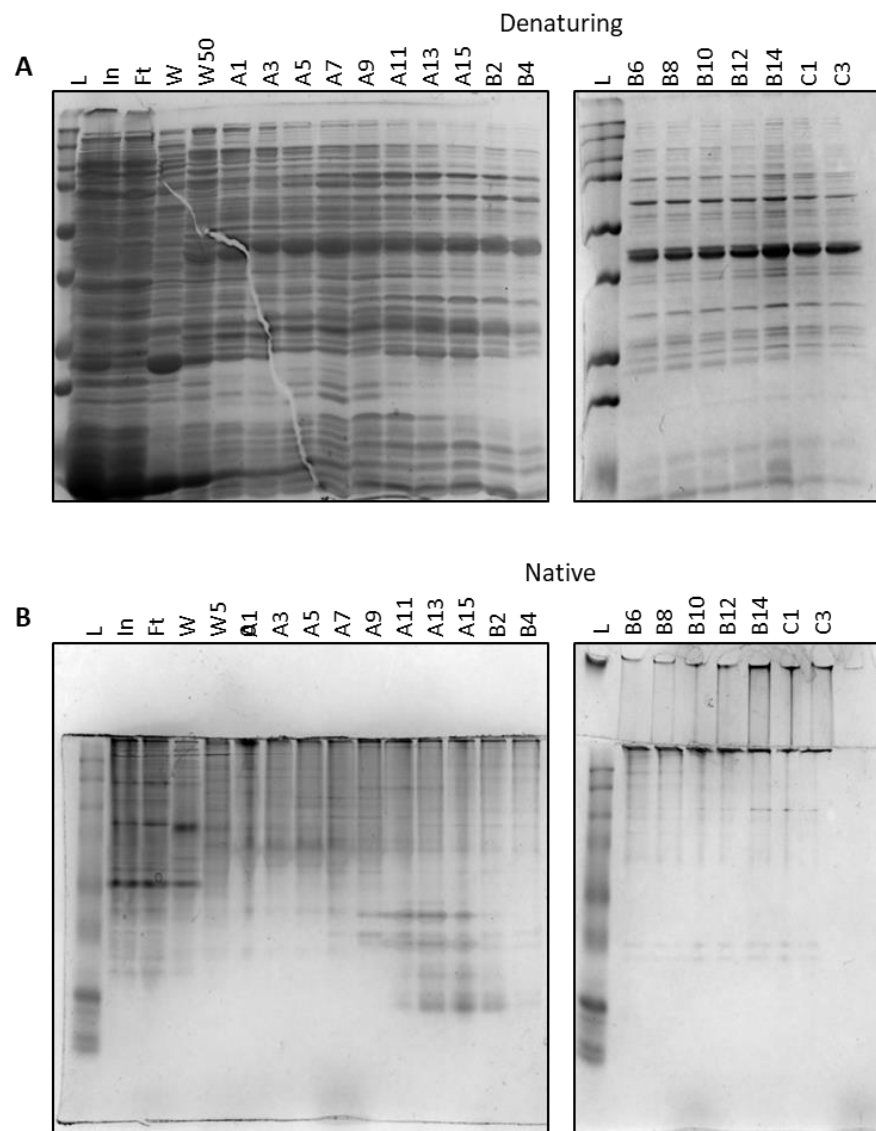


Figure 9-6 IREVIP1 NTA cHis enters denaturing but not native gels
 IREVIP1 NTA cHis purification fractions were run on denaturing (top) or native (bottom) polyacrylamide gels. Samples were analysed by Coomassie. IREVIP1 cHis was able to enter denaturing gels yet unable to enter native gels.

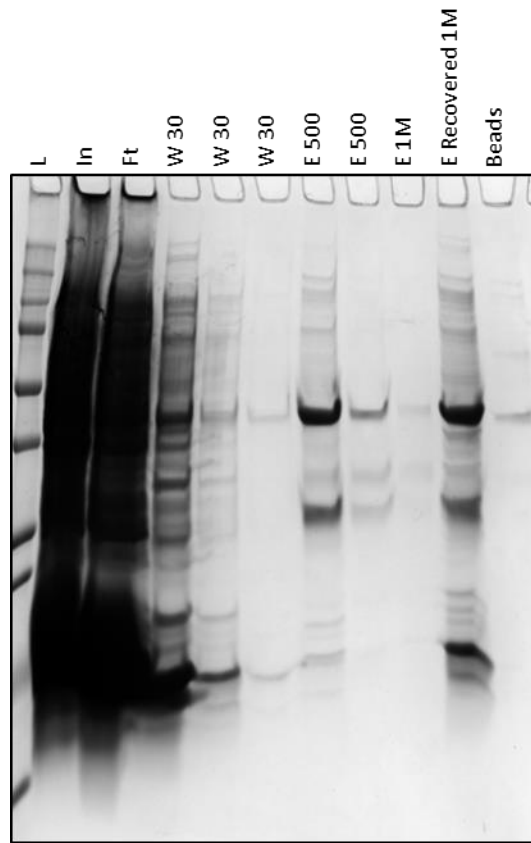


Figure 9-7 IREVIP1 NTA cHis purification under denaturing conditions.
 IREVIP1 NTA cHis was purified in batch under denaturing conditions (8 M urea) using Ni-NTA beads. The beads were also washed with 1% Triton X-100, 800 mM NaCl and 30 mM imidazole. Refolding was carried out on the beads prior to elution with 500 mM and 1 M imidazole. A proportion of beads were lost in the washes, however, these were recovered by centrifugation and subsequently eluted and pooled. Fractions were analysed by Coomassie.

9.1 Publications

During my DPhil, I have contributed towards the following publications:

Ramadan, Kristijan, Swagata Halder, **Katherine Wiseman**, and Bruno Vaz. 2017. "Strategic Role of the Ubiquitin-Dependent Segregase p97 (VCP or Cdc48) in DNA Replication." *Chromosoma* 126 (1). Chromosoma: 17–32. doi:10.1007/s00412-016-0587-4.