CHARACTERIZATION OF APLF IN NON-HOMOLOGOUS END-JOINING

BY

PURNATA V. SHIRODKAR

A thesis submitted in conformity with the requirements for the degree of Master of Science

Graduate Department of Medical Biophysics,
University of Toronto

© Copyright by Purnata V. Shirodkar (2011)
Characterization of APLF in Non-homologous End-joining

Purnata V. Shirodkar  
Master of Science  
Department of Medical Biophysics,  
University of Toronto  
2011

ABSTRACT

APLF (Aprataxin and Polynucleotide kinase-Like Factor), a novel protein with a forkhead-associated (FHA) domain and two poly(ADP-ribose)-binding zinc fingers (PBZ), interacts with core non-homologous end-joining (NHEJ) repair factors, Ku and XRCC4-DNA ligase IV, and facilitates NHEJ. However, how APLF functions in NHEJ is undefined. This thesis demonstrates that the Ku-binding domain on APLF is mapped to amino acid residues 180-200, where conserved amino acid residue W189 strongly contributes to the APLF-Ku interaction. Remarkably, the APLF-Ku interaction is involved in the nuclear localization of APLF. Furthermore, we demonstrate that the N-terminal region (amino acids 1-200), containing the XRCC4-Ligase IV and Ku binding domains, is required for APLF-dependent NHEJ. Collectively, these findings suggest that Ku contributes to APLF nuclear localization, and that once APLF is retained in the nucleus, the N-terminal portion of APLF, which facilitates interactions with the core NHEJ proteins Ku and XRCC4-DNA ligase IV, is required for efficient NHEJ.
ACKNOWLEDGEMENTS

First and foremost, I offer my sincerest gratitude to my supervisor, Dr. Anne Koch, who has supported me throughout my graduate degree with her guidance, patience and passion for research. My invaluable experiences working in her lab has enriched my growth as a student and researcher. I am also grateful to my committee members Dr. Brent Derry, Dr. Daniel Durocher and Dr. Hitoshi Okada for their helpful comments and suggestions. I would like to thank Lily Meng for her extraordinary patience and assistance with scientific experiments, and for sharing good laughs and stories with me. I am also indebted to my colleagues, Kellie Jacks and Rich McCulloch, for their support and helpful suggestions. Above all, I would like to thank Amanda Fenton for being there for me from the very beginning to the end. She has not only provided me with a constant oasis of ideas and encouragement regarding my scientific work, but a lifelong friendship as well. I am deeply grateful for all of the special people in my life who have always believed in me and supported me through the good and difficult times. Finally, I thank my parents for their encouragement and support throughout all of my studies.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract.................................................................................................................. ii</td>
</tr>
<tr>
<td>Acknowledgements..................................................................................................... iii</td>
</tr>
<tr>
<td>Table of Contents...................................................................................................... iv</td>
</tr>
<tr>
<td>List of Figures.......................................................................................................... vi</td>
</tr>
<tr>
<td>List of Symbols and Abbreviations.............................................................................. viii</td>
</tr>
</tbody>
</table>

## CHAPTER 1: INTRODUCTION

1.1 Genomic Instability and DNA Double-strand Break Repair...........................................2
1.2 Non-homologous End-Joining..........................................................................................4
1.3 DNA End-Processing...................................................................................................... 9
1.4 Polynucleotide Kinase and Aprataxin in DNA Repair....................................................10
1.5 A unique subclass of Forkhead-Associated Domains....................................................12
1.6 Aprataxin and Polynucleotide kinase Like Factor..........................................................13
1.7 APLF is implicated in Non-homologous End-Joining..................................................15
1.8 The Importance of the Spatio-temporal Distribution of Proteins in DNA Repair....17
1.9 The Classical Nuclear Transport Mechanism..............................................................18
1.10 Non-Classical Nuclear Transport Mechanisms..........................................................21
1.11 Study Aims and Summary of Findings........................................................................23

## CHAPTER 2: MATERIALS & METHODS

2.1 Cloning and plasmid constructions...............................................................................27
2.2 Protein expression and purification...............................................................................29
2.3 Cell culture...................................................................................................................30
2.4 Transfections................................................................................................................31
2.5 Antibodies....................................................................................................................31
2.6 Preparation of cell extracts, immunoprecipitations, pull-downs and immunoblotting .................................................................................................................................31
2.7 Immunofluorescence microscopy..................................................................................32
2.8 Plasmid integration assay............................................................................................33
CHAPTER 3: RESULTS

3.1 The Ku-binding domain on APLF spans amino acids 180 to 200 ........................................36
3.2 APLF nuclear localization is disrupted in the absence of Ku ...........................................38
3.3 APLF is predominantly localized to the nucleus in the absence of DNA-PKcs or XRCC4 ..........................................................44
3.4 APLF is predominantly localized to the nucleus in the absence of XRCC1 ....................46
3.5 APLF nuclear localization is not dependent on the APLF FHA domain ...................................47
3.6 PNK and APTX are predominantly localized to the nucleus in the absence of Ku ...........49
3.7 IR-induced DNA damage does not stimulate the nuclear localization of APLF ........50
3.8 APLF nuclear localization is dependent on an interaction with Ku ........................................53
3.9 APLF depends on an interaction with Ku to be localized to the nucleus and interact with the XRCC4-DNA ligase IV complex .............................................................54
3.10 Tryptophan 189 of APLF strongly contributes to the nuclear localization of APLF ....57
3.11 Conserved Tryptophan 189 of APLF is required for the interaction of APLF with Ku ..........................................................59
3.12 Amino acid residues 569 to 598 within the Ku80 C-terminus region strongly contribute to the nuclear localization of APLF ........................................................................62
3.13 APLF interacts with Ku in the absence of the Ku80 C-terminus region .................64
3.14 The N-terminal region of APLF is required for APLF-dependent NHEJ .......................66

CHAPTER 4: DISCUSSION & FUTURE DIRECTIONS

4.1 The Ku-binding domain on APLF is mapped to amino acids 180-200 ................................72
4.2 The APLF-Ku interaction is important for the nuclear localization of APLF .................73
4.3 Other core NHEJ proteins and APLF-interacting proteins do not strongly contribute to the nuclear localization of APLF ........................................................................79
4.4 The FHA and PBZ domains are not required for the nuclear localization of APLF ...80
4.5 Mapping the APLF-binding site on Ku ..................................................................................82
4.6 The Function of APLF in NHEJ .......................................................................................84
4.7 Other functional roles and consequences of the APLF-Ku interaction ....................86
4.8 Summary .................................................................................................................................89

REFERENCES ....................................................................................................................................92
# LIST OF FIGURES

## CHAPTER 1: INTRODUCTION

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>General organization of the DNA damage response pathway</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>General model for mammalian NHEJ</td>
<td>8</td>
</tr>
<tr>
<td>1.3</td>
<td>Structural organization of the Ku70 and Ku80 proteins</td>
<td>9</td>
</tr>
<tr>
<td>1.4</td>
<td>Structural organization of the PNK, APTX and APLF proteins</td>
<td>13</td>
</tr>
<tr>
<td>1.5</td>
<td>Structural organization and interacting partners of APLF</td>
<td>15</td>
</tr>
<tr>
<td>1.6</td>
<td>Schematic model of the classical nuclear transport mechanism</td>
<td>20</td>
</tr>
</tbody>
</table>

## CHAPTER 3: RESULTS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The peptide region spanning amino acids 180 to 200 in APLF is evolutionary conserved within vertebrates</td>
<td>37</td>
</tr>
<tr>
<td>3.2</td>
<td>APLF&lt;sup&gt;Δ180-200&lt;/sup&gt; is pancellular in human cells</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>APLF is pancellular in Ku80 mutant XRS-5 CHO cells</td>
<td>40</td>
</tr>
<tr>
<td>3.4</td>
<td>Ectopically expressed APLF associates with endogenous Ku in CHO-K1 cells</td>
<td>41</td>
</tr>
<tr>
<td>3.5</td>
<td>The nuclear localization of endogenous Ku is not altered in the absence of APLF</td>
<td>42</td>
</tr>
<tr>
<td>3.6</td>
<td>APLF exhibits “predominantly nuclear”, “pancellular” and “predominantly cytoplasmic” subcellular localization phenotypes in Ku80 mutant XRS-5 cells</td>
<td>42</td>
</tr>
<tr>
<td>3.7</td>
<td>APLF is pancellular or predominantly cytoplasmic in the absence of Ku</td>
<td>44</td>
</tr>
<tr>
<td>3.8</td>
<td>APLF is pancellular or predominantly cytoplasmic in the absence of Ku, but is predominantly nuclear in the absence of other NHEJ proteins</td>
<td>45</td>
</tr>
<tr>
<td>3.9</td>
<td>APLF is predominantly nuclear in the absence of XRCC1</td>
<td>47</td>
</tr>
<tr>
<td>3.10</td>
<td>The APLF FHA domain is not required for the nuclear localization of APLF</td>
<td>48</td>
</tr>
<tr>
<td>3.11</td>
<td>PNK and APTX are predominantly nuclear in the absence of Ku</td>
<td>50</td>
</tr>
<tr>
<td>3.12</td>
<td>IR-induced DNA damage does not stimulate the nuclear localization of APLF in the presence or absence of Ku</td>
<td>51</td>
</tr>
<tr>
<td>3.13</td>
<td>IR-induced DNA damage does not stimulate the nuclear localization of APLF in the presence or absence of the Ku-binding domain</td>
<td>52</td>
</tr>
<tr>
<td>3.14</td>
<td>APLF requires the Ku-binding domain (amino acid residues 181-200) to be localized to the nucleus</td>
<td>54</td>
</tr>
<tr>
<td>3.15</td>
<td>APLF requires the Ku-binding domain (amino acid residues 181-200) to localize to the nucleus and thus interact with the nuclear XRCC4-DNA ligase IV complex</td>
<td>55</td>
</tr>
</tbody>
</table>
3.16 The absence of the Ku-binding domain on APLF does not affect the ability of APLF to interact with the XRCC4-DNA ligase IV complex ................................................................. 57

3.17 Conserved Tryptophan 189 of APLF strongly contributes to the nuclear localization of APLF .............................................................................................................. 59

3.18 Conserved residues within the APLF Ku-binding domain are required for the APLF-Ku interaction ................................................................................................. 60

3.19 Conserved Tryptophan 189 of APLF is required for the interaction of APLF with Ku ........................................................................................................................ 62

3.20 Amino acid residues 569 to 598 within the Ku80 C-terminus region strongly contribute to the nuclear localization of APLF ........................................................................ 64

3.21 APLF associates with Ku80 in the absence of the Ku80 C-terminus region ............................... 65

3.22 Plasmid integration assay ........................................................................................................ 69

3.23 The N-terminal region of APLF, containing the XRCC4-Ligase IV and Ku binding domains (spanning amino acids 1-200), is required for APLF-dependent NHEJ ........ 70

CHAPTER 4: DISCUSSION & FUTURE DIRECTIONS

4.1 Proposed model for APLF nuclear localization and recruitment to sites of DSBs ............................... 91
LIST OF SYMBOLS AND ABBREVIATIONS

γ-H2AX  Gamma-H2AX (histone variant H2AX phosphorylated at Ser139)
°C  Degree Celsius (unit of temperature)
ng, µg, mg, g  Nano-, micro-, milli, -gram (g) (unit of mass)
nl, µL, mL, L  Nano-, micro-, milli, -litre (L) (unit of volume)
nm, µm, mm, m  Nano-, micro-, milli, -metre (m) (unit of length)
nM, µM, mM, M  Nano-, micro-, milli, -molar (M) (unit of concentration)
1x PI  Complete protease inhibitor cocktail
3’  Downstream end of a DNA chain
5’  Upstream end of a DNA chain
53BP1  p53 binding protein 1
ADP  Adenosine diphosphate
AOA1  Ataxia oculomotor apraxia type 1
APLF  Aprataxin and polynucleotide kinase-like factor
APTX  Aprataxin
ATM  Ataxia telangiectasia mutated kinase
ATP  Adenosine triphosphate
ATR  Ataxia telangiectasia and Rad-3 related
BARD1  BRCA1-associated RING domain protein 1
BAX  Bcl-2-associated X
B-NHEJ  Backup Nonhomologous end-joining
BRCA1  Breast cancer type 1 susceptibility protein
C2orf13  Chromosome 2 within open reading frame 13
CHO  Chinese hamster ovary
CK2  Casein kinase 2
CPT  Camptothecin
CSA  Cockayne syndrome group A
CSB  Cockayne syndrome group B
C-terminal  Downstream end of a polypeptide chain (carboxy termini)
DAPI  4′,6-diamidino-2-phenylindole fluorescent DNA stain
DDB  DNA binding protein
DDR  DNA damage response
DMEM  Dulbecco’s modified Eagle’s medium
DMSO  Dimethyl sulfoxide
DNA  Deoxyribonucleic acid
DNA-PK  DNA dependent protein kinase holoenzyme
DNA-PKcs  DNA dependent protein kinase catalytic subunit
DSB  DNA double strand break
EDTA  Ethylene diamine tetracetic acid
eCFP  Enhanced cyan fluorescent protein
eGFP  Enhanced green fluorescent protein
EGTA  Ethylene glycol tetracetic acid
EMSA  Electrophoretic mobility shift assays
ERCC1  Excision repair cross-complementing 1
EtBr  Ethidium bromide
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA</td>
<td>Fanconi Anemia</td>
</tr>
<tr>
<td>FAA</td>
<td>Fanconi anemia A protein</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FHA</td>
<td>Forkhead associated domain</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanine triphosphate</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase protein tag</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray (unit of radiation)</td>
</tr>
<tr>
<td>H2O2</td>
<td>Hydrogen peroxide</td>
</tr>
<tr>
<td>hEXO1</td>
<td>Human exonuclease 1</td>
</tr>
<tr>
<td>HIT</td>
<td>Histidine triad domain</td>
</tr>
<tr>
<td>hMLH1</td>
<td>Human mutL homolog 1</td>
</tr>
<tr>
<td>hMSH2</td>
<td>Human mutS homolog 2</td>
</tr>
<tr>
<td>HR</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>IR</td>
<td>Ionizing radiation</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>Ku</td>
<td>Ku70/Ku80 heterodimer</td>
</tr>
<tr>
<td>LMB</td>
<td>Leptomycin B</td>
</tr>
<tr>
<td>MMR</td>
<td>Mismatch repair</td>
</tr>
<tr>
<td>MMS</td>
<td>Methyl methanesulfonate</td>
</tr>
<tr>
<td>MRN</td>
<td>MRE11-RAD50-NBS1</td>
</tr>
<tr>
<td>NAD</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NER</td>
<td>Nucleotide excision repair</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear exit signal</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization signal</td>
</tr>
<tr>
<td>NPC</td>
<td>Nuclear pore complex</td>
</tr>
<tr>
<td>NT</td>
<td>Non-targeting (siRNA)</td>
</tr>
<tr>
<td>N-terminal</td>
<td>Upstream end of a polypeptide chain (amino-termini)</td>
</tr>
<tr>
<td>NHEJ</td>
<td>Nonhomologous end-joining</td>
</tr>
<tr>
<td>PALF</td>
<td>Polynucleotide kinase and Aprataxin-Like Factor</td>
</tr>
<tr>
<td>PAR</td>
<td>Poly (ADP) ribose</td>
</tr>
<tr>
<td>PARP1</td>
<td>Poly (ADP) ribose polymerase 1</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PBZ</td>
<td>Poly (ADP) ribose-binding zinc finger</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction (to amplify DNA)</td>
</tr>
<tr>
<td>PIA</td>
<td>Plasmid integration assay</td>
</tr>
<tr>
<td>PIKK</td>
<td>Phosphoinositide 3-kinase related kinases</td>
</tr>
<tr>
<td>PNK</td>
<td>Polynucleotide Kinase</td>
</tr>
<tr>
<td>PVDVF</td>
<td>Polyvinylidene difluoride</td>
</tr>
<tr>
<td>Ran</td>
<td>RAS-related nuclear protein</td>
</tr>
<tr>
<td>Ran-GTP</td>
<td>RAS-related nuclear protein in its GTP bound state</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNai</td>
<td>RNA interference</td>
</tr>
<tr>
<td>SAP</td>
<td>SAF-A/B, Acinus and PIAS</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA duplexes</td>
</tr>
<tr>
<td>SSB</td>
<td>DNA single-strand break</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stanned DNA</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian virus 40 of large T-antigen</td>
</tr>
<tr>
<td>TdT</td>
<td>Terminal deoxynucleotidyltransferase</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet light</td>
</tr>
<tr>
<td>V(D)J</td>
<td>Variable (diversity) joining recombination</td>
</tr>
<tr>
<td>vWA</td>
<td>von Willebrand A</td>
</tr>
<tr>
<td>WCE</td>
<td>Whole cell extract</td>
</tr>
<tr>
<td>WRN</td>
<td>Werner syndrome protein</td>
</tr>
<tr>
<td>WT</td>
<td>Wild-type</td>
</tr>
<tr>
<td>Xip1</td>
<td>XRCC1-interacting protein 1</td>
</tr>
<tr>
<td>XLF</td>
<td>XRCC4-like factor</td>
</tr>
<tr>
<td>XPD</td>
<td>Xeroderma pigmentosum group D</td>
</tr>
<tr>
<td>XPF</td>
<td>Xeroderma pigmentosum group F</td>
</tr>
<tr>
<td>XRCC</td>
<td>X-ray cross-complementation</td>
</tr>
<tr>
<td>XRCC1</td>
<td>X-ray repair cross-complementing protein 1</td>
</tr>
<tr>
<td>XRCC4</td>
<td>X-ray repair cross-complementing protein 1</td>
</tr>
<tr>
<td>ZF</td>
<td>Zinc finger</td>
</tr>
</tbody>
</table>
CHAPTER 1:

INTRODUCTION
1.1 Genomic Instability and DNA Double-strand Break Repair

DNA repair and maintenance of genome stability are crucial to cellular and organismal function, and defects in these processes have been implicated in cancer. Of the various types of DNA damage that exist within the mammalian cell, the DNA double strand break (DSB) - which results from the disruption of the phosphodiester backbone on both strands of the DNA double helix - is conceivably the most deleterious [1, 2]. Indeed the cell encounters 10-100 endogenous DSBs per nucleus per day, and even one DSB is sufficient to kill a cell if it inactivates an essential gene or triggers apoptosis [3]. DSBs can be induced by extrinsic agents such as ionizing radiation (IR) and certain chemotherapeutic drugs, or intrinsic events such as the production of reactive oxygen species during cellular metabolism, DNA replication from stalled replication forks at sites of single-strand breaks (SSB), and programmed rearrangements during V(D)J recombination which facilitates antibody and T-cell receptor diversity during lymphoid cell development [1, 4, 5]. If cells fail to resolve DSBs correctly, this can give rise to the loss, amplification and translocation of chromosomal material, thus leading to genomic instability and, ultimately, an enhanced rate of carcinogenesis [4, 6].

