

---

Electronic Thesis and Dissertation Repository

---

7-25-2017 12:00 AM

## Structural and Functional Characterization of Non-Homologous End Joining Factors

Huasheng Wang  
*The University of Western Ontario*

Supervisor  
Dr Murray Junop  
*The University of Western Ontario*

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science  
© Huasheng Wang 2017

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Biochemistry Commons](#)

---

### Recommended Citation

Wang, Huasheng, "Structural and Functional Characterization of Non-Homologous End Joining Factors" (2017). *Electronic Thesis and Dissertation Repository*. 4887.  
<https://ir.lib.uwo.ca/etd/4887>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

DNA double strand breaks represent the most toxic form of DNA damage. In mammals, non-homologous end-joining (NHEJ) is the primary DNA repair pathway for such damage, preventing both carcinogenesis and accelerated aging. Structural understanding of this repair pathway has received considerable attention, but has been significantly limited by the inability to obtain structures of higher order nucleoprotein complexes. A main obstacle in this respect has been difficulty in obtaining highly purified proteins, sufficient for structural determination. Improved protein expression and purification methods developed in this thesis permitted several NHEJ complexes to be selected for structural studies. Among these, Ku70-DNA and Ku70-DNA-PAXX yielded promising preliminary results. In depth optimization for crystal growth was performed and resulted in a full-length PAXX homodimer structure as well as low-resolution diffraction data for a novel Ku70-DNA complex. The PAXX structure confirmed prior suggestions that the C-terminal region of PAXX is highly disordered.

## Keywords

DNA double strand breaks, Ku70, Non-homologous end joining, PAXX, Protein expression, Protein purification, Protein structure, X-ray crystallography

## Acknowledgments

I owe my utmost gratitude to Dr Murray Junop for granting me the opportunity to explore one of the most vital mechanism for genome integrity. Dr Junop has been a wonderful supervisor, an irreplaceable mentor, provided me with insights into research and aspects of life, all the while tolerating my peculiar sense of humour, and without whom this thesis simply would not be possible. My committee advisors Dr Brian Shilton and Dr Hong Ling's expert opinions, and for offering me much needed guidance to shape my research direction. The kindness and refreshing advices from Drs David Litchfield, Derek McLachlin and Caroline Schild-Poulter were very much appreciated as they accompanied my learning experiences. I thank Dr Alba Guarné and Jeremy Caron for inspiring the crystallographer in me, Dr Mac Mok and soon-to-be Drs Robert Szabla for software assistance in data processing, Chris Brown and Sam Chu for technical assistance in experiments, and the rest of the Junop lab for the brainstorming and cheerful banters, and all aforementioned profusely for allowing me to tap into your wisdom and expertise.

I would also like to thank my family; Western's varsity fencing team, Carol, Brad, and teammates, whose love and support brought me constant joy and strength; teammates and classmates from McMaster University, Bayview Secondary and Walter Murray Collegiate, always there for me through the high and lows. M.Sc. has been a challenging journey, but among the stress and struggles rise a realistic sense of self-efficacy, and skills that are only earned after enormous efforts.

# Table of Contents

Abstract .....	ii
Acknowledgments.....	iii
Table of Contents .....	iv
List of Tables .....	vii
List of Figures .....	viii
List of Appendices .....	x
Abbreviations .....	xi
Chapter 1 .....	1
1 Introduction .....	1
1.1 DNA Double Strand Breaks.....	1
1.2 DNA Double Strand Break Repair Pathways .....	4
1.2.1 Homologous Recombination .....	6
1.2.2 Classical Non-Homologous End Joining.....	6
1.2.3 Alternative Non-Homologous End Joining .....	11
1.3 NHEJ Core Factors .....	11
1.3.1 DNA Ligase IV .....	11
1.3.2 XRCC4.....	14
1.3.3 XLF .....	17
1.3.4 Ku70/80.....	18
1.3.5 PAXX.....	24
1.4 Thesis Objectives .....	25
Chapter 2.....	26
2 Methods.....	26
2.1 Plasmid Constructs.....	26

2.2 Protein Expression and Purification.....	27
2.2.1 DNA Ligase IV/XRCC4 (LX4).....	27
2.2.2 XLF.....	28
2.2.3 Ku70/80.....	28
2.2.4 Ku70.....	29
2.2.5 PAXX.....	30
2.3 Protein Quantification.....	31
2.4 SDS-PAGE.....	31
2.5 Western Blot.....	31
2.6 SEC-MALS Analysis.....	32
2.7 DNA Ligation Assay.....	32
2.8 DNA Substrates for Crystallography.....	33
2.9 Electrophoretic Mobility Shift Assay.....	34
2.10 Crystallography.....	34
2.10.1 Crystallization of LX4, LX4-DNA, LX4-XLF-DNA.....	35
2.10.2 Crystallization of Ku70, Ku70-DNA, Ku70-DNA-PAXX.....	35
2.10.3 Optimization of Ku70-DNA and Ku70-DNA-PAXX Crystals.....	36
2.11 X-ray Diffraction Data Collection.....	36
2.12 Structural Determination and Refinement.....	37
Chapter 3.....	38
3 Results.....	38
3.1 Improvements in Production of NHEJ Factors.....	38
3.1.1 Purification of LX4, XLF and Ku70/80.....	38
3.1.2 Purification of Ku70 and PAXX.....	45
3.2 Purified LX4 Retains Ligation Activity.....	49
3.3 Ku70 Forms a Stable Dimer in Solution.....	51

3.4 Ku70 Homodimer Exhibits DNA Binding Activity .....	53
3.5 Crystallization Screening of LX4, LX4-DNA and LX4-XLF-DNA .....	55
3.6 Crystallization of Ku70 and Ku70-DNA .....	58
3.7 Optimization of Ku70-DNA Crystals Improved Diffraction Data Quality .....	63
3.8 Crystallization of Ku70-DNA-PAXX Generated PAXX Crystals .....	67
3.9 PAXX Structure Determined from Full-length PAXX.....	71
Chapter 4.....	73
4 Discussion .....	73
4.1 Overcoming Practical Limits to Studying NHEJ Complex Structures .....	74
4.2 Crystallographic Strategies to Facilitate High Quality Diffraction Data.....	76
4.3 Methods to Complement Crystal Structure Determination .....	79
4.4 Implications for NHEJ Complex Interactions.....	80
4.5 Outstanding Questions for NHEJ in Cancer Treatment.....	82
References.....	83
Appendices.....	100
Curriculum Vitae .....	113

## List of Tables

Table 1. Comparison of prokaryotic and eukaryotic NHEJ homologs .....	9
Table 2. Summary of purification yield of NHEJ factors .....	48
Table 3. Components of promising Ku70-DNA crystallization conditions .....	62
Table 4. Optimization of MCSG I #95 for Ku70-loop16 .....	66
Table 5. Comparison of Ku70-DNA and Ku70/80-DNA crystal parameters .....	66
Table 6. Summary of PAXX Crystallization and X-ray diffraction .....	69
Table 7. Comparison of experimental and referenced PAXX crystal parameters .....	70

## List of Figures

Figure 1. Overview of the causes and response of DNA double strand breaks.....	3
Figure 2. Overview of DNA repair pathway preference.....	5
Figure 3. Overview of NHEJ general mechanism and complex assembly.....	10
Figure 4. Structure of DNA Ligase IV catalytic domain in complex with an Artemis peptide .....	13
Figure 5. Crystal structures of XLF and XRCC4-DNA Ligase IV <sup>BRCT</sup> .....	15
Figure 6. Filament assembly of XRCC4 and XLF .....	16
Figure 7. Domains of Ku70 and Ku80.....	22
Figure 8. Crystal structure of Ku70/80 .....	23
Figure 9. Representative purification of LX4 complex .....	40
Figure 10. Purification of XLF .....	42
Figure 11. Ku70/80 Autoinduction profile .....	43
Figure 12. Purification of Ku70/80.....	44
Figure 13. Purification of Ku70 .....	46
Figure 14. Purification of PAXX.....	47
Figure 15. DNA ligation assay showing LX4 activity.....	50
Figure 16. SEC-MALS data showing stable dimer of Ku70 in solution .....	52
Figure 17. Analysis of Ku70 DNA binding activity with loop16 DNA substrate.....	54
Figure 18. Unoptimized LX4 crystals.....	57



Figure 19. Typical Ku70 non-diffracting microcrystals .....	60
Figure 20. Initial hits of Ku70-DNA crystals .....	60
Figure 21. DNA crystals obtained from Ku70-DNA complex crystallization .....	61
Figure 22. Diffraction patterns of the initial hits of Ku70-DNA .....	61
Figure 23. Progression of Ku70-loop16 crystal morphology through iterations of optimization .....	65
Figure 24. Progression of Ku70-DNA crystal X-ray diffraction data quality .....	65
Figure 25. Initial hit of Ku70-DNA-PAXX complex .....	68
Figure 26. Images of PAXX crystals in two different crystallization conditions .....	68
Figure 27. Single image of X-ray diffraction pattern from PAXX data collection .....	70
Figure 28. Highlighted Structural Information of PAXX.....	72
Figure 29. Idealized states of solutes .....	78