To ensure survival and maintain genome integrity, cells respond to DSBs by activating a complex DNA damage response pathway that impinges on a variety of cellular events, including cell cycle control, DNA repair and programmed cell death [7-9] (Figure 1.1). This pathway provides a mechanism for amplifying a signal from a sensor, which recognizes the DSB, through a transduction cascade to a series of downstream effector molecules. For the most part, the initial response upon
encountering a DSB is to slow progression through the cell cycle, so as to provide time for repair [7, 9]. However, if faced with excessive or unrepairable damage, it may be more prudent for the cell to undergo apoptosis, thus preventing mutations from being propagated to daughter cells [3, 9].

There are two distinct mechanisms for mammalian DNA DSB repair: homologous recombination (HR) and non-homologous end-joining (NHEJ) [5]. In HR, which is operative only in the S and G2 phases of the cell cycle where a homologous DNA template is available for sequence-dependent repair, the DNA ends of the DSB are first resected by nucleases [4, 5]. The resulting single-stranded DNA ends invade the double-helix of an intact, homologous chromosome, and are extended by DNA polymerase, which faithfully copies the information from the intact template [1]. In contrast, NHEJ does not rely on extensive homologies between the two recombining ends of a DSB, but instead directly re-joins the two ends in a sequence-independent manner [4].
Figure 1.1 - General organization of the DNA damage response pathway. Exogenous or endogenous agents can induce DSBs in the genome. The presence of a DSB is recognized by a sensor, which transmits the signal to a series of downstream effector molecules through a transduction cascade to activate signaling mechanisms for cell death, cell cycle arrest and DNA repair. Mammalian DSBs are repaired by either HR or NHEJ. Failure to repair the DSB may ultimately lead to genomic instability.

1.2 Non-homologous End-Joining

Functioning in all phases of the cell cycle, NHEJ is the predominant repair mechanism for DSBs induced by exogenous factors, such as IR, in mammalian cells [6]. NHEJ is also fundamental in repairing the programmed DSBs generated during V(D)J recombination, which is an essential process for development of a functional immune system [10]. Much of our knowledge of this repair pathway, including the identification of core NHEJ proteins, has been established through the characterization of IR-sensitive mutant Chinese hamster cell lines. In particular, four
X-ray cross-complementation (XRCC) groups of these cell lines were found to share similar phenotypes, including acute radiosensitivity, defective DSB repair and impaired V(D)J recombination [11]. Subsequently, the human genes XRCC4, XRCC5, XRCC6 and XRCC7, which encode the proteins XRCC4, Ku80, Ku70 and DNA-PKcs, respectively, were identified to complement the defects in their corresponding mutant cell lines, thus highlighting the importance of these factors in DSB repair [11, 12]. Since then, it has been established that NHEJ is critical for IR resistance, DSB repair and V(D)J recombination, and that the key components of the NHEJ apparatus include the DNA-dependent protein kinase (DNA-PK) - which is composed of the DNA-end binding Ku70/Ku80 (Ku) complex and the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) - and the XRCC4-DNA ligase IV (X4L4) complex which catalyzes DNA end-joining [11, 13]. Although all the steps of NHEJ are still not fully understood, the following model of this repair pathway suggests a sequential order of events (Figure 1.2).

The initial step in NHEJ is the binding of Ku to the broken DNA ends. Ku, an extremely stable heterodimeric complex of Ku70 (70 kDa) and Ku80 (80 kDa), is an abundant nuclear protein [14, 15]. The Ku70 and Ku80 subunits share sequence similarity over the length of their proteins, thus suggesting that they diverged from a common ancestor [14-16]. Each Ku subunit is comprised of three regions: (1) an amino-terminal von Willebrand A domain which facilitates dimerization of the two Ku subunits; (2) a central core domain which binds to DNA; and (3) a divergent carboxy-terminal region [15, 17] (Figure 1.3). In Ku70, this latter region contains a SAP (named after three proteins containing this structural motif: SAF-A/B, Acinus and PIAS)
domain that has DNA-binding capability, whereas the longer carboxy-terminal region of Ku80 has been shown to bind to DNA-PKcs [15, 18].

Although Ku is implicated in numerous cellular processes, including transcription, apoptosis and telomere maintenance, it is primarily known for its role in DNA repair [19-22]. Ku binds to double-stranded DNA ends with high affinity, and the three-dimensional crystal structure of this protein reveals that the heterodimer forms an asymmetrical ring-like structure - with Ku70 and Ku80 each forming half of the ring - that completely encircles the DNA helix [19, 23]. The contact between Ku and the DNA backbone occurs in a structure-specific and sequence-independent manner, whereby the positively-charged amino acid residues within the central core domain of Ku interact with the negatively-charged phosphodiester backbone of DNA [23]. The function of Ku-binding to DNA ends is thought to protect the ends from nucleolytic degradation, align and maintain the ends in close proximity, as well as to recruit DNA-PKcs to the DSB which results in activation of its kinase function [13, 19].

DNA-PKcs is a serine/threonine kinase classified within the family of phosphatidyl-inositol-3-kinase related kinases (PIKKs), which also includes ataxia-telangiectasia mutated (ATM) and ATM- and Rad3-related (ATR), both of which function in DNA damage-induced signaling [13, 24]. Although DNA-PKcs can bind to DNA in the absence of Ku, its dissociation constant decreases 100-fold in the presence of Ku-bound DNA [25]. Upon binding of DNA-PKcs, Ku inwardly translocates along the DNA in an ATP-independent manner whilst still maintaining close contact with DNA-PKcs, thus forming a stable synaptic complex in which the two DNA ends are held together [16]. During synapsis, end-processors modify non-complementary or damaged
ends that cannot be ligated directly [26]. The activated DNA-PK complex also acquires the ability to phosphorylate itself and other DSB repair factors, including the XRCC4-DNA ligase IV complex, which is important for their activation and recruitment to the DSB [4, 27].

The final step in NHEJ is ligation of the DNA ends by the XRCC4-DNA ligase IV complex, which exists as a tightly-bound complex in the stoichiometric ratio of 2:1 [28]. XRCC4, a nuclear phosphoprotein that is devoid of intrinsic catalytic activity, is absolutely required for the stabilization and activation of DNA ligase IV [29, 30]. Another recently identified core NHEJ protein, XRCC4-like factor (XLF, also known as Cernunnos), has also been shown to interact with and stimulate the XRCC4-DNA ligase IV complex [31, 32]. Moreover, interaction of the X4L4 complex with Ku has been shown to be necessary for efficient ligation of DNA ends [33].
Figure 1.2 - General model for mammalian NHEJ. (1) DSBs can be induced in the genome by either a DNA damaging agent, such as IR, or as a consequence of V(D)J recombination. (2) The Ku70/Ku80 heterodimer (Ku) recognizes the DSB and binds the free DNA ends. (3) Ku translocates inwards from the DNA end in an ATP-independent manner. DNA-PKcs is recruited to DNA-bound Ku to form the active DNA-PK holoenzyme. A protein-protein interaction between the two DNA-PKcs molecules on the opposing DNA ends bridges the two broken DNA ends together during synapsis. (4) If the DSB contains unligatable ends, the ends are processed by specific end-processors to generate 5’ phosphate groups and 3’ hydroxyl groups. (5) The XRCC4-Ligase IV complex is recruited to the DNA-PK complex, and catalyzes end-joining. There is dissociation of the NHEJ repair factors once the DSB is repaired.
Figure 1.3 - Structural organization of the Ku70 and Ku80 proteins. Both proteins contain an amino-terminal von Willebrand domain (vWA), a central core domain, and a divergent carboxy-terminal domain (CT). The nuclear localization signal (NLS) for each protein resides within the CT. Ku70 contains a DNA-binding SAP (SAF-A/B, Acinus and PIAS) domain at the extreme carboxy-terminus, while Ku80 has a longer CT domain that contains a region required for interaction with DNA-PKcs (PK).

1.3 DNA End-Processing

In addition to the NHEJ core components, several accessory proteins have been implicated in NHEJ, particularly in the end-processing step, which occurs prior to ligation. Many DSB-inducing agents often generate DNA ends that cannot undergo direct ligation as they contain non-ligatable end groups, such as 3’-phosphates, 5’-hydroxyls, incompatible overhangs and damaged DNA backbone sugars or bases [3, 26]. Therefore, these complex DNA termini must be modified by a host of end-processing enzymes in order to restore the ligatable 5’ phosphate and 3’-hydroxyl moieties. One end-processor implicated in NHEJ is Artemis, a 5’-3’ exonuclease that complexes with and is phosphorylated by DNA-PKcs [34]. These two events stimulate the endonucleolytic activity of Artemis, which enables this nuclease to process 5’ and 3’ overhangs in NHEJ, and cleave hairpin intermediates during VDJ recombination. Another group of end-processors thought to participate in NHEJ are DNA polymerases of the Pol X family, including pol μ, pol λ, and terminal deoxynucleotidyltransferase (TdT), which synthesize nucleotides to fill in gaps and overhangs at DSBs [35].
syndrome protein (WRN) is an enzyme with both helicase and exonuclease activity, and can physically interact with Ku and the DNA-PK complex, which strongly stimulates WRN exonuclease activity on DNA ends [36]. Although Ku is presumed to only bind to DNA ends and recruit other repair proteins to sites of DSBs, intriguingly, it has recently been reported that Ku has enzymatic activity and is directly involved in end-processing, where it functions as a 5’-deoxyribose-5-phosphate lysase to efficiently repair DSBs with associated abasic sites by NHEJ [37]. Polynucleotide kinase (PNK) possesses both 5’ DNA kinase and 3’ DNA phosphatase activities, and functions to phosphorylate 5’-hydroxyl groups and remove 3’-phosphate groups at IR-induced DNA breaks in both single-strand break (SSB) and DSB repair pathways [38-40]. Aprataxin (APTX) catalyzes the release of adenylate groups from 5’ DNA ends in both SSB and DSB repair pathways [41-44].

1.4 Polynucleotide Kinase and Aprataxin in DNA Repair

Both PNK and APTX play a critical role in processing damaged DNA termini. PNK is a monomeric and nuclear protein with three conserved domains: (1) an amino-terminal forkhead-associated (FHA) domain which is a specific phosphothreonine-binding module; (2) a 5’ DNA kinase domain that catalyzes phosphorylation of 5’ DNA ends; and (3) a 3’ DNA phosphatase domain that mediates dephosphorylation of 3’ DNA ends [45, 46] (Figure 1.4). PNK is an integral component of both SSB and DSB repair, and human cells depleted of this end-processor exhibit a pronounced sensitivity to exogenous DNA-damaging agents, such as IR, slower repair kinetics of DNA strand breaks, and an increased frequency of spontaneous mutation, hence
emphasizing the importance of PNK for DNA repair [47]. PNK is involved in SSB repair through its interaction with XRCC1 - a protein that associates with DNA ligase III in a manner that is analogous to the interaction of XRCC4 and DNA ligase IV in NHEJ - while PNK is involved in DSB repair through its interaction with the core NHEJ component XRCC4 [38-40]. It has been shown that XRCC1 and XRCC4 are phosphorylated by the serine/threonine casein kinase 2 (CK2), and that the FHA domain of PNK specifically recognizes and interacts with these phosphorylated forms of XRCC1 and XRCC4, thereby directing PNK to the repair sites of SSBs and DSBs, respectively [38-40, 48].

APTX, a nuclear protein defective in the neurodegenerative disorder ataxia oculomotor apraxia type 1 (AOA1), has three functional domains: (1) an amino-terminal FHA domain that shares extensive homology to the PNK FHA domain; (2) a central Histidine Triad (HIT) domain that catalyzes the release of adenylate groups covalently attached to 5′phosphate termini; and (3) a carboxy-terminal C2H2-type zinc finger motif which mediates DNA-binding [42] (Figure 1.4). Cell lines derived from AOA1 patients exhibit hypersensitivity to DNA-damaging agents such as hydrogen peroxide and camptothecin, thus implicating APTX in DNA damage repair [49]. APTX is recruited to sites of SSBs and DSBs in a manner similar to that of PNK. The FHA domain of APTX interacts with CK2-phosphorylated XRCC1 and XRCC4, substantiating a role for APTX in the repair of SSBs and DSBs, respectively [43, 50].
1.5 A unique subclass of Forkhead-Associated Domains

FHA domains are present in a wide-range of proteins from both prokaryotes and eukaryotes, suggesting that these domains are involved in many different cellular processes [51]. Specifically, in eukaryotes, many FHA domain-containing proteins play critical roles in DNA repair, cell cycle checkpoints or transcriptional regulation. Whilst FHA domains - typically ranging from 60-150 amino acids in length - do not share extensive sequence similarity, they share similar secondary and tertiary structures, consisting of a sandwich of two anti-parallel β-sheets [52, 53]. A loop structure between these β-sheets contains conserved residues that coordinate phosphopeptide binding with proteins phosphorylated by serine/threonine kinases [51]. Whereas other characterized FHA domains recognize specific sequences that are C-terminal to the phosphothreonine residue, the PNK and APTX FHA domains exhibit preferences for sequences that are N-terminal to the phosphothreonine [40, 45, 46, 51]. In particular, positively-charged lysine and arginine residues found within the binding pocket of the FHA domains of PNK and APTX both display specificity for acidic residues surrounding the phosphothreonine on XRCC1 and XRCC4 [40, 45, 48]. As such, PNK and APTX are divergent members of the FHA domain family [40].

Recently, a protein encoded by the open reading frame 13 of chromosome 2 (C2orf13) was identified as another member of this FHA subgroup [40]. This protein was termed Aprataxin and Polynucleotide kinase-Like Factor (APLF), but also goes by the name of Polynucleotide kinase and Aprataxin-Like Factor (PALF) or XRCC1-interacting protein 1 (Xip1) [40, 54, 55]. Sequence analysis of the FHA domains of PNK, APTX and APLF demonstrate that they are all closely related [56]. On account of
the functional similarities between the FHA domains of PNK and APTX, whereby the FHA domain-mediated interactions with CK2-phosphorylated XRCC1 and XRCC4 permit PNK and APTX to participate in DNA repair, it was hypothesized that APLF, which possesses a FHA domain similar to those of PNK and APTX, also plays a role in DNA repair.

Figure 1.4 - Structural organization of the PNK, APTX and APLF proteins. All three proteins contain an amino-terminal FHA domain. PNK contains a phosphatase domain which catalyzes phosphorylation of 5’ DNA ends, and a kinase domain that dephosphorylates 3’ DNA ends. APTX contains a central Histidine Triad (HIT) domain that catalyzes the release of adenylate groups covalently attached to 5’phosphate termini, and a carboxy-terminal zinc finger motif (ZF) that mediates binding to DNA. Only PNK and APTX possess an identifiable nuclear localization signal (NLS). APLF contains two tandem poly-ADP-ribose-binding zinc finger (PBZ) domains at the extreme carboxy-terminus, which mediate interactions with PARP1 and PAR.

1.6 Aprataxin and Polynucleotide kinase Like Factor

APLF is an evolutionary conserved protein with homologues identified in mammals, rodents, chicken and sea urchin [56]. Like PNK and APTX, the amino-terminal FHA domain of APLF has been found to direct interactions with the CK2-phosphorylated threonine residue of both XRCC1 and XRCC4, thus implicating APLF in SSB and DSB repair [54, 56, 57]. However, unlike PNK and APTX, APLF has neither an
identifiable catalytic domain nor a classical nuclear localization signal (NLS), and is also devoid of intrinsic DNA binding ability (Figure 1.4) [56].

In addition to possessing the amino-terminal FHA domain, APLF also contains two tandem poly-ADP-ribose (PAR)-binding zinc finger (PBZ) domains within the carboxy-terminal region [54, 56, 57] (Figure 1.5). These domains are highly conserved across different species, and are present in proteins as either single or two tandem motifs [56-58]. Several of these proteins have putative roles in DNA damage repair and signaling. Interestingly, the APLF tandem PBZ domains have been found to be essential for the rapid recruitment of APLF to sites of DNA damage through direct interactions with poly(ADP-ribose) polymerase 1 (PARP1) and PAR [57, 59, 60]. PARP1, the most abundant enzyme of the PARP superfamily, plays a key role in the DNA damage response [61]. Indeed, this nuclear enzyme has a well-established role in the repair of SSBs by base-excision repair, and has also been recently implicated in DSB repair [62, 63]. PARP1 rapidly binds to DNA strand breaks, and signals the presence of the lesions by synthesizing large amounts of negatively-charged PAR from nicotinamide adenine dinucleotide (NAD') on itself and proximal acceptor proteins [63]. Hyper-poly(ADP-ribosyl)ated PARP1 results in its reduced affinity for DNA, and thus its release from the DNA lesions [61, 63]. This, in turn, provides the necessary area for the assembly of repair proteins at the sites of DNA breaks in a PAR-binding dependent manner. Thus, the ability of APLF to interact with PARP1 and PAR may prove to be a novel component of signaling in the DNA damage response pathway [57, 59, 60].
Another interaction that has been identified is the APLF-Ku interaction. In contrast to PNK and APTX, which are unable to directly interact with Ku, APLF interacts with Ku in an FHA- and PBZ-independent manner, and can bind to Ku while Ku is bound to DNA (Figure 1.5) [54, 56].

![Figure 1.5 - Structural organization and interacting partners of APLF. APLF interacts with XRCC4-DNA ligase IV and XRCC1-DNA ligase III in a FHA-dependent manner, however these are likely mutually exclusive interactions. APLF also associates with Ku independently of the FHA and PBZ domains, undergoes ATM-dependent phosphorylation at S116, and interacts with PARP1 via the first PBZ domain.](image)

1.7 APLF is implicated in Non-homologous End-Joining

A role of APLF in NHEJ is implicated by a number of findings. In addition to its physical interactions with core NHEJ components, namely the XRCC4-DNA ligase IV complex and Ku, APLF has also been shown to undergo phosphorylation following IR at serine residue 116 by ATM, a serine/threonine kinase of the PIKK family (Figure 1.5)
[56]. This ATM-dependent phosphorylation event further implicates APLF in DSB repair signaling since ATM is one of the main transducers of the DSB response pathway [24]. For instance, ATM has been shown to regulate the activity of DNA end-processor Artemis by phosphorylation in a NHEJ subpathway that repairs a small subset (~15%) of IR-induced DSBs with slow kinetics [64-66]. These slow repair kinetics are thought to represent the repair of DSBs that are located within heterochromatin, and since ATM signaling modifies chromatin structure, ATM may be specifically required for the repair of such DSBs [64, 66]. Thus, like Artemis, ATM may regulate APLF activity via phosphorylation, but it remains to be determined how this post-translational modification attributes to the role of APLF in DSB repair and the DNA damage response. Another implication for a role of APLF in DSB repair is that depletion of APLF from human cells subjected to IR is associated with impaired DSB repair kinetics [56, 57]. Furthermore, down-regulation of APLF in human cells has been shown to result in defective random genomic integration of foreign DNA - a process known to be dependent on the NHEJ pathway - thus suggesting that APLF functions in NHEJ to facilitate end-joining [56].

Although these findings implicate APLF in NHEJ, the mechanistic function of APLF in this repair pathway remains to be elucidated. Moreover, APLF is unlikely to be restricted only to the NHEJ pathway, but instead may also function in DNA repair pathways other than NHEJ, as suggested by the ability of APLF to interact with non-NHEJ proteins, including SSB-protein XRCC1 and DNA-strand break sensor PARP1, which is implicated in both SSB and DSB repair. Interestingly, both PARP-1 and XRCC1 have been suggested to play a role in alternative or back-up NHEJ (B-NHEJ) pathways
when classical NHEJ is compromised, such as in the absence of core NHEJ components [67-70]. B-NHEJ is thought to use components of other DNA repair pathways, such as XRCC1 and PARP1, to repair DSBs, albeit in a relatively slow and inefficient manner [67, 68]. By virtue of its ability to interact with both of these proteins, APLF may also participate in this form of repair. However, while we acknowledge the existence of non-NHEJ roles of APLF in DNA repair, the focus of this thesis pertains to the characterization of APLF in the context of DSB repair and the NHEJ pathway.

1.8 The Importance of the Spatio-temporal Distribution of Proteins in DNA Repair

The spatial and temporal distribution of repair complexes at sites of DNA damage is of particular importance for DNA repair and maintaining the integrity of the genome. To function at these sites, repair proteins must migrate into the cell nucleus after they are translated in the cytoplasm since their target DNA substrates are located in the nucleus [71]. Consequently, the spatio-temporal ability of repair proteins to localize to the nucleus may influence how rapidly they are recruited to DNA lesions and how readily they form complexes with other repair proteins [72].

Whilst APLF is implicated in nuclear NHEJ, the mechanism by which APLF is localized to the nucleus is unclear. Unlike PNK and APTX, APLF has no identifiable NLS, yet APLF has been shown to localize predominantly to the nucleus [56]. Furthermore, the nuclear distribution of APLF with its NHEJ-interacting partners, Ku and XRCC4-DNA ligase IV, and whether or not these interacting complexes are mutually-exclusive or co-dependent remains to be determined. As such, studies
focusing on the mechanism of APLF nuclear localization might provide further insight into the functional role of APLF in NHEJ.

In eukaryotic cells, the nucleus is separated from the cytoplasm by the nuclear envelope, a double membrane system [71]. As a result of this spatial segregation, efficient transport mechanisms must exist so that repair proteins are able to localize to sites of DNA damage. The nucleocytoplasmic translocation of proteins occurs through nuclear pore complexes (NPCs), which are perforated in the nuclear envelope and are the sole sites of exchange between the nucleus and cytoplasm [71, 73]. NPCs, composed of greater than a hundred different proteins, termed nucleoporins, form a 10 nm central channel which can accommodate the passive or active transport of proteins [71, 73]. Passive transport permits proteins smaller than approximately 60 kDa to diffuse through NPCs in an energy-independent manner [71]. APLF, for instance, is a small protein of 57 kDa and, for this reason, may be able to passively diffuse into the nucleus. In contrast, larger proteins cannot diffuse freely through NPCs; these proteins are actively and selectively transported across the nuclear envelope [71, 72].

1.9 The Classical Nuclear Transport Mechanism

The active import and export of proteins occurs by the classical nuclear transport mechanism, which is tightly regulated by a large, evolutionary conserved superfamily of transport factors - the karyopherins (Figure 1.6) [73]. Two main classes within this family are importins and exportins, of which the importin α/β heterodimer and exportin-1, respectively, are best characterized [72, 74]. Proteins to be
transported into or out of the nucleus are bound by importins and exportins, respectively, which recognize specific sequences called nuclear targeting signals [73]. Proteins destined to enter the nucleus contain an NLS, whereas proteins that exit the nucleus contain a nuclear exit signal (NES) [71-73]. Both of these oligopeptide sequences do not follow strict consensus sequences, but there are clear similarities between most. Classical NLSs are either monopartite (containing a single stretch of basic residues) or bipartite (containing two clusters of basic residues separated by a 10-12 amino acid linker) [72, 75]. These stretches of basic amino acids - comprised primarily of lysine and arginine residues - interact and confer electrostatic selectivity with the negatively-charged acidic residues that line the binding pocket of the α subunit of importin α/β [75]. The prototypic NES, on the other hand, which is recognized by exportin-1, is not well-characterized, and at best, has been defined as being leucine-rich [73]. Both nuclear targeting signals remain intact in the proteins, thereby permitting multiple rounds of nucleocytoplasmic transport [71].

During nuclear import, importin α/β binds to the NLS-containing protein in the cytoplasm, and then this complex translocates through the NPC via transient interactions between nucleoporins and the β subunit of Importin α/β [74]. Once in the nucleus, dissociation of the complex is dependent on an interaction with the guanine-nucleotide binding RAS-related nuclear protein (Ran) in its GTP bound state (Ran-GTP) [71, 74]. After dissociation, the importins are recycled back to the cytoplasm to be used for another round of import [71]. Nuclear export occurs in an analogous manner, whereby NES-containing proteins inside the nucleus are recognized by Exportin-1 and
Ran-GTP [73]. Subsequent to translocation into the cytoplasm, hydrolysis of Ran-GTP to Ran-GDP dissociates the complex [71].

Figure 1.6 - Schematic model of the classical nuclear transport mechanism. Nuclear import begins with NLS-containing proteins binding to the importin-α subunit of the importin-α/β heterodimer in the cytoplasm. The resulting complex then docks to the cytoplasmic side of the nuclear pore complex (NPC) via the importin-β subunit, and is translocated to the nuclear side where Ran-GTP binds to importin-β, resulting in the disassembly of the importin-α/β complex and the NLS-containing protein. During nuclear export, NES-containing proteins inside the nucleus are recognized by Exportin-1 and Ran-GTP. Upon translocation through the NPC into the cytoplasm, dissociation of the complex is facilitated by hydrolysis of Ran-GTP to Ran-GDP, which is regulated by Ran-GAP.
1.10 Non-Classical Nuclear Transport Mechanisms

Although nuclear targeting signals - NLS and NES motifs - mediate the majority of protein import and export, it has become increasingly apparent that not all proteins follow this canonical transport mechanism and, as such, nucleocytoplasmic transport may be facilitated by non-classical or hitherto-unknown mechanisms. Some examples of non-conventional transport mechanisms described in the literature include: direct binding of proteins to components within the nucleus, such as nucleoporins; cytoskeletal-assisted transport, whereby proteins exploit the cell's microtubular network and molecular motor proteins, such as dynein, to facilitate movement along microtubules toward the nucleus; and stimulation of calcium-mediated signal transduction pathways to regulate nuclear transport [72, 76, 77].

Post-translational modifications or protein binding conformations, which mask or expose nuclear targeting signals, are another means of regulating nuclear transport [72]. Tumour suppressor proteins p53 and breast cancer type 1 susceptibility protein (BRCA1) have all been shown to undergo such transport regulation [78]. In p53, DNA-damaged induced phosphorylation of one of its NESs inhibits its nuclear export, while ubiquitination has been shown to augment its nuclear export [78, 79]. BRCA1 contains two NLSs and a NES, yet additional mechanisms regulate its nuclear transport. The BRCA1 amino-terminal RING-finger domain, which is the binding site for BRCA1-associated RING domain protein 1 (BARD1), has been shown to function as a non-classical NLS, whereby BARD1 binds BRCA1 via the RING domain, and shuttles BRCA1 along with itself into the nucleus [80, 81]. Moreover, the NES of BRCA1 lies within the RING domain, and association with BARD1 masks the BRCA1 NES, thereby preventing
the nuclear export of BRCA1 [80]. Therefore, the BRCA1-BARD1 interaction may have important implications for permitting BRCA1 to both localize and retain within the nucleus, especially under DNA damage conditions.

Protein co-import, or piggyback mechanisms, is another novel way of regulating the nuclear accumulation of proteins, whereby NLS-deficient or NLS-proficient proteins complex with NLS-proficient proteins in the cytoplasm and then translocate together into the nucleus [72]. Indeed, this mode of nuclear translocation has been implicated for many repair proteins. Intriguingly, heterodimerization of the Ku70 and Ku80 subunits has been shown to facilitate the nuclear entry of Ku [82]. Although, each subunit can translocate to the nucleus via their individual NLS, Ku70 and Ku80 can also translocate to the nucleus in a manner that is independent of its own NLS and that requires interaction with the other subunit. This suggests that independent nuclear targeting of Ku70 and Ku80 may facilitate the functioning of each subunit in cellular processes that are independent of each other, while the nuclear translocation of both subunits together as a heterodimer may fulfill cell functions that are dependent on each other. Furthermore, the presence of these redundant NLS signals may exist to ensure nuclear import and, consequently, efficient DSB repair activity. Similarly, mismatch repair (MMR) protein human exonuclease 1 (hEXO1) has been reported to depend on complex formation with either human mutL homolog 1 (hMLH1) or human mutS homolog 2 (hMSH2) - both of are involved in MMR and possess NLSs - to localize to the nucleus when its own NLS is dysfunctional [83]. The heterotrimeric protein complex consisting of MRE11, RAD50 and NBS1 (MRN) is involved in DSB repair.
and the DNA damage response, and it has been suggested that NLS-proficient NBS1 is required for the nuclear localization of the entire MRN complex [84].

As mentioned above, NLS-deficient proteins may also adopt piggyback mechanisms with NLS-proficient proteins to localize to the nucleus. For example, there are a number of NLS-deficient proteins involved in nucleotide excision repair (NER) that are thought to depend on interactions with other NLS-proficient NER proteins to localize to the nucleus in a DNA damage-dependent manner. These proteins include the 127 kDa subunit of the human damage-specific DNA binding protein (DDB) (p127), excision repair cross-complementing 1 (ERCC1), Xeroderma pigmentosum group D (XPD) and Cockayne syndrome group A (CSA) [85-88]. Interestingly, p48, ERCC1 and CSA are all proteins smaller than 60 kDa, and are thus sufficiently small enough to passively diffuse through the NPC into the nucleus, yet they have been shown to rely on interactions with proteins involved in their respective repair pathway to localize to the nucleus [88]. Such a mode of nuclear localization may be important for recruiting and modulating the presence of repair factors in specific repair pathways. Collectively, there appears to exist a wide-range of non-classical mechanisms for the nucleocytoplasmic transport of proteins, and it would be interesting to determine whether APLF, which has no identifiable NLS or NES, utilizes any of these mechanisms.

1.11 Study Aims and Summary of Findings

In this study, we undertook a detailed characterization of the role of APLF in DSB repair and the NHEJ pathway. APLF is comprised of a highly conserved amino-
terminal FHA and two carboxy-terminal PBZ domains, and is able to interact with the core NHEJ components, XRCC4 and Ku, which is consistent with a role for APLF in NHEJ. Therefore, to tease apart the role of APLF in this repair pathway, we sought to determine the functional significance of its domains - the FHA-, PBZ- and Ku-binding domains. While the FHA domain-mediated interaction between APLF and XRCC4 has been extensively characterized, the APLF-Ku interaction is less defined, so we focused on mapping and determining the functional consequence of the APLF-Ku interaction. In addition, given that downregulation of APLF is associated with defective NHEJ, we developed a cell-based NHEJ assay to delineate which domains, and thus which APLF interactions, permit APLF to promote efficient NHEJ.

The results presented in this thesis demonstrate that the Ku-binding domain on APLF is mapped to amino acid residues 180-200, where conserved amino acid residue W189 strongly contributes to the APLF-Ku interaction. Remarkably, the APLF-Ku interaction is critical for the nuclear localization of APLF. Disruption of the nuclear localization of APLF also disrupts interaction with the XRCC4-DNA ligase IV complex. Moreover, in preliminary mapping studies characterizing the APLF-binding site on the Ku80 subunit of Ku, we show that the C-terminal portion of Ku80 does not mediate the APLF-Ku interaction. Lastly, we demonstrate that the N-terminal region (amino acids 1-200), containing the XRCC4-Ligase IV and Ku binding domains, is required for APLF-dependent NHEJ. Collectively, these results suggest that localization of APLF to the nucleus, where NHEJ takes place, occurs in a Ku-dependent manner, and that once APLF is retained in the nucleus, the N-terminal portion of APLF, which facilitates
interactions with the core NHEJ proteins XRCC4-DNA ligase IV and Ku, is required for efficient NHEJ.
CHAPTER 2:

MATERIALS & METHODS
2.1 Cloning and plasmid constructions

The following constructs were previously generated and used in another study [56]. The human APLF open reading frame was PCR-amplified from the human cDNA IMAGE clone ID 6042653 (Open Biosystems) and TOPO-cloned into the pcDNA3.1/V5-His mammalian expression vector (Invitrogen) to generate pcDNA3.1-V5/His-APLF (APLF-V5). QuikChange site-directed mutagenesis (Stratagene) was then used to create an arginine-to-alanine substitution at residue 27 of APLF, generating pcDNA3.1-V5/His-APLR27A (APLR27A-V5). To generate pcDNA3.1-V5/His-APLFΔ180-200 (V5-APLFΔ180-200), mutagenesis was used to remove amino acid residues 180-200 from pcDNA3.1-V5/His-APLF. The amplified APLF cDNA was further modified using PCR by the introduction of EcoRI restriction sites at the 5’ and 3’ ends, and then cloned in-frame into the EcoRI site of the bacterial expression vector pGEX4T3 (Amersham) to generate pGEX4T3-APLF (GST-APLF).