## List of Appendices

Appendix 1. List of oligonucleotides.....	100
Appendix 2. Structure of annealed DNA substrates .....	102
Appendix 3. Comparison of XLF expression and purification using different methods.....	106
Appendix 4. Ku70-DNA bound in solution.....	107
Appendix 5. List of crystallization conditions for Ku70 and X-ray screening outcomes.....	108
Appendix 6. List of crystallization conditions for Ku70-DNA and X-ray screening outcomes .....	109
Appendix 7. Optimization of Wizard I #10 and MCSG II #80 for Ku70-loop16 crystal growth and X-ray diffraction .....	111
Appendix 8. Ku70-DNA, Ku70-DNA-PAXX crystallization conditions .....	112

## Abbreviations

**5'dRP/AP site** 5' deoxyribosephosphate / apurinic and apyrimidinic site

**53BP1** p53 binding protein 1

**a-NHEJ** alternative Non-Homologous End Joining

**AP** apurinic/apyrimidinic

**APS** Advanced Photon Source

**ATM** ataxia telangiectasia mutated

**ATR** ataxia telangiectasia and Rad3 related

**Bax** B-cell lymphoma 2 associated X

**bp** base pair

**BRCT** breast cancer 1 C Terminus

**c-NHEJ** classical Non-Homologous End Joining

**CDK** cyclin dependent kinase

**CHK1** checkpoint kinase 1

**CHK2** checkpoint kinase 2

**Cryo-EM** cryo-electron microscopy

**CtIP** C-terminus binding protein interacting protein

**cv** column volume

**DNA-PKcs** DNA-dependent protein kinase catalytic subunit

**DSB** double strand break

**DTT** dithiothreitol

**EDTA** ethylenediaminetetraacetic acid

**EtBr** ethidium bromide

**H2AX** histone H2A, member X

**HEPES** 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid

**IPTG** isopropyl  $\beta$  -D-1-thiogalactopyranoside

**Ku70/80** Ku70 / Ku80 heterodimer

**LB** Lysogenic Broth-Lennox

**LX4** DNA Ligase IV / XRCC4 complex

**MRN** Mre11, Rad50, and Nbs1 complex

**Ni column**  $\text{Ni}^{2+}$  charged immobilized metal affinity column

**NLS** nuclear localization signal

**NMR** nuclear magnetic resonance spectroscopy

**PARP1** Poly [ADP-ribose] polymerase 1

**PAXX** Paralog of XRCC4 and XLF

**PBS** phosphate buffered saline

**PEG** polyethylene glycol

**PIKK** phosphoinositide 3-kinase-related kinase

**PMSF** phenylmethanesulphonyl fluoride

**PVDF** polyvinylidene difluoride

**RAG** recombination activating gene

**RbCl** rubidium chloride

**RPA** replication protein A

**SAP domain** SAF-A/B, Acinus and PIAS domain

**SDS-PAGE** sodium dodecyl sulphate polyacrylamide gel electrophoresis

**TBS-T** Tris buffered saline with Tween-20

**TCEP** Tris(2-carboxyethyl)phosphine hydrochloride

**V(D)J** Variable, Diversity and Joining

**vWA** von Willebrand A

**XLF** XRCC4-like factor

**XRCC4** X-ray Repair Cross-Complementing protein 4

## Chapter 1

### 1 Introduction

Failure to properly repair DNA damage affects both cancer formation and treatment. In normal cells, the lack of timely DNA repair leads to genomic instability and is considered one of the most important hallmarks of cancer development. Conversely, the increased DNA damage resistance in cancer cells in response to chemotherapeutic agents often originates from upregulated DNA repair genes (1). Such altered physiology of cancer cells opens the potential for new strategies of cancer treatment, which typically utilize DNA damaging agents to exploit the elevated levels of replication in cancerous cells. Conventional radio- and chemotherapy causes indiscriminate DNA damage to all cells in the body, although preference to killing is directed towards cancerous cells due to their rapidly dividing nature. In tumours that have undergone relapse, the upregulated DNA repair capacity of cancer cells decreases the effectiveness of further DNA damaging chemotherapeutics. Continued treatment has a negative effect on patient well-being, and leads to heightened mutation rates and further resistance toward therapy. Thus, targeting the prominent DNA repair pathway in tumour cells provides an effective way of re-sensitizing tumours to DNA damaging radio- and chemotherapies (2). Understanding the mechanisms of DNA repair assists such effort by providing necessary insight for development of small molecules that modulate cellular responses to DNA damage.