Quikchange site-directed mutagenesis was also used to create the R182A, R184A, W189A, M190A and L191A point mutations within pcDNA3.1-V5/His-APLF (APLR182A-V5, APLR184A-V5, APLW189A-V5, APLM190A-V5 and APLL191A-V5, respectively). To generate pcDNA3.1-V5/His-NLS-APLF (NLS-APLF-V5), pcDNA3.1-V5/His-NLS-APLFΔ180-200 (NLS-APLFΔ180-200-V5) and pcDNA3.1-V5/His-APLFW189A (NLS-APLFΔW189A-V5), the simian virus 40 (SV40) large T-antigen nuclear localization signal (MAPKKKRKVK) was inserted before the APLF sequence by mutagenesis of the appropriate pcDNA3.1-V5/His-APLF construct. pcDNA3.1-V5/His-APLF was digested with BamHI and Xhol, and ligated in-frame into pcDNA 3.1-Zeo-HA2.B (Invitrogen) to generate pcDNA 3.1-Zeo-HA2.B-APLF. To then generate pcDNA 3.1-Zeo-HA2.B-APLF1.
pcDNA 3.1-Zeo-HA2.B-APLF$^{1-160}$ and pcDNA 3.1-Zeo-HA2.B-APLF$^{1-200}$ (HA-APLF$^{1-160}$, HA-APLF$^{1-180}$ and HA-APLF$^{1-200}$, respectively), a stop codon was inserted by mutagenesis after amino acid residues 160, 180 and 200, respectively, within pcDNA 3.1-Zeo-HA2.B-APLF. The pSUPER.retro.neo+GFP plasmid was used for expressing the APLF RNA interference (RNAi) sequence GAAGAAATCTGCAAAGATA [57]. A pSUPER RNAi-resistant APLF rescue construct harbouring six different nucleotides in the RNAi target sequence GAGGAGATTGTAGGAC (changed nucleotides are underlined) was generated by mutagenesis (Stratagene). Quikchange site-directed mutagenesis (Stratagene) was then used to create the following RNAi-resistant APLF mutants within the pSUPER RNAi-resistant APLF: arginine-to-alanine substitution at residue 27 (RNAi-resistant APLF$^{R27A}$), the tryptophan-to-alanine substitution at amino acid residue 189 along with the SV40 large T-antigen nuclear localization signal (MAPKKKRKV) insertion before the APLF sequence (RNAi-resistant NLS-APLF$^{W189A}$), the stop codon after amino acid residue 200 (RNAi-resistant NLS-APLF$^{1-200}$), and the cysteine-to-glycine substitutions at both of the zinc-coordinating cysteine residues in the first (amino acid residues 379 and 385) and second (amino acid residues 421 and 427) PBZ motifs (RNAi-resistant APLF$^{ZF1+2m}$). The sequences encoding for the human Ku80 open reading frame were PCR-amplified from the human cDNA IMAGE clone ID: 4555162 (Open Biosystems) and TOPO-cloned in-frame into the EcoRI site of p3XFLAG-CMV-14 (Sigma) to generate p3XFLAG-CMV-14-Ku80 (3xFlag-Ku80). The sequence encoding the human Ku80 open reading frame was excised from p3XFLAG-CMV-14-Ku80 using EcoRI, TOPO-cloned into the EcoRI site of pGEX4T3 (Amersham), and pulled in-frame by site-directed mutagenesis to generate pGEX4T3-Ku80 (GST-Ku80). To then generate
pGEX4T3-Ku80\textsuperscript{1-569}, pGEX4T3-Ku80\textsuperscript{1-598} and pGEX4T3-Ku80\textsuperscript{1-720} (GST-Ku80\textsuperscript{1-569}, GST-Ku80\textsuperscript{1-598} and GST-Ku80\textsuperscript{1-720}, respectively), a stop codon was inserted by mutagenesis after amino acid residues 569, 598 and 720, respectively, within pGEX4T3-Ku80. pGEX4T3-Ku80 was digested with XhoI and BamH1, and ligated in-frame into pEGFP-C1 (Clontech) to generate pEGFP-C1-Ku80 (eGFP-Ku80). To then generate pEGFP-C1-Ku80\textsuperscript{1-569}, pEGFP-C1-Ku80\textsuperscript{1-598} and pEGFP-C1-Ku80\textsuperscript{1-720} (eGFP-Ku80\textsuperscript{1-569}, eGFP-Ku80\textsuperscript{1-598} and eGFP-Ku80\textsuperscript{1-720}, respectively), a stop codon was inserted by mutagenesis after amino acid residues 569, 598 and 720, respectively, within pEGFP-C1-Ku80. To generate pBABE-puro-eCFP, pECFP-C1 (Clontech) was digested with ApaLI and AflIII, blunt-ended and ligated in-frame into the EcoRI site of pBABE-puro (Clontech).

The following constructs were previously generated and used in another study [40]. The human APTX cDNA was PCR-amplified from a human lymphocyte cDNA library and TOPO-cloned into the pcDNA3.1-V5/His vector to generate pcDNA3.1-V5/His-APTX (APTX-V5). The human PNK cDNA was PCR-amplified from a human lymphocyte cDNA library and TOPO-cloned into the pcDNA3.1-V5/His vector to generate pcDNA3.1-V5/His-PNK (PNK-V5). All of the plasmid constructs used in this study were verified by sequence analysis.

2.2 Protein expression and purification

GST-APLF recombinant protein was produced in \textit{E.coli} BL21(DE3)/pLysS (Novagen). Transformed bacteria were grown to an OD\textsubscript{600} of 0.6, and expression was induced by addition of 1 mM isopropyl-\textbeta-D-1-thiogalactopyranoside (IPTG, Sigma) for 3 hrs at 37\textdegree C. The bacteria were then pelleted by centrifugation, washed with
phosphate-buffered saline (PBS), and re-centrifuged. For purification, the cell pellet was resuspended in extraction buffer (50 mM Tris-Hcl pH 7.5, 150 mM NaCl, 1 mM Dithiothreitol (DTT), 1% Triton X-100, 1X Complete protease inhibitor cocktail (Roche), and disrupted by sonication. They lysates were clarified by centrifugation at 16,000 x g at 4°C for 20 minutes. The supernatant was collected and incubated with glutathione sepharose 4B beads (Amersham) for 2 hours at 4°C with gentle mixing. The beads were then washed and the protein eluted with extraction buffer containing 20 mM glutathione. The glutathione was then removed and the purified protein was exchanged into a suitable buffer through three sequential rounds of dialysis using Slide-A-Lyzer dialysis cassettes (Pierce). Unless otherwise specified, all chemicals were purchased from Sigma Aldrich.

2.3 Cell culture

Human embryonic kidney (HEK) 293T and U2OS cell lines were cultured in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and antibiotics. The following Chinese hamster ovary (CHO) cell lines were cultured in Alpha Modified Eagle Medium supplemented with 10% fetal bovine serum (FBS) and antibiotics: parental CHO-K1, Ku80 mutant XRS-5, DNA-PKcs mutant XRC-1, XRCC4 mutant XR-1, and XRCC1 mutant EMC-11. To stably knock down APLF, U2OS cells were transfected with 2 μg of either empty pSUPER vector or pSUPER vector encoding the APLF RNAi sequence GAAGAAATCTGCAAAGATA [57], and then selected with 800 μg/ml of G418 (Gibco). Clonal U2OS cell lines were isolated and then maintained in DMEM supplemented with 10% FBS and 200 μg/ml of G418. All cell lines were grown at 37°C with a humidified atmosphere containing 5% CO₂.
2.4 Transfections

Transient transfections of plasmid DNA were performed with the Effectene transfection kit (Qiagen) according to the manufacturer’s instructions.

2.5 Antibodies

Commercial antibodies used in this study were from Serotec (anti-XRCC4 rabbit polyclonal and anti-DNA ligase IV rabbit polyclonal), Cedarlane (anti-Ku80 rabbit polyclonal), Cell Signaling (anti-Ku70 rabbit polyclonal), Invitrogen (anti-V5 mouse monoclonal), Upstate (anti-HA mouse monoclonal), Santa Cruz Biotechnology (anti-GFP rabbit polyclonal), Sigma (anti-FLAG M2 mouse monoclonal) and Abcam (anti-Tubulin mouse monoclonal). Secondary antibodies for immunoblotting were from Jackson ImmunoResearch (goat anti-mouse and goat anti-rabbit), and secondary antibodies for immunofluorescence microscopy were from Invitrogen (goat anti-mouse Alexa 488).

2.6 Preparation of cell extracts, immunoprecipitations, pull-downs and immunoblotting

For whole cell extract (WCE) preparation, cells were rinsed once in cold PBS, and lysed on ice in 50mM Tris-HCL pH 7.5, 100mM NaCl, 1% Triton X-100, 0.5mM DTT, 0.5mM EDTA, 1x complete mini protease inhibitor cocktail (Roche), 1mM sodium orthovanadate, 40mM β-Glycerophosphate, and 50mM sodium fluoride. Lysates were clarified by centrifugation at 13,200 rpm for thirty minutes at 4°C. When indicated, lysates were pre-treated with 50 μg/ml of ethidium bromide (EtBr, Invitrogen) for thirty minutes on ice.
For immunoprecipitations (IPs), clarified WCEs were incubated with 1 μg of the relevant antibody on ice for sixty minutes with occasional gentle agitation, followed by the addition of 30 μl of protein A-coupled sepharose beads (ThermoFisher) and incubation for three hours at 4°C with nutation. Immunoprecipitates were washed thrice with lysis buffer, resuspended in SDS-PAGE sample buffer.

For pull-down assays, clarified WCEs were incubated with 1 μg of the indicated GST-fusion protein immobilized on glutathione sepharose 4B beads. The complexes were washed three times with lysis buffer, and resuspended in 30 μl sample buffer.

For immunoblotting, samples were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and proteins were transferred to polyvinylidene difluoride (PVDF) membrane (Millipore) using a Trans-Blot SD semi-dry electrophoretic transfer cell (Bio-rad). Membranes were immunoblotted with the primary antibody and relevant secondary antibody as indicated. Detection was performed using SuperSignal West Pico enhanced chemiluminescence (Pierce).

2.7 Immunofluorescence microscopy

Cells were plated into 35mm dishes containing glass coverslips (Fisher) and incubated for twenty-four hours at 37°C in a tissue culture incubator. Cells in these plates were mock- or irradiated with 10 Gy of IR and incubated at 37°C for the indicated time points prior to fixation and immunostaining. Coverslips were then immunostained, with all procedures performed at room temperature, all solutions prepared in PBS, and with PBS washes performed between each step. Cells were fixed with 4% paraformaldehyde for 10min, permeabilized by incubation with 0.5% Triton X-100 (Sigma) and 3% BSA (Sigma) for 10 min, and then blocked with 1.5% normal goat
serum (Sigma) for 1h. Cells were immunostained with the indicated antibody for 2h, and then incubated with either goat anti-mouse Alexa 488 or goat anti-rabbit Alexa 488 secondary antibody (Invitrogen) for 1h. Coverslips were then mounted on slides with Vectashield containing DAPI nucleic stain (Vector Laboratories). Slides were examined using a Zeiss AxioObserver microscope with a Plan-Neofluar 40X objective. Image captured was achieved using a Roper Scientific CoolSnap HQ CCD camera. For subcellular localization analyses, 100 cells for each condition were analyzed in three independent experiments. Images presented in this paper were obtained using Image J and were adjusted for brightness and contrast only.

2.8 Plasmid integration assay

Random plasmid integration assay was performed essentially as described [31, 56] with some modifications. U2OS cells stably depleted in APLF were transfected with empty vector, pSUPER RNAi-resistant APLF, RNAi-resistant APLF<sup>R27A</sup>, RNAi-resistant NLS-APLF<sup>W189A</sup>, RNAi-resistant NLS-APLF<sup>1-200</sup>, or RNAi-resistant APLF<sup>ZF1+2m</sup> and incubated for 48h at 37°C. Throughout experimentation, the expression of these APLF proteins in the APLF-depleted U2OS cells were verified by immunoblotting. Subsequently, cells were transfected with linearized pBABE-puro-eCFP plasmid DNA containing incompatible 5’ overhangs and a puromycin resistance cassette. Twenty-four hours later, cells were replated at low density in selective media containing 2 μg/ml puromycin, and incubated for 10 days at 37°C. Colonies were then stained with Coomassie Blue dye and counted. A fraction of the transfected cells was monitored for nuclear eCFP expression to normalize the data for transfection efficiency. The relative plasmid integration of the APLF-depleted cell lines reconstituted with RNAi-
resistant APLF\textsuperscript{WT} was set at 100% integration. The error bars represent the standard error of the mean. Three independent experiments were performed in triplicate. Transient transfections were performed with the Effectene transfection kit (Qiagen) according to the manufacturer’s instructions.
CHAPTER 3:

RESULTS
3.1 The Ku-binding domain on APLF spans amino acids 180 to 200

We have previously demonstrated that endogenous APLF interacts with Ku under basal conditions in vivo and with DNA-bound Ku in vitro [56]. Moreover, the interaction of APLF with Ku is both FHA- and zinc finger-independent, and the Ku-binding region localizes to amino acid residues 100-263 on APLF - a region located between the FHA domain and tandem zinc finger motifs [56].

Mutational analyses using truncated APLF mutants in an immunoprecipitation assay further demonstrated that amino acid residues 1-200 are sufficient for binding to Ku (Chloe Macrae, unpublished data). Since the FHA domain spans amino acid residues 1-100, and we have previously shown that the APLF-Ku interaction is FHA-independent, the Ku-binding domain was thus further narrowed down to amino acid residues 100-200. Within this region on APLF, five different twenty amino acid regions were looped out (APLF\textsuperscript{Δ101-120}, APLF\textsuperscript{Δ121-140}, APLF\textsuperscript{Δ141-160}, APLF\textsuperscript{Δ161-180}, APLF\textsuperscript{Δ181-200}) to determine whether any of these APLF mutants disrupted the APLF-Ku interaction. We found that removal of amino acid 181-200 diminished binding to Ku, suggesting that this region is required for maximal binding to Ku (Chloe Macrae, unpublished data).

We noted that the amino acid sequence spanning amino acids 180 to 200 is evolutionary conserved within APLF vertebrate homologues, thus highlighting the importance of this region as a putative Ku-binding domain in APLF (Figure 3.1). Since the removal of this region on APLF resulted in a diminished interaction with Ku, and Ku is a core NHEJ component, we hypothesized that this would somehow affect the role of APLF in NHEJ.
Figure 3.1 - The peptide region spanning amino acids 180 to 200 in APLF is evolutionary conserved within vertebrates. Alignment of human APLF and selected homologues by MultAlin [89]. Invariant residues are identified by red text, while residues with greater than 50% consensus are identified by blue text.

To begin characterizing the functional significance of the diminished interaction between APLF and Ku, we first sought to examine its subcellular localization. To assess the subcellular localization of APLFΔ181-200, we transfected human cells with either APLFΔ18-200-V5, or wild-type APLF (APLFWT-V5) as a positive control. After fixation, the cells were immunostained with DAPI nucleic acid stain to visualize the nucleus, and with anti-V5 antibody to visualize the APLF proteins. As judged by indirect immunofluorescence microscopy, APLFΔ181-200-V5 is “pancellular” (present in the nucleus and cytoplasm) compared to the predominantly nuclear localization of APLFWT-V5, suggesting that the diminished interaction between APLF and Ku affects the nuclear retention of APLF (Figure 3.2).
Figure 3.2 - APLF$^{Δ180-200}$ is pancellular in human cells. U2OS cells ectopically expressing APLF$^{WT}$-V5 or APLF$^{Δ181-200}$-V5 were fixed and immunostained with anti-V5 (APLF) antibody and DAPI nuclear stain, and the images were merged.

3.2 APLF nuclear localization is disrupted in the absence of Ku

To further characterize this putative dependence of APLF on Ku for its nuclear localization, as well as to ensure that the pancellular phenotype exhibited by APLF$^{Δ181-200}$ was not resulting from misfolding of the APLF protein due to removal of the twenty amino acid peptide region (amino acids 181 to 200), we next examined the subcellular localization of wild-type APLF in the absence of Ku. We initially tried to deplete Ku in human U2OS via small interfering RNA (siRNA) either by targeting Ku80 alone, or Ku70 and Ku80 in conjunction, however in both cases, we were unable to reduce Ku expression to acceptable levels, most likely because Ku is such an abundant protein (data not shown). Fattah et al. recently generated Ku-depleted human cell lines, however the null cell lines were not viable and the heterozygous inactivation of Ku70 and Ku80 resulted in haploinsufficient phenotypes, including cell proliferation defects and high levels of gross chromosomal rearrangements [90]. Therefore, we
decided to utilize the \textit{Ku80} mutant X-ray-sensitive complementation group 5 (XRS-5) Chinese hamster ovary (CHO) cell line. The XRS-5 cell line is characterized by low levels of wild-type Ku80 transcript, and since Ku70 is unstable in the absence of Ku80, these cells also lack Ku70 protein, thereby rendering the XRS-5 cell line null for Ku [91]. Thus, we examined APLF\textsuperscript{WT}-V5 in XRS-5 cells, or parental CHO-K1 cells as a positive control, and determined that ectopically expressed APLF is pancellular in the absence of Ku, compared to the nuclear localization of APLF in the parent CHO-K1 cell line, as judged indirectly by immunofluorescence microscopy (Figure 3.3). We also engineered a simian virus 40 (SV40) large T-antigen nuclear localization signal (NLS) onto V5-tagged APLF (NLS-APLF\textsuperscript{WT}-V5) and examined the subcellular localization of NLS-APLF\textsuperscript{WT}-V5 in XRS-5 cells to determine whether we could recapitulate the nuclear phenotype of APLF seen in the CHO-K1 cell line. We found that NLS-APLF\textsuperscript{WT} is predominantly nuclear in XRS-5 cells, suggesting that APLF depends on the presence of an NLS to be localized to the nucleus in the absence of Ku.
Figure 3.3 - APLF is pancellular in mutant XRS-5 CHO cells. CHO-K1 (parental) or XRS-5 (Ku80 mutant) CHO cells ectopically expressing APLF-V5 or NLS-APLF-V5 were fixed and immunostained with anti-V5 (APLF) antibody and DAPI nuclear stain, and the images were merged.

Since APLF predominantly localized to the nucleus in CHO-K1 cells, we wanted to ensure that APLF associates with endogenous Ku in these cells. We could not examine the interaction of endogenous APLF with Ku in this cell line since the anti-APLF antibody that we have is specific to human APLF and not to hamster APLF, thus we utilized ectopically expressed V5-tagged APLF. Reciprocal co-immunoprecipitation experiments were performed whereby CHO-K1 cells were transfected with either empty vector or APLF-V5, and the whole cell extracts were immunoprecipitated with either anti-V5 or anti-Ku80 antibodies to recover APLF along with Ku, or vice versa. We detected interactions between APLF and endogenous Ku, indicating that APLF-V5 interacts with Ku in CHO cells (Figure 3.4). In parallel, we also wanted to ensure that
the nuclear localization of endogenous Ku was not altered in the absence of APLF. We previously established a human U2OS cell line stably depleted in APLF (APLF-depleted cells) using RNA interference (RNAi), and a corresponding stable U2OS control cell line expressing wild-type APLF. These cells were fixed, immunostained with anti-Ku80 antibody, and examined by indirect immunofluorescence microscopy. Figure 3.5 demonstrates that Ku is localized to the nucleus in APLF-depleted cells, suggesting that Ku expression is not affected in the absence of APLF.

Figure 3.4 - Ectopically expressed APLF associates with endogenous Ku in CHO-K1 cells. Whole cell extracts (WCE) from parental Chinese hamster ovary CHO-K1 cells ectopically transfected with empty vector or APLF-V5 were immunoprecipitated with either anti-V5 (APLF) or anti-Ku80 antibodies, and the resulting protein complexes were immunoblotted with anti-Ku80 or anti-V5 (APLF) antibodies as indicated.
Figure 3.5 - The nuclear localization of endogenous Ku is not altered in the absence of APLF. The U2OS cell lines either expressing wild-type APLF (Vector control) or stably depleted in APLF by RNAi (APLF-depleted cells) were fixed and immunostained with anti-Ku80 antibody and DAPI nuclear stain, and the images were merged.

In addition to predominantly “pancellular” subcellular localizations of APLF in XRS-5 cells, we also found that APLF exhibited “predominantly cytoplasmic” or “predominantly nuclear” phenotypes, examples of which are shown in Figure 3.6.

Figure 3.6 - APLF exhibits “predominantly nuclear”, “pancellular” and “predominantly cytoplasmic” subcellular localization phenotypes in mutant XRS-5 CHO cells. XRS-5 (Ku80 mutant) CHO cells ectopically expressing APLF-V5 were fixed and immunostained with anti-V5 (APLF) antibody.
Given that the predominantly nuclear localization of APLF is altered in Ku80 mutant XRS-5 cells, we wanted to quantify the subcellular localization of APLF in this cell line. To do so, we examined the subcellular localization of ectopically expressed APLF\textsuperscript{WT}\textsuperscript{-V5} in either XRS-5 cells, or CHO-K1 cells as a positive control. We performed three independent experiments, each time examining 100 cells to determine the subcellular localization status (predominantly nuclear, pancellular or predominantly cytoplasmic) of APLF. We observed that APLF undergoes a great loss in nuclear localization and exhibits large increases in pancellular and predominantly cytoplasmic phenotypes in XRS-5 cells, compared to a predominantly nuclear localization in CHO-K1 cells (Figure 3.7). We then sought to determine whether reconstitution of the Ku80 mutant XRS-5 cell line with wild-type Ku80 would restore APLF predominantly to the nucleus. We examined XRS-5 cells co-expressing 3x-flag-tagged Ku80 (Ku80-3xFLAG) and APLF\textsuperscript{WT}\textsuperscript{-V5}, and determined that APLF was predominantly localized to the nucleus, as judged indirectly by immunofluorescence microscopy (Figure 3.7).
Figure 3.7 - APLF is pancellular or predominantly cytoplasmic in the absence of Ku. CHO cell lines (parental CHO-K1, Ku80 mutant XRS-5 and Ku80 mutant XRS-5 complemented with wild-type Ku80) ectopically expressing APLF-V5 were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean. The reconstitution of Ku80 in XRS-5 cells was confirmed by immunoblotting whole cell extracts from XRS-5 cells complemented with wild-type Ku80 (inset).

3.3 APLF is predominantly localized to the nucleus in the absence of DNA-PKcs or XRCC4

Since APLF is implicated in the NHEJ pathway, we wanted to determine whether or not APLF depended on the presence of other core NHEJ components in addition to Ku to be localized to the nucleus. We thus examined the subcellular localization of APLF in the absence of the core NHEJ proteins DNA-PKcs and XRCC4. We have previously shown that APLF does not interact with DNA-PKcs, but does interact with the XRCC4-DNA ligase IV complex in an FHA-dependent manner [56]. We ectopically expressed APLFWT-V5 in either DNA-PKcs mutant XRC-1 or XRCC4 mutant
XR-1 CHO cells. Figure 3.8 demonstrates that APLF is predominantly localized to the nucleus in the absence of either DNA-PKcs or XRCC4, as judged indirectly by immunofluorescence microscopy. These data are compared to the results of Figure 3.7 where APLF is highly pancellular and predominantly cytoplasmic in the absence of Ku. Taken together, these results suggest that APLF depends on an interaction with the core NHEJ component Ku, and not other core NHEJ components, to be localized to the nucleus.

Figure 3.8 - APLF is pancellular or predominantly cytoplasmic in the absence of Ku, but is predominantly nuclear in the absence of other NHEJ proteins. DNA-PKcs mutant XRC-1 and XRCC4 mutant XR-1 CHO cell lines ectopically expressing APLF-V5 were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean. These data are compared to the subcellular localization of APLF\textsuperscript{WT} in parental CHO-K1 and Ku80 mutant XRS-5 cell lines shown previously (Figure 3.7).
3.4 APLF is predominantly localized to the nucleus in the absence of XRCC1

APLF has also been found to interact with the SSB repair protein XRCC1 [55]. Interestingly, it has been reported that the FHA-domain-mediated interaction of APLF with XRCC1 promotes the nuclear localization of APLF in the presence of oxidative DNA damage [57]. Therefore, we sought to examine the subcellular localization of ectopically expressed APLF\textsuperscript{WT}-V5 in the XRCC1 mutant EMC-11 CHO cell line. Compared to the subcellular localization status of APLF in the parental CHO-K1 cell line, it appears that there is a slight loss in the nuclear localization of APLF in the absence of XRCC1, but this loss is not as extensive as seen in the absence of Ku (Figure 3.8). These results suggest that there may be multiple ways in which APLF is retained within the nucleus, possibly through interactions with XRCC1, as well as Ku.

We also wanted to check the status of XRCC1 and Ku expression in the XRS-5 and EMC-11 cell lines, respectively, to ensure that each mutant CHO cell line was not altered in the expression of the other respective repair factor, and thus not indirectly affecting the nuclear localization of APLF. To do so, Ku80 mutant XRS-5 and XRCC1 mutant EMC-11 whole cell extracts, alongside human U2OS and hamster parental CHO-K1 whole cell extracts as positive controls, were harvested, and then immunoblotted with anti-Ku80 and anti-XRCC1 antibodies. We observed that XRCC1 expression does not seem to be affected in the Ku80-deficient XRS-5 cell line, suggesting that the loss in nuclear localization of APLF in the XRS-5 cell line is largely due to a loss in the expression of Ku (Figure 3.9). However, Ku80 expression in the XRCC1-deficient EMC-11 cell line seems to be diminished, suggesting that this phenomenon may be
contribution to the slight loss in APLF nuclear localization seen in this cell line (Figure 3.9).

Figure 3.9 - APLF is predominantly nuclear in the absence of XRCC1. XRCC1 mutant EMC-11 CHO cells ectopically expressing APLF-V5 were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean. These data are compared to the subcellular localization of APLF<sup>WT</sup> in parental CHO-K1 and Ku80 mutant XRS-5 cell lines shown previously (Figure 3.7). Human U2OS, parental CHO-K1, Ku80 mutant XRS-5 and XRCC1 mutant EMC-11 whole cell extracts (WCE) were harvested, and then immunoblotted with anti-XRCC1, anti-Ku80, and anti-tubulin antibodies (inset).

3.5 APLF nuclear localization is not dependent on the APLF FHA domain

Previously, it has been shown that APLF interacts with both XRCC4 and XRCC1 in an FHA-dependent manner [54, 56, 57]. Given that the results of Figures 3.7 and 3.8 have demonstrated that the nuclear localization of APLF is not greatly affected in the absence of either XRCC4 or XRCC1, we hypothesized that the FHA arginine-to-
alanine substitution (R27A) within APLF, which abolishes the phosphothreonine-binding ability of the FHA domain to interact with XRCC4 and XRCC1, would not affect the nuclear localization of APLF. To verify this, CHO-K1 cells expressing V5-tagged APLF containing the R27A substitution in the FHA domain (APLF<sup>R27A</sup>-V5) were fixed, and then immunostained with anti-V5 antibody. As judged indirectly by immunofluorescence microscopy, APLF<sup>R27A</sup>-V5 is predominantly localized to the nucleus, a phenotype similar to the ones seen with wild-type APLF in the parental CHO-K1, XRCC4 mutant XR-1 and XRCC1 mutant EMC-11 cell lines (Figure 3.10).

![Graph showing subcellular localization](image)

**Figure 3.10 - The APLF FHA domain is not required for the nuclear localization of APLF.** Parental CHO-K1 cells ectopically expressing APLF<sup>R27A</sup>-V5 were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean. These data are compared to the subcellular localization of APLF<sup>WT</sup> in parental CHO-K1, XRCC4 mutant XR-1 and XRCC1 mutant EMC-11 cell lines shown previously (Figures 3.7-3.9).
3.6 PNK and APTX are predominantly localized to the nucleus in the absence of Ku

Given that APLF is the third member of the unique subclass of FHA domains comprised of PNK and APTX, we then sought to characterize the subcellular localization of both these proteins in the context of Ku. We hypothesized that PNK and APTX would not depend on the presence of Ku to be localized to the nucleus for two reasons. Firstly, PNK and APTX each possess a nuclear localization signal, and thus are imported to the nucleus through the classical import pathway, and secondly, unlike APLF, PNK and APTX do not have direct interactions with Ku. To test this theory, we examined *Ku80* mutant XRS-5 cells, alongside parental CHO-K1 cells as a positive control, expressing V5-tagged PNK or APTX. In CHO-K1 cells, PNK and APTX, like APLF, were both predominantly localized to the nucleus (Figure 3.11). However, in contrast to the pronounced loss of APLF nuclear localization in the XRS-5 cell line, PNK and APTX were still predominantly localized to the nucleus even in the absence of Ku (Figure 3.11). These results demonstrate that PNK and APTX do not depend on the presence of Ku to be localized to the nucleus.
Figure 3.11 - PNK and APTX are predominantly nuclear in the absence of Ku. Parental CHO-K1 or Ku80 mutant XRS-5 CHO cells ectopically expressing PNK-V5 and APTX-V5 were fixed, immunostained with anti-V5 (PNK and APTX) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean. These data are compared to the subcellular localization of APLF^{WT} in CHO-K1 and XRS-5 cells shown previously (Figure 3.7).

3.7 IR-induced DNA damage does not stimulate the nuclear localization of APLF

Since the APLF-Ku interaction exists under basal conditions, and we have thus far examined the putative Ku-dependent APLF nuclear localization under basal conditions, we next sought to determine whether the presence of DNA damage affects the nuclear localization of APLF in the presence of absence of Ku. To do so, CHO-K1 or XRS-5 cells expressing wild-type APLF were mock- or irradiated with 10Gy IR, and fixed either 15, 30 or 60 minutes after irradiation, and then subsequently
immunostained with anti-V5 antibody. As judged indirectly by immunofluorescence microscopy, APLF is predominantly localized to the nucleus in CHO-K1 cells, whereas APLF is largely pancellular and cytoplasmic in XRS-5 cells under no IR conditions (Figure 3.7). These patterns of APLF nuclear localization in both cell lines do not appear to change following IR at various time-points, suggesting that IR-induced DNA damage does not stimulate the nuclear localization of APLF either in the presence or absence of Ku (Figure 3.12).

![Figure 3.12](image)

**Figure 3.12** - IR-induced DNA damage does not stimulate the nuclear localization of APLF in the presence or absence of Ku. Parental CHO-K1 or Ku80 mutant XRS-5 CHO cells ectopically expressing APLF-V5 and APTX-V5 were mock- or irradiated with a 10 Gy dose and fixed at 15, 30 or 60 minutes post-IR. Cells were subsequently immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from two independent experiments, and error bars represent the standard error of the mean.

As a further control to confirm that APLF subcellular localization in the presence or absence of Ku is not affected by IR-induced DNA damage, we also
examined the subcellular localization of wild-type APLF and the Ku-binding domain mutant APLF$^{Δ181-200}$ in human U2OS cells. U2OS cells expressing APLF-V5 or APLF$^{Δ181-200}$-V5 were mock- or irradiated with 10Gy IR, and fixed either 15, 30 or 60 minutes after irradiation, and then subsequently immunostained with anti-V5 antibody. As Figure 3.13 demonstrates, APLF is predominantly localized to the nucleus whereas APLF$^{Δ181-200}$, which has a diminished interaction with Ku, is highly pancellular with some cytoplasmic phenotypes (Figure 3.13). Again, these subcellular localization patterns remain the same even after IR-induced DNA damage.

Figure 3.13 - IR-induced DNA damage does not stimulate the nuclear localization of APLF in the presence or absence of the Ku-binding domain. Human U2OS cells ectopically expressing either APLF-V5 or APLF$^{Δ180-200}$-V5 mock- or irradiated with a 10 Gy dose and fixed at 15, 30 or 60 minutes post-IR. Cells were subsequently immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from two independent experiments, and error bars represent the standard error of the mean.
3.8 APLF nuclear localization is dependent on an interaction with Ku

We have previously demonstrated that Ku interacts with amino acids 180 to 200 on APLF (Chloe Macrae, unpublished data), and have now shown that this interaction is critical for the nuclear localization of APLF. To further assess the dependency of APLF to interact with Ku, and thus be localized to the nucleus, we generated a series of progressively truncated HA-tagged APLF proteins (HA-APLF\textsuperscript{1-200}, HA-APLF\textsuperscript{1-180} and HA-APLF\textsuperscript{1-160}), which were then, expressed in human U2OS cells, fixed and immunostained with anti-HA antibody. We hypothesized that HA-APLF\textsuperscript{1-200}, which contains the Ku-binding domain of amino acids 180 to 200, would be predominantly localized to the nucleus, whereas HA-APLF\textsuperscript{1-180} and HA-APLF\textsuperscript{1-160}, which lose the Ku-binding domain, would have reduced nuclear localizations. As predicted, Figure 3.14 demonstrates that HA-APLF\textsuperscript{1-200} predominantly localizes to the nucleus, while both HA-APLF\textsuperscript{1-180} and HA-APLF\textsuperscript{1-160} are highly pancellular. Interestingly, the ability of HA-APLF\textsuperscript{1-200} to localize predominantly to the nucleus even in the absence of the C-terminal portion of APLF demonstrates that the tandem zinc fingers are not required for the nuclear localization of APLF (Figure 3.14). Collectively, these results confirm that APLF requires the Ku-binding domain (amino acid residues 180-200) to be localized to the nucleus.
Figure 3.14 - APLF requires the Ku-binding domain (amino acid residues 181-200) to be localized to the nucleus. U2OS cells ectopically expressing either APLF<sup>WT</sup>-V5, APLF<sup>1-200</sup>-V5, APLF<sup>1-180</sup>-V5 or APLF<sup>1-160</sup>-V5 were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from three independent experiments, and error bars represent the standard error of the mean.