This chapter delineates the current research in the field of DNA damage and repair pathways, with a focus towards the structures and mechanisms of DNA double strand breaks, repair, and interactions of repair factors.

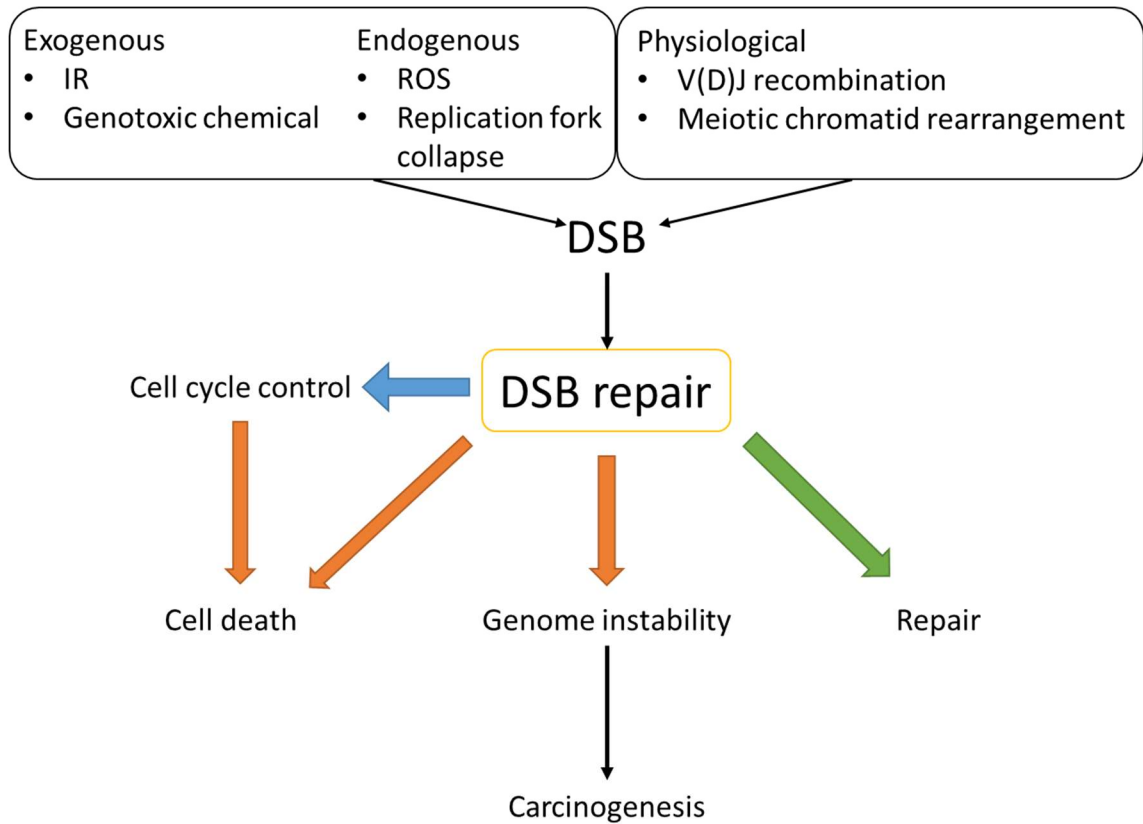
#### 1.1 DNA Double Strand Breaks

Genome stability constitutes a crucial aspect safeguarding human health and longevity. Our genome is constantly bombarded by DNA damaging events, including those resulting from ordinary cell processes (3). Ionizing radiation and other clastogenic agents, frequently used in cancer treatments, generate DNA damage (4). Of these, perhaps the best studied is ionizing radiation. Due to the heterogeneity in energy deposition, ionizing

radiation creates pockets of radicals in the aqueous environment of a cell, either through direct ionization of DNA molecules (5) or formation of hydroxyl radicals, that subsequently react with and damage DNA (6–8). This results in a complex variety of types of damage to the DNA, including single- and double-strand breaks, base and deoxyribose damage, and DNA-protein crosslinks (9, 10). In addition to exogenous sources of damage, cellular respiration also generates reactive oxygen species as a side product of the electron transport chain, and damages DNA in a similar manner (11, 12).