3.9 APLF depends on an interaction with Ku to be localized to the nucleus and interact with the XRCC4-DNA ligase IV complex

We then performed co-immunoprecipitation experiments with the same truncated HA-tagged APLF proteins to determine how the loss of nuclear localization, resulting from the loss of the Ku-binding domain, affected the ability of APLF to interact with other core NHEJ proteins, namely XRCC4 and DNA ligase IV. HEK293T cells were transfected with either empty vector, HA-APLF, HA-APLF<sup>1-200</sup>, HA-APLF<sup>1-180</sup>
or HA-APLF1-160, and whole cell extracts were immunoprecipitated with anti-HA antibodies to recover APLF and any interacting proteins. The resulting protein complexes were then immunoblotted with anti-DNA ligase IV, anti-XRCC4, anti-Ku80 and anti-HA (APLF) antibodies. We detected interactions of APLF1-200 with Ku, XRCC4 and DNA-ligase IV (Figure 3.15). However, HA-APLF1-180 and HA-APLF1-160 did not complex with Ku, XRCC4 and DNA ligase IV (Figure 3.15). These results suggest that the absence of the Ku-binding domain on APLF not only disrupts the ability of APLF to interact with Ku, but with the other nuclear NHEJ proteins of XRCC4 and DNA ligase IV as well, possibly because APLF has lost its predominantly nuclear subcellular localization.

Figure 3.15 - APLF requires the Ku-binding domain (amino acid residues 181-200) to localize to the nucleus and thus interact with the nuclear XRCC4-DNA ligase IV complex. Whole cell extracts (WCE) from HEK293T cells ectopically transfected with empty vector, HA-APLFWT, HA-APLF1-200, HA-APLF1-180 or HA-APLF1-160 were immunoprecipitated with anti-HA (APLF) antibody and the resulting protein complexes were immunoblotted with DNA ligase IV, XRCC4, Ku80, HA (APLF) or tubulin antibodies as indicated.
We were interested in further examining the inability of APLF to interact with the XRCC4-DNA ligase IV complex in the absence of the Ku-binding domain (amino acids 180 to 200). We wanted to ensure that the removal of the Ku-binding domain on APLF, which results in a diminished interaction with Ku, did not also impair the function of the FHA domain, and thus the ability of APLF to interact with the XRCC4-DNA ligase IV complex. To do so, a simian virus 40 (SV40) large T-antigen nuclear localization signal (NLS) was engineered onto V5-tagged APLF\textsuperscript{Δ181-200} (NLS-APLF\textsuperscript{Δ181-200}-V5) since we previously showed that APLF depends on the presence of an NLS to be localized to the nucleus when the interaction with Ku is disrupted (Figure 3.3). Given that NLS-APLF\textsuperscript{Δ181-200} is completely localized to the nucleus, we wanted to determine whether APLF could now interact with XRCC4 and DNA ligase IV. We performed co-immunoprecipitation assays by ectopically expressing either empty vector, NLS-APLF\textsuperscript{Δ181-200}-V5, APLF-V5 as a positive control, or APLF\textsuperscript{Δ181-200}-V5 as a negative control, in HEK293T cells. The whole cell extracts were immunoprecipitated with anti-V5 antibodies to recover APLF and any interacting proteins, and the resulting protein complexes were then immunoblotted with anti-Ku80, anti-XRCC4, anti-DNA ligase IV and anti-V5 (APLF) antibodies. We observed that while both APLF\textsuperscript{Δ181-200}-V5 and NLS-APLF\textsuperscript{Δ181-200} were unable to bind to Ku, only NLS-APLF\textsuperscript{Δ181-200} was able to associate with XRCC4 and DNA ligase IV (Figure 3.16). These results demonstrate that the FHA domain-mediated interaction of APLF\textsuperscript{Δ181-200} with the XRCC4-DNA ligase IV complex is intact, and that the stability of APLF\textsuperscript{Δ181-200} is not compromised even when the interaction between APLF and Ku is disrupted.
Figure 3.16 - The absence of the Ku-binding domain on APLF does not affect the ability of APLF to interact with the XRCC4-DNA ligase IV complex. Whole cell extracts (WCE) from HEK293T cells ectopically transfected with empty vector, APLF-V5, APLFΔ181-200-V5 or NLS-APLF-V5 were immunoprecipitated with anti-V5 (APLF) antibody and the resulting protein complexes were immunoblotted with Ku80, XRCC4, DNA ligase IV, V5 (APLF) or tubulin antibodies as indicated.

3.10 Tryptophan 189 of APLF strongly contributes to the nuclear localization of APLF

We next wanted to determine whether we could narrow the site of Ku interaction on APLF to a single amino acid residue within the Ku-binding domain. To do so, we created APLF-V5 constructs encoding for alanine substitutions at five different evolutionary conserved sites (Arg-182, Arg-184, Trp-189, Met-190 and Lys-191) singly or in combination (Figure 3.1) to yield the V5-tagged APLF mutant proteins APLF<sup>R182A</sup>-V5, APLF<sup>R184A</sup>-V5, APLF<sup>W189A</sup>-V5, APLF<sup>M190A</sup>-V5, APLF<sup>L191A</sup>-V5, APLF<sup>R(182,184)A</sup>-V5,
and APLF^{VWM(189-191)AAAV5}, respectively. The first five proteins were ectopically expressed alongside APLF-V5 in U2OS cells, and then fixed, immunostained with anti-V5 (APLF) antibody, and characterized by immunofluorescence microscopy. As demonstrated in Figure 3.16, the APLF mutants harbouring the alanine substitution at arginine residues 182 or 184, tryptophan 189, or methionine 190, exhibited loss in nuclear localization to varying degrees. However, the substitution of the tryptophan residue resulted in the greatest loss of APLF nuclear localization, suggesting that this site within the Ku-binding domain is important for mediating an interaction with Ku (Figure 3.17). The APLF mutant protein with an alanine substitution at lysine 191 behaved similarly to wild-type APLF, indicating that this site is not essential for the nuclear localization of APLF, and thus the APLF-Ku interaction (Figure 3.17).
Figure 3.17 - Conserved Tryptophan 189 of APLF strongly contributes to the nuclear localization of APLF. XRS-5 cells ectopically expressing APLF-V5 or APLF-V5 mutants harbouring alanine substitutions at residues 182, 184, 189, 190 or 191 (APLF<sup>R182A</sup>-V5, APLF<sup>R184A</sup>-V5, APLF<sup>W189A</sup>-V5, APLF<sup>M190A</sup>-V5 and APLF<sup>L191A</sup>-V5) were fixed, immunostained with anti-V5 (APLF) antibody, and quantified by immunofluorescence microscopy. Data represent the mean of 100 counted cells from two independent experiments, and error bars represent the standard error of the mean.

3.11 Conserved Tryptophan 189 of APLF is required for the interaction of APLF with Ku

Given that we observed loss of nuclear localization to varying degrees with the five different APLF mutants (Figure 3.17), we next sought to determine whether these mutants exhibited diminished interactions with Ku. Co-immunoprecipitation assays were performed by ectopically expressing either empty vector, APLF-V5 as a positive
control, APLFΔ180-200-V5 as a negative control, APLF^{R(182, 184)}A-V5, APLF^{WML(189-191)AAA-V5}APLF^{R182A-V5}, APLF^{R184A-V5}, APLF^{W189A-V5}, APLF^{M190A-V5} or APLF^{L191A-V5} in HEK293T cells. The whole cell extracts were immunoprecipitated with anti-V5 antibodies to recover APLF and any interacting proteins, and the resulting protein complexes were then immunoblotted with anti-Ku80 and anti-V5 (APLF) antibodies. We observed that both APLF^{R(182, 184)}A-V5 and APLF^{WML(189-191)AAA-V5} were unable to bind to Ku (Figure 3.18). The APLF mutants harbouring the alanine substitution at arginine residues 182 or 184, tryptophan 189, or methionine 190 exhibited either a diminished or a loss of interaction with Ku - phenotypes which were consistent with the general trend of loss in nuclear localization exhibited by each respective mutant (Figures 3.17-3.18). APLF^{L191A-V5}, on the other hand, interacted with Ku similar to wild-type APLF (Figure 3.18).

Figure 3.18 - Conserved residues within the APLF Ku-binding domain are required for the APLF-Ku interaction. Whole cell extracts (WCE) from HEK293T cells ectopically transfected with empty vector, APLF-V5, APLFΔ180-200-V5 or APLF-V5 mutants harbouring alanine substitutions at residues 182, 184, 189, 190 or 191 (APLF^{R(182, 184)}A-V5, APLF^{WML(189-191)AAA-V5}, APLF^{R182A-V5}, APLF^{R184A-V5}, APLF^{W189A-V5}, APLF^{M190A-V5} and APLF^{L191A-V5}) were immunoprecipitated with anti-V5 (APLF) antibody and the resulting protein complexes were immunoblotted with Ku80, V5 (APLF) or tubulin antibodies as indicated.
Because we observed that the APLF tryptophan-to-alanine substitution at residue 189 resulted in a loss of interaction with Ku and the greatest loss of APLF nuclear localization - similar to the phenotype seen with the removal of the entire APLF Ku-binding domain (APLF_{Δ181-200}) - we next wanted to determine whether the interaction with Ku was disrupted even when APLF_{W189A} is completely localized to the nucleus. To do so, we created a derivative of APLF_{W189A-V5} that harboured an SV40 large T-antigen NLS (NLS-APLF_{W189A-V5}), and then performed co-immunoprecipitation assays by ectopically expressing either empty vector, APLF-V5 as a positive control, APLF_{Δ180-200-V5}, APLF_{W189A-V5}, or NLS-APLF_{W189A-V5} in HEK293T cells. The whole cell extracts were immunoprecipitated with anti-V5 antibodies to recover APLF and any interacting proteins, and the resulting protein complexes were then immunoblotted with anti-Ku80 and anti-V5 (APLF) antibodies. As Figure 3.19 demonstrates, NLS-APLF_{W189A-V5} is unable to associate with Ku, suggesting that Trp-189 is critical for mediating an interaction with Ku.
Figure 3.19 - Conserved Tryptophan 189 of APLF is required for the interaction of APLF with Ku. Whole cell extracts (WCE) from HEK293T cells ectopically transfected with empty vector, APLF-V5, APLF_{181-200-V5} or APLF-V5 and NLS-APLF-V5 mutants harbouring alanine substitutions at residues 189 (APLF_{W189A-V5} and NLS-APLF_{W189A-V5}) were immunoprecipitated with anti-V5 (APLF) antibody and the resulting protein complexes were immunoblotted with Ku80, V5 (APLF) or tubulin antibodies as indicated.

3.12 Amino acid residues 569 to 598 within the Ku80 C-terminus region strongly contribute to the nuclear localization of APLF

To further investigate the APLF-Ku interaction, we next focused our attention on mapping the APLF-binding site on Ku. APLF has been shown to directly interact with the Ku80 subunit, and not the Ku70 subunit, of Ku [54]. Although the two Ku subunits possess highly conserved regions of homology over the length of their proteins, the carboxy-terminal domain of Ku80 is longer and divergent from the Ku70 subunit (Figure 1.3). Indeed, it is within this region where DNA-PKcs has been found to interact [15, 16]. Accordingly, we wondered whether this region could also be important for mediating a specific interaction with APLF. To address this question, we
created a series of C-terminal deletions on Ku80. We generated an eGFP-tagged Ku80 protein, which was then used to create the truncated eGFP-tagged Ku80 constructs (eGFP-Ku80<sub>1-720</sub>, eGFP-Ku80<sub>1-598</sub> and eGFP-Ku80<sub>1-569</sub>). All of these truncated Ku80 proteins have been previously reported to stably exist in association with the Ku70 subunit to form the Ku heterodimer, as well as to associate with DNA ends [15, 92]. Since we previously showed that reconstitution of the Ku80 mutant XRS-5 cell line with wild-type Ku80 restored the nuclear localization of APLF (Figure 3.7), we now sought to determine whether reconstitution of this cell line with any of these Ku80 C-terminal mutants resulted in a disruption of the APLF-Ku interaction, and thus a disruption in APLF nuclear localization. To do so, we co-transfected XRS-5 cells with APLF-V5 and either eGFP-Ku80, eGFP-Ku80<sub>1-720</sub>, eGFP-Ku80<sub>1-598</sub> or eGFP-Ku80<sub>1-569</sub>, and then fixed, immunostained with anti-V5 (APLF) antibody and DAPI nuclear stain, and imaged the cells by immunofluorescence microscopy. We analyzed anti-V5 (APLF) expression only in cells expressing the eGFP Ku80 constructs. As demonstrated in Figure 3.20, the expression of eGFP-Ku80<sub>1-569</sub> resulted in a significant loss of APLF nuclear localization, whereas the expression of eGFP-Ku80<sub>1-720</sub> and eGFP-Ku80<sub>1-598</sub> both resulted in a predominantly nuclear localization of APLF. These results suggest that the region lying between amino acids 569 and 598 on Ku80 is important for mediating an interaction with APLF, and is thus critical for the nuclear localization of APLF.
3.13 APLF interacts with Ku in the absence of the Ku80 C-terminus region

Because the nuclear localization of APLF appears to be dependent on the region encompassing amino acids 569 to 598 within the Ku80 C-terminus, we wanted to determine whether or not this region on Ku80 is required for directing interactions with APLF. To assess this, we performed pull-down assays utilizing purified recombinant GST-fusion protein of full-length APLF (GST-APLF). The GST-APLF protein was immobilized on glutathione-sepharose beads and incubated with whole cell
extracts from XRS-5 cells transfected with empty vector, eGFP-Ku80 or eGFP-Ku80\(^{1-569}\) to recover any interacting proteins. The resulting protein complexes were then immunoblotted with anti-Ku70 and anti-Ku80 antibodies. As demonstrated in Figure 3.21, the C-terminal region of Ku80 is not required for the association of Ku80 with APLF, as eGFP-Ku80\(^{1-569}\) was able to bind to APLF in a manner comparable to eGFP-Ku80. Moreover, although we observed that the introduction of both eGFP-Ku80 and eGFP-Ku80\(^{1-569}\) into the XRS-5 cells restored Ku70 protein levels, the cellular levels and association of Ku70 with eGFP-Ku80\(^{1-569}\) was diminished slightly compared to eGFP-Ku80 (Figure 3.21).

![Figure 3.21 - APLF associates with Ku80 in the absence of the Ku80 C-terminus region. Pull-down assays were performed utilizing purified recombinant GST-APLF protein immobilized on glutathione-sepharose beads. Beads were incubated with whole cell extracts (WCE) from XRS-5 cells ectopically expressing empty vector, eGFP-Ku80\(^{WT}\), or eGFP-Ku80\(^{1-569}\) in the presence of 50 \(\mu\)g/ml of ethidium bromide. The resulting protein complexes were immunoblotted with anti-Ku70 or anti-Ku80 antibodies as indicated.](image-url)
3.14 The N-terminal region of APLF is required for APLF-dependent NHEJ

Given that we have previously shown that APLF interacts with core NHEJ components, specifically Ku and the XRCC4-DNA ligase IV complex, and that down-regulation of APLF in human cells results in an NHEJ defect, we next sought to define the mechanism of APLF involvement in NHEJ [56]. Since protein function can largely be inferred by the presence of its structural domains, we focused on systematically delineating which APLF domains – be it the FHA, Ku-binding or zinc finger domains – are sufficient for the function of APLF in NHEJ. To achieve this, we employed a cell-based NHEJ assay – termed random plasmid integration assay (PIA) – that serves as a global measure of the efficiency of NHEJ. This assay was established on the concept that NHEJ is the predominant mechanism for the nonhomologous integration of foreign DNA into endogenously occurring double-strand breaks (DSBs) within the genome. Foreign DNA is recognized and modified by NHEJ components before becoming integrated into the genome of cells in culture, and the efficiency of this random genome integration is measured by colony formation. However, the downregulation of a core NHEJ component results in an NHEJ defect, whereby there is decreased integration of foreign DNA into the genome, and thus decreased colony formation. We previously demonstrated that the transient depletion of APLF using siRNA in human cells markedly decreased plasmid integration, thus implicating APLF to facilitate NHEJ repair [56]. We sought to further investigate this NHEJ defect by carrying out reconstitution experiments using various APLF mutants in an APLF-depleted cell line to determine whether or not they can mediate the full rescue of
the NHEJ defects, thereby indicating the significance of these mutations on APLF function.

To conduct the reconstitution experiments, we previously established a human U2OS cell line stably depleted in APLF (APLF-depleted cells) using RNA interference (RNAi) and also generated a V5-tagged construct encoding an RNAi-resistant wild-type APLF (RNAi-resistant APLF\textsuperscript{WT}) protein. Reconstitution of the APLF-depleted cell line with RNAi-resistant APLF\textsuperscript{WT} completely rescues APLF expression in the APLF-depleted cell line. We next generated various RNAi-resistant APLF mutants to determine whether or not they are able to rescue the NHEJ defect when reconstituted in the APLF-depleted cell line. If any one of these mutants is unsuccessful in rescuing the NHEJ defect, then we can delineate which specific APLF domains, and thus which APLF interactions with other DNA repair proteins, are critical for the function of APLF in NHEJ. To achieve this, we targeted specific sites on APLF: (1) the FHA-inactivating arginine-to-alanine substitution at residue 27 to disrupt binding to the XRCC4-DNA ligase IV complex (RNAi-resistant APLF\textsuperscript{R27A}); (2) the tryptophan-to-alanine substitution at residue 189 to disrupt binding to Ku, along with the engineered SV40-large T-antigen NLS to allow APLF to localize to the nucleus even when the APLF-Ku interaction is disrupted (RNAi-resistant NLS-APLF\textsuperscript{W189A}); (3) the truncated APLF protein corresponding to amino acid residues 1-200 within the N-terminal region of APLF, which encompasses the FHA and Ku-binding domains, but loses the C-terminal tandem zinc fingers (RNAi-resistant NLS-APLF\textsuperscript{1-200}); and (4) the zinc finger-inactivating cysteine-to-glycine substitutions at residues 379, 385, 421 and 427 to disrupt the functionality of these two motifs (RNAi-resistant APLF\textsuperscript{ZF1+2m}). We hypothesized that
the N-terminal portion of APLF, which directs interactions with core NHEJ components, namely Ku and the XRCC4-DNA ligase IV complex, will be sufficient to reconstitute APLF-dependent NHEJ. In contrast, we hypothesized that the C-terminal portion of APLF, which contain the zinc fingers, may be important for a role of APLF in other DNA repair activities (for instance, through the zinc-finger-mediated interaction with PARP-1) and thus will be dispensable for NHEJ function.

To perform the PIA, linearized plasmid DNA containing a puromycin resistance marker was transfected into APLF-depleted cells transfected with empty vector, RNAi-resistant APLF<sup>WT</sup>, RNAi-resistant NLS-APLF<sup>WT</sup>, RNAi-resistant APLF<sup>R27A</sup>, RNAi-resistant NLS-APLF<sup>W189A</sup>, RNAi-resistant APLF<sup>1-200</sup>, or RNAi-resistant APLF<sup>ZF1+2m</sup>. Cells were subsequently plated in selection media containing puromycin to allow for integration of the plasmid DNA into the genome of cells in culture, which was measured by the formation of puromycin-resistant colonies (Figure 3.22). Colony counts indicate the efficiency of plasmid DNA integration, and thus NHEJ efficiency, so the number of colonies formed when APLF-depleted cells were reconstituted with RNAi-resistant APLF<sup>WT</sup> was set at 100% integration. As expected, reconstitution with RNAi-resistant APLF<sup>R27A</sup> or RNAi-resistant NLS-APLF<sup>W189A</sup> resulted in significant decreases of plasmid integration frequency, while reconstitution with RNAi-resistant APLF<sup>1-200</sup> or RNAi-resistant APLF<sup>ZF1+2m</sup> was able to rescue the NHEJ defect comparable with RNAi-resistant APLF<sup>WT</sup> (Figure 3.23). Interestingly, reconstitution with RNAi-resistant NLS-APLF<sup>WT</sup> resulted in greater plasmid integration compared to RNAi-resistant APLF<sup>WT</sup> (Figure 3.23). This observation is consistent with fact that the presence of an NLS completely localizes APLF to the nucleus - where NHEJ takes place.
- whereas without an NLS, approximately 5% of wild-type APLF is pancellular and not localized completely to the nucleus, and thus cannot participate in nuclear NHEJ (Figure 3.17). Taken together, these results suggest that the N-terminal region, containing the XRCC4-Ligase IV and Ku binding domains, is required for APLF-dependent NHEJ.

Figure 3.22 - Plasmid integration assay. Linearized plasmid DNA, containing an antibiotic resistant marker (puromycin), can be integrated into the genome of mammalian cells in culture via NHEJ. The efficiency of plasmid DNA integration serves as a direct measure of NHEJ efficiency.
Figure 3.23 - The N-terminal region of APLF, containing the XRCC4-Ligase IV and Ku binding domains (spanning amino acids 1-200), is required for APLF-dependent NHEJ. U2OS cells stably depleted in APLF were transfected with empty vector, RNAi-resistant APLF<sup>WT</sup>, RNAi-resistant NLS-APLF<sup>WT</sup>, RNAi-resistant APLF<sup>R27A</sup>, RNAi-resistant NLS-APLF<sup>W189A</sup>, RNAi-resistant APLF<sup>1-200</sup>, or RNAi-resistant APLF<sup>ZF1+2m</sup>, re-transfected with linearized plasmid DNA containing a puromycin resistance cassette 48 hours later, and then re-plated in triplicate at low density in media containing puromycin 24 hours later. The resulting puromycin-resistant colonies were stained and quantified 10 days later. The relative plasmid integration with RNAi-resistant APLF<sup>WT</sup> (striped bar) was set at 100% integration. The data shown in this representative experiment is the mean number of colonies from three plates (n=3), with error bars representing the standard deviation.
4.1 The Ku-binding domain on APLF is mapped to amino acids 180-200

Given that the APLF-Ku interaction is likely important for the function of APLF in NHEJ, we sought to better understand the mechanism of this interaction. The results presented in this thesis show that the Ku-binding domain on APLF is mapped to amino acids 181-200 - a region that lies between the FHA and tandem PBZ domains. The importance of this amino acid region is highlighted by the fact that it is evolutionary conserved within APLF vertebrate homologues, and indeed, removal of this region on APLF resulted in a diminished interaction with Ku. Moreover, surprisingly, we showed that APLF nuclear localization is disrupted either in the absence of the Ku-binding domain, or when expressed in Ku-deficient mammalian CHO cells. APLF nuclear localization was rescued either upon reconstitution of Ku in the Ku-deficient cells, or upon addition of a NLS onto APLF, which circumvented the requirement for an interaction with Ku.

We further narrowed down the site of interaction with Ku to include an evolutionary conserved amino acid residue, tryptophan 189, which resides within the Ku-binding domain of APLF. Substitution of this residue resulted in both a loss of interaction with Ku and a loss of APLF nuclear localization - these resulting phenotypes were similar to the ones exhibited when the entire Ku-binding domain was removed from APLF. The distinguishing characteristics of the tryptophan residue may be of importance in its ability to mediate an interaction with the Ku protein. This non-polar hydrophobic residue contains a large planar and aromatic side chain - an indole functional group - which can participate in stabilizing protein folds through hydrogen bond interactions. Moreover, it is likely that other residues, such as arginine
184 and methionine 190, that surround Trp-189 within the Ku-binding domain, also contribute to the interaction with Ku, since substitution of these residues resulted in abolished or decreased interactions with Ku, and significant losses in APLF nuclear localization. To further examine whether the APLF Ku-binding domain is sufficient to direct interactions with Ku, it will be valuable to assess whether the isolated peptide fragment encompassing amino acids 181-200 can interact with Ku in pull-down assays. Additionally, if the isolated peptide directly competes with intact APLF protein for binding to Ku in a titration assay, it would further support the notion that this region of APLF is required for Ku binding. If the isolated Ku-binding domain is found to be sufficient for binding to Ku, then it would be important to determine the strength of binding and the stoichiometry of the peptide in complex with Ku using direct binding studies, such as isothermal titration calorimetry. Structural studies can also be used to determine the structural basis for the interaction between the APLF Ku-binding domain and Ku.

4.2 The APLF-Ku interaction is important for the nuclear localization of APLF

In contrast to APLF, PNK and APTX both possess a NLS, and are thus transported into the nucleus via the classical import pathway, and neither interacts with Ku. Consistent with this, we demonstrated that both these proteins are able to completely localize to the nucleus in the absence of Ku. Therefore, it seems that a Ku-dependent mode of nuclear localization is unique to APLF. It is conceivable that APLF may be able to piggyback onto Ku - which has an NLS and is thus imported into the nucleus via the classical nuclear transport pathway - in order to be localized to
the nucleus and participate in NHEJ or other DNA repair pathways [14]. However, it is not clear what proportion of endogenous Ku is engaged in complex with APLF under basal conditions, and whether the entire proportion of these complexes is important for mediating the nuclear localization of APLF. To address these questions, the subcellular distribution of APLF-Ku complexes can be quantitatively evaluated using biochemical cell fractionation and gel-filtration chromatography assays.

Although our findings implicate a role for Ku in mediating the nuclear localization of APLF, we cannot exclude the possibility that APLF may also be able to passively diffuse into the nucleus from the cytoplasm. Our results demonstrated that a small proportion of APLF is localized to the nucleus in Ku-deficient hamster cells under basal conditions, suggesting that APLF may be able to passively diffuse into the nucleus even in the absence of Ku. Indeed the molecular weight cut-off range for passive diffusion of proteins into the nucleus is 50-60 kDa, and APLF falls into this range with a molecular weight of 57 kDa [71]. Interestingly, there are many reported examples of small molecular weight and NLS-deficient repair proteins that primarily rely on interactions with NLS-containing proteins, which are involved in their respective repair pathway, to bring them into the nucleus in a DNA-damage dependent manner [85-87]. This may represent a level of spatio-temporal regulation that is more efficient than simple diffusion - the pre-formed complex of the NLS-deficient and NLS-proficient proteins in the cytoplasm can be targeted to the nucleus, and thus to sites of damage, where they can directly integrate into repair complexes in a timely manner. This may then enable the proteins to perform a concerted function relating to their respective DNA repair pathway in the nucleus. APLF,
however, may not follow this type of regulatory system since we demonstrated that nuclear localization of either APLF$^{\Delta 180-200}$ in human cells or wild-type APLF in Ku-deficient hamster cells is not up-regulated in the presence of IR, a DNA-damaging agent that triggers the NHEJ pathway. We also have preliminary results indicating that there are no differences in wild-type APLF nuclear localization upon the treatment of the DNA damaging agents bleomycin and hydrogen peroxide ($H_2O_2$) in human cells (data not shown). Thus, based on these results, it appears that Ku-mediated nuclear localization of APLF occurs under basal conditions, and is not altered in the presence of DNA damage. These data conflict with a study by Iles et al., which found that APLF is largely pancellular, and cytoplasmic in XRCC1-deficient EM9 CHO cells under basal conditions, but upon reconstitution of these cells with XRCC1, APLF nuclear localization increases [57]. Moreover, $H_2O_2$ treatment further stimulated the nuclear localization of APLF when co-expressed with XRCC1 in these cells. In contrast, when we examined the XRCC1-deficient EMC-11 cell line under basal conditions, APLF was found to be predominantly nuclear. Collectively, these conflicting observations raise an interesting question about the nuclear localization of APLF in different cell types and organisms. It is possible that certain cell types respond to DNA damage stimuli in different ways, thus, it would be important to further characterize the nuclear localization of APLF in XRCC1-deficient EM9 cells under basal and $H_2O_2$ damage conditions.

The availability of more than one nuclear import mechanism - passive diffusion and piggyback import - may also exist to ensure that repair proteins are still able to localize to the nucleus when one import mechanism is downregulated. Furthermore,
the APLF-Ku interaction may serve to stabilize APLF, thus up-regulating the nuclear entry of APLF compared to passive diffusion of APLF alone. Such situations may exist with APLF, hence it would be interesting to determine whether there are distinct subcellular pools of APLF that are either able to passively diffuse into the nucleus, or need to bind to Ku to be actively transported into the nucleus. This can be assessed using nuclear import assays. In one such assay, cells are permeabilized with a detergent, such as digitonin, which removes essential cytosolic factors, like importin-αβ and Ran-GTP, which are required for efficient in vitro nuclear import [93, 94]. These cells can be incubated with fluorescently tagged APLF, supplemented with or without fluorescently tagged Ku, cell cytosol (containing essential transport factors), and an energy source (such as GTP), and then visualized by confocal microscopy. If APLF accumulates in the nucleus in the absence of Ku, cytosol and an energy-source, then APLF is likely able to passively diffuse into the nucleus, whereas if APLF accumulates in the nucleus only in a Ku-, cytosol- and energy-dependent manner, then APLF is likely transported into the nucleus through interaction with Ku via the classical nuclear import pathway. Alternatively, if APLF is able to accumulate to the nucleus under both scenarios, then APLF may depend on both passive diffusion and Ku-interactions to be localized to the nucleus. It would also be interesting to determine whether these modes of nuclear localization are maintained in equilibrium at all times or are cell-cycled regulated.

Intriguingly, the putative Ku-dependent mode of APLF nuclear import suggests a novel function for Ku since this has not been extensively described before in the literature. One line of evidence implicating Ku in regulating the subcellular
distribution of its interacting partners is in a study that identified that the Ku70 subunit of Ku can bind to pro-apoptotic protein BAX in the cytoplasm, and thus inhibit BAX-mediated apoptosis by preventing its relocalization to the mitochondria [21]. Furthermore, as previously mentioned, complex formation between Ku70 and Ku80 plays an important role in regulating the targeting of the Ku heterodimer to the nucleus under certain conditions, presumably when the functional heterodimer complex is required to mediate certain nuclear processes such as NHEJ [82]. However, the lack of any prior research highlighting a Ku-dependent mode of nuclear import may simply be due to the fact that no one has ever thought to characterize this aspect of Ku function. As such, future studies detailing the mode of nuclear localization of other Ku-interacting proteins, and whether they, like APLF, depend on an interaction with Ku to localize to the nucleus, would be of great interest. Indeed, in a way, the APLF Ku-binding domain serves as an alternative nuclear localization element for APLF, and it is possible that this domain is found in other proteins where it serves a similar purpose. Thus, we looked to see if the human APLF Ku-binding domain sequence was present in other Ku-interacting proteins [80]. Interestingly, there are some sequence similarities - particularly with the tryptophan, arginine and methionine residues that we have shown to be important for APLF-Ku binding - in the human species of WRN and DNA-PKcs. Ku-binding regions have already been mapped to WRN, and one of these regions contains the sequence similarities that we identified based on homology to the APLF Ku-binding domain [95]. The Ku-binding region on DNA-PKcs, on the other hand, has yet to be defined, but structural studies suggest that Ku binds to the central region of DNA-PKcs, which is where the APLF Ku-binding
sequence similarities are found [96]. Thus, it would be of interest to follow-up on our findings by conducting mutagenesis studies targeting our detected sites of sequence similarities in WRN and DNA-PKcs to determine if these respective regions are related to Ku-binding. Additionally, if the isolated APLF Ku-binding domain is found to be sufficient for Ku interaction, this domain can be fused to a cytoplasmic NLS-deficient protein to determine whether the protein acquires the ability to interact with Ku and thus localize to the nucleus. In this way, we can determine whether the APLF Ku-binding domain functions as an alternative nuclear localization element.

Furthermore, the nuclear export of APLF warrants examination. It is unknown if APLF possesses a functional NES, which is characterized as a leucine-rich peptide region that is recognized by Exportin-1 [73]. Sequence analysis of APLF suggests that there are some regions that contain closely-spaced leucine residues, particularly in the amino terminal portion of APLF. Therefore, it will be advantageous to first determine whether APLF has a functional NES. A useful agent enabling the identification of NES-proficient proteins is the Exportin-1 specific inhibitor, leptomycin B (LMB), which disrupts the interaction between Exportin-1 and NES-proficient proteins, resulting in the nuclear retention of the NES-proficient protein [97]. For example, if the highly pancellular APLF\textsuperscript{R184A} mutant responds to LMB in human cells and accumulates in the nucleus, then APLF may depend on nuclear export via Exportin-1. As such, it would then prove fruitful to conduct mutagenesis studies targeting APLF leucine sites to identify the NES. Conversely, if APLF\textsuperscript{R184A} does not respond to LMB, and still exhibits a loss in nuclear localization, then it is likely that APLF does not act through the Exportin-1-mediated nuclear export pathway. In that
case, it may be that APLF depends on passive diffusion, Ku, or other interacting NES-proficient partners to shuttle out of the nucleus.

4.3 Other core NHEJ proteins and APLF-interacting proteins do not strongly contribute to the nuclear localization of APLF

While our results strongly suggest that Ku is the major protein responsible for APLF nuclear localization, we cannot exclude the possibility that multiple APLF interactions may contribute to its nuclear localization. It is unlikely that other core NHEJ proteins promote APLF nuclear localization since we demonstrated that APLF maintains a predominantly nuclear localization in the absence of either core NHEJ protein DNA-PKcs or XRCC4. These results reflect the fact that DNA-PKcs does not directly interact with APLF and, therefore, would not be able to piggyback APLF into the nucleus if such a mechanism is adopted for the nuclear import of APLF, whereas the APLF-XRCC4 interaction may facilitate other aspects of APLF function - most likely pertaining to NHEJ - once APLF is in the nucleus.

However, APLF nuclear localization may be mediated in part by XRCC1 as APLF exhibited a slight loss in nuclear localization in the absence of this SSB protein under basal conditions, though this loss was not as extensive as seen in the absence of Ku. As mentioned previously, this finding is inconsistent with the Iles et al. report that XRCC1 promotes the nuclear localization of APLF under both basal and DNA damage conditions; however, this discrepancy may derive from the use of different cell lines [57]. Strikingly, we observed that Ku80 expression is diminished in the XRCC1-deficient EMC-11 cell line, suggesting that this phenomenon may be contributing to the slight loss in APLF nuclear localization in this cell line. It is difficult to explain this
observation since Ku and XRCC1 participate in separate repair pathways – DSB and SSB repair, respectively. There are detailed characterizations of various CHO cell lines in terms of whether the loss of expression in one repair factor affects the expression of another factor that functions in the same repair pathway. For example, there are almost undetectable levels of DNA ligase IV in XRCC4-deficient XR-1 CHO cells, suggesting that XRCC4 stabilizes and regulates the cellular levels of the DNA ligase IV protein [30]. However, there are no examples in the literature pertaining to the effect of the absence of one repair factor on the expression of another factor functioning in a different pathway. Thus, it would be worthwhile to gain more insight into our finding by assessing the expression status of core NHEJ proteins in the EMC-11 cell line, and correlate expression levels with functional repair assays.

Ultimately, it is conceivable that multiple protein interactions, whether through Ku, XRCC1 or other unknown APLF-interacting partners, contribute to APLF nuclear localization, and it would be interesting to determine what proportions and under what conditions these APLF-interacting partners are responsible for the nuclear localization of APLF.

4.4 The FHA and PBZ domains are not required for the nuclear localization of APLF

We observed that the first 200 amino acid residues of APLF, which direct interactions with XRCC4-DNA ligase IV and Ku, but lack the C-terminal PBZ domains, did not impair APLF nuclear localization, suggesting that the PBZ domains contribute to another aspect of APLF function. Indeed consistent with this notion, the PBZ
domains have been shown to be essential for the PARP-1- and PAR-dependent recruitment of APLF to sites of DNA damage [57, 59, 60].

In addition, we observed that the highly conserved amino-terminal FHA domain is not required for the nuclear import of APLF and, as such, may instead serve to facilitate the participation of APLF in DSB and SSB repair - through interactions with XRCC4 and XRCC1, respectively - once APLF is in the nucleus. Since APLF has no identifiable catalytic activity, it is possible that APLF functions as an adaptor protein to facilitate interactions between the core NHEJ proteins. Indeed, Ku has been shown to bind to XRCC4 and DNA-ligase IV, which helps mediate the recruitment of the XRCC4-DNA ligase IV complex to DSBs in NHEJ [33]. Thus, it is conceivable that either a pre-formed APLF-Ku complex may facilitate the recruitment of the XRCC4-DNA ligase IV complex to DSBs or, alternatively, Ku may facilitate the recruitment of a pre-formed APLF-XRCC4-DNA ligase IV complex to DSBs. The recruitment and formation of these NHEJ protein complexes at a DSB may not follow a strict order, but may depend on the different spatio-temporal events of the current state of the cell. Future work detailing the formation of APLF complexes with these core NHEJ proteins may provide insight into the spatio-temporal activities of NHEJ, which can be further assessed with recruitment studies. For instance, we can use live cell imaging to investigate whether the lack of an interaction with Ku affects the recruitment kinetics of APLF to sites of laser-induced DNA damage.
4.5 Mapping the APLF-binding site on Ku

To further investigate the APLF-Ku interaction, we sought to map the APLF-binding site on the Ku80 subunit of the Ku heterodimer, which has been shown to directly interact with APLF [54]. We found that the region encompassing amino acid residues 569 to 598 on Ku80 is critical for the nuclear localization of APLF, suggesting that loss of this Ku80 region disrupts the interaction with APLF. However, upon further examination, we found that the absence of this C-terminal region on Ku80 does not impair the physical interaction between APLF and Ku, as judged by in vitro pull-down assays. The observation that Ku70 levels are slightly diminished in the presence of Ku80\textsuperscript{1–569} in the pull-down suggests that the absence of the Ku80 carboxy-terminus region may affect the ability of Ku70 and Ku80\textsuperscript{1–569} to form an intact, functional heterodimer. This may explain why APLF undergoes a loss in nuclear localization in the presence of Ku80\textsuperscript{1–569} even when the APLF Ku-binding domain on Ku80 remains intact. Thus, APLF may be dependent on binding to the intact Ku heterodimer to localize to the nucleus. Although Weterings et al. have previously demonstrated that loss of the Ku80 carboxy-terminus does not affect binding of the Ku heterodimer to DNA, suggesting that the stability of Ku is intact, they alternate between the use of the Ku80\textsuperscript{1–598} and Ku80\textsuperscript{1–569} mutants in their study [92]. As such, the subtle changes in Ku70 expression may have been overlooked while using Ku80\textsuperscript{1–598} instead of Ku80\textsuperscript{1–569} in certain experiments. Indeed we found that Ku80\textsuperscript{1–569}, and not Ku80\textsuperscript{1–598}, disrupted the nuclear localization of APLF, so it will be important to first determine if there is a difference in Ku70 expression between Ku80\textsuperscript{1–598} and Ku80\textsuperscript{1–569} reconstituted in XRS-5 cells and used in a pull-down assay with purified APLF. If Ku70
levels are only diminished in the presence of Ku80\textsuperscript{1-569}, then the region encompassing amino acids 569 to 598 on Ku80 may be required for the stability of the Ku heterodimer. To further examine this, we can create a Ku70 double mutant lacking both its NLS and region of Ku80 heterodimerization, thus abrogating its ability to localize to the nucleus independent of Ku80 and to interact with Ku80, respectively. This Ku70 mutant, along with wild-type Ku80, can be reconstituted into Ku-deficient XRS-5 cells to determine whether the loss of intact Ku70, but not Ku80, affects APLF nuclear localization. Similarly, we can also create a Ku80 double mutant targeting both its nuclear and heterodimerization function, and express it alongside wild-type Ku70 in XRS-5 cells. If APLF is unable to localize to the nucleus in the presence of either double mutant, this would suggest that heterodimerization of Ku is important for mediating the nuclear entry of APLF.

It is also possible that another unknown factor is affected by the absence of the Ku80 C-terminal region, and thus indirectly disrupts the nuclear localization of APLF even though the APLF-binding domain on Ku80 still remains intact. We are currently examining whether purified APLF and Ku80\textsuperscript{1-569} can directly interact in pull-down assays, which will assist in clarifying this possibility. Moreover, it is conceivable that the APLF-binding site on Ku80 does not reside within the divergent C-terminal region, but instead resides within the amino-terminal or core portion of Ku80. Although the amino-terminal and core regions of the Ku70 and Ku80 subunits share high structural similarity, they are low in overall sequence identity (~15%), so it is possible that the APLF-binding site on Ku80 lies somewhere in these regions [16].
Hence, it will be important to conduct additional mapping studies to delineate the APLF-binding site and further characterize the APLF-Ku interaction [19].

4.6 The Function of APLF in NHEJ

We found that the amino-terminal portion of APLF, which directs interactions with the core NHEJ proteins, XRCC4-DNA ligase IV and Ku, through the FHA and Ku-binding domains, respectively, is sufficient for APLF-dependent NHEJ, as measured by the cell-based PIA. We also demonstrated that the carboxy-terminal portion of APLF, which contains the PBZ domains, is likely dispensable for NHEJ function and instead important for a role of APLF in other DNA repair activities, for instance through the PBZ-mediated interactions with PARP-1 and PAR. It is interesting to note that PARP1 activity varies depending on the type of DNA ends, and has preferences for blunt ends, 5’ or 3’ single-base overhangs and 3’ phosphoglycolate [70, 98]. Since we utilized a DNA substrate with incompatible 5’ two-base overhangs in the PIA, it is conceivable that PARP1 was not activated in the presence of this DNA substrate and, thus, was unable to modify and/or interact with APLF - this may explain why we did not see a NHEJ defect with the APLF PBZ mutant. Therefore, it may be important to assess whether different DNA substrates with different termini modifications affect the efficiency of APLF-dependent NHEJ.

Although the results from the PIA provide a useful and comprehensive readout of which APLF domains seem to be important for a role of APLF in NHEJ, we can improve this assay by including more controls. For instance, an additional positive control to compare the decreased plasmid integration, and thus deficiency of NHEJ,
to the APLF-depleted cell lines would be useful. The downregulation of DNA ligase IV, which is a core component of NHEJ, will be expected to result in such a NHEJ defect. Thus, we are currently working on generating a stably-depleted DNA-ligase IV U2OS cell line using DNA ligase IV-specific RNAi.

Though we now have a clue as to which APLF domains and which APLF-interacting partners allow APLF to play a role in NHEJ, the molecular mechanism has not been elucidated for these interactions. As mentioned previously, APLF may serve as an adaptor to facilitate end-joining in NHEJ through interactions with both XRCC4 and Ku. Once at the DSB site, APLF may stabilize interactions between Ku and XRCC4-DNA ligase IV, thus promoting efficient end-joining. To further characterize APLF-dependent end-joining, we can use in vivo plasmid end joining assays with reporter plasmids that contain the firefly luciferase gene. Linearized forms of these plasmids with different termini modifications (such as blunt ends and incompatible overhangs), which will require end-processing activity before they can be re-ligated, can be transfected into wild-type and APLF-depleted cells to assess end-joining. The luciferase gene will only be expressed upon recirculization of the plasmid, indicating that end-joining occurred. If APLF-depleted cells demonstrate reduced end-joining frequency compared to the control wild-type cell line, this infers that APLF enhances the efficiency of end-joining. It would also be interesting to determine whether reconstitution of the APLF-depleted cell line with various APLF mutants, such as APLF<sub>1-200</sub>, APLF<sup>R27A</sup> and NLS-APLF<sup>W189A</sup>, are able to rescue the end-joining defect. In addition, we can conduct in vitro biochemical reconstitution assays using purified recombinant APLF (wild-type and the APLF<sub>1-200</sub>, APLF<sup>R27A</sup> and APLF<sup>W189A</sup> mutants) and
core NHEJ proteins (including the Ku heterodimer, DNA-PKcs, and the XRCC4-DNA ligase IV complex) to assess whether APLF enhances the re-joining of radiolabelled DSBs with different termini modifications.

The role of APLF in end-joining can also be assessed in the context of the APLF-Ku interaction. As previously mentioned, Ku has been recently reported to have enzymatic activity and function as a 5’-deoxyribose-5-phosphate lysase to efficiently repair DSBs with associated abasic sites by NHEJ [37]. It would be interesting to assess whether APLF enhances the efficiency of Ku lyase activity in this DSB repair pathway. We can carry out in vitro end-joining and lyase reactions, whereby radiolabelled DNA substrates with abasic sites located near their 5’ termini are incubated with recombinant purified Ku, in the presence or absence of purified APLF (wild-type and APLF<sup>W189A</sup> mutant). For example, if the presence of purified recombinant APLF, and not APLF<sup>W189A</sup>, enhances the end-joining of DNA substrates, then we can infer that the APLF-Ku interaction enhances the enzymatic ability of Ku to resolve abasic DSBs in the NHEJ pathway.

4.7 Other functional roles and consequences of the APLF-Ku interaction

It is possible that in addition to promoting APLF nuclear localization and APLF function in NHEJ, the APLF-Ku interaction has other functional consequences. There may exist distinct subcellular pools of APLF-Ku complexes that are important for facilitating the nuclear localization of APLF, the involvement of APLF in NHEJ, or the involvement of APLF in other Ku-mediated cellular processes (such as apoptosis and telomere maintenance), respectively [20, 21]. These scenarios may be mutually
exclusive, or may operate in concert. We can assess this using biochemical and biophysical microscopy studies.

We can first determine if the loss of interaction between these two proteins has an effect on cell survival in a clonogenic survival assay. APLF-depleted cells reconstituted with RNAi-resistant APLF\(^{WT}\), RNAi-resistant APLF\(^{W189A}\) or RNAi-resistant NLS-APLF\(^{W189A}\) can be treated with different DNA-damaging agents, such as IR, camptothecin (CPT) and methyl methanesulfonate (MMS) that have previously been shown to sensitize APLF-depleted cells [55]. If the APLF-Ku interaction is important for cell survival, then RNAi-resistant NLS-APLF\(^{W189A}\), which is localized to the nucleus but cannot interact with Ku, will be expected to demonstrate increased sensitivity to these DNA damaging agents. If an interaction with Ku is not important for cell survival, then RNAi-resistant NLS-APLF\(^{W189A}\) will likely not display sensitivity to the DNA-damaging agents.

The APLF-Ku interaction may be further regulated by PARP1. While PARP1 has been firmly implicated in base excision and single strand break repair, evidence of a role for PARP1 in DSB repair - in both HR and NHEJ - has recently emerged [62, 70]. In HR, PARP1 is proposed to bind to DNA ends in direct competition with Ku and suppress the effects of Ku-mediated NHEJ on HR [62]. Alternatively, when core components of the NHEJ pathway are absent, particularly in the absence of Ku, it has been proposed that PARP1 is recruited for DSB repair in B-NHEJ [70]. APLF may be important in regulating the interactions and activities of Ku and PARP1 since both of these proteins not only interact with APLF, but with each other as well [101-103]. It will be important to examine whether the APLF-Ku and APLF-PARP1 complexes are mutually
exclusive, or whether a multi-protein complex of APLF-Ku-PARP1 can form in the cell. This can be assessed in biochemical studies, such as gel filtration, co-immunoprecipitation and pull-down assays. Electrophoretic mobility shift assays (EMSA) can also be used to examine whether APLF affects the competition between Ku and PARP1 for DNA binding. We have previously shown that APLF does not have intrinsic double-stranded DNA-binding abilities, but can complex with Ku-bound DNA [56]. Moreover, as previously mentioned, PARP1 activity varies depending on the type of DNA ends [70, 98]; hence, it will be important to use PARP1-specific types of DNA substrates (blunt ends, 5’ or 3’ single-base overhangs and 3’ phosphoglycolate) in this experiment. Increasing amounts of Ku or PARP1 can be incubated with the radiolabelled DNA substrates, followed by the addition of PARP1 or Ku, respectively, alone or in combination with APLF. If the presence of APLF allows or enhances the ability of PARP1 to bind to DNA ends with a higher affinity than Ku, then it is likely that the APLF-PARP1 interaction out-competes Ku for DNA binding. Conversely, if Ku, in the presence of APLF, binds to DNA ends more robustly than PARP1, then the APLF-Ku interaction may out-compete PARP-1 for binding of DNA ends. If the absence of APLF does not affect the DNA-binding activities of either Ku or PARP1, then APLF may not facilitate the binding of these proteins to DNA ends.

It has previously been shown that depletion of APLF from human cells subjected to IR is associated with impaired DSB repair kinetics [56, 57]. This defect in DSB repair is likely associated with a defect in NHEJ since APLF interacts with core NHEJ proteins, however, it is also conceivable that this DSB repair defect is associated with the other mammalian DSB repair pathway, HR, which would implicate a role of
APLF in HR. To investigate this, we can use a recombination GFP reporter assay with APLF-depleted cells to monitor gene conversion events, which are carried out by HR in response to DSBs. As previously described [100], when DSBs are introduced into the GFP reporter, thus inactivating the gene, only a HR repair event will restore functional GFP expression. If APLF is required for HR, then APLF-depleted cells will be expected to exhibit reduced levels of HR repair, while reconstitution of these cells with RNAi-resistant APLF may correct this repair defect. Another functional assay to examine the role of APLF in HR would be to compare the clonogenic survival amongst wild-type, NHEJ-defective and HR-defective cell lines that are treated with either non-targeting (NT) or APLF-specific siRNA, upon exposure to different DNA damaging agents. In theory, if one DSB repair pathway is non-functional, then the additional depletion of repair factors in the same pathway would likely not result in further defects. However, if a component in another DSB repair pathway were compromised, then an additive DSB repair defect resulting in a more severe phenotype would be observed. If APLF is strictly involved in NHEJ, then its depletion in NHEJ-defective cells should not differ from those treated with NT siRNA. However, if APLF is also involved in HR, then depletion of APLF in NHEJ-defective cells may lead to increased hypersensitivity.

4.8 Summary

The work presented in this thesis demonstrates for the first time that the Ku-binding domain on APLF is mapped to amino acids 180-200, and that the interaction between APLF and Ku is critical for the nuclear localization of APLF. We suggest a
model whereby APLF localizes to the nucleus in a Ku-dependent manner, and once in the nucleus, the N-terminal portion of APLF is required for facilitating interactions with the core NHEJ proteins Ku and XRCC4-DNA ligase IV to promote efficient APLF-dependent NHEJ.
Figure 4.1: Proposed model for APLF nuclear localization and recruitment to sites of DSBs. Once APLF localizes to the nucleus in a Ku-dependent manner, APLF may be recruited to sites of DSBs either in a PARP-1- and PAR-dependent, Ku-dependent or/and XRCC4-dependent manner. Once at the DSB site, APLF may stabilize interactions between Ku and XRCC4-DNA ligase IV, thus promoting efficient end-joining in NHEJ.
REFERENCES