DNA double strand breaks occur when single-strand breaks are generated on complementary strands of a DNA helix within close distance (~10 base pairs, bp). In dividing primary mammalian fibroblasts, it is estimated that DSBs occur at a rate of about ten per day per cell (13, 14). Such damage can result in small local alterations to DNA sequence as well as larger chromosomal loss or rearrangements if the broken DNA ends fail to remain in close proximity for repair (Figure 1).

Due to the severity of such damage, even a single unrepaired DSB can lead to replication arrest and cell death (15). When a substantial number of cells experience senescence or apoptosis due to extensive DNA damage, it can lead to tissue atrophy resulting in eventual organ failure. DNA repair mechanisms are also involved in telomere maintenance, and failure in these two cell functions mirrors effects of accelerated aging (16–18). Indeed, the frequency of chromosome aberrations increases with age, suggesting that the rate of aging is related to DNA repair capacity (19). Equally undesirable, mis-repaired DSBs cause chromosomal mutations including sequence alterations, large scale deletions and inappropriate joining of DNA ends, resulting in further genomic instability through loss of tumour suppressor genes or activation of oncogenes. Ensuring an effective DNA repair system is a fundamental condition for human survival. Thus, efforts to understand DNA repair mechanisms are crucial for gaining insight to understand the basis for its failure, and ultimately formulate ways to modulate the repair process for improved cancer treatments.



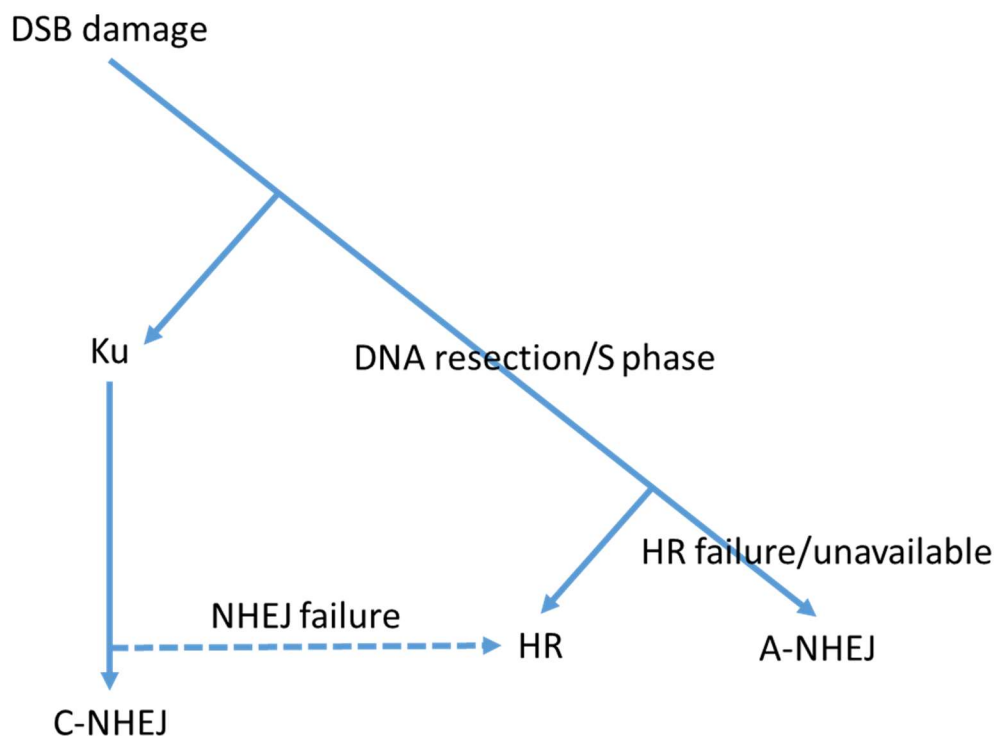
**Figure 1. Overview of the causes and response of DNA double strand breaks.** The blue arrow denotes parallel responses; orange arrows lead to outcomes when the repair fails; the green arrow leads to successful repair. Adapted from (18).

## 1.2 DNA Double Strand Break Repair Pathways

DNA repair provides the path for all organisms to maintain genome stability against a plethora of damaging agents. The two major DSB repair pathways in mammals are homologous recombination (HR) and non-homologous end joining (NHEJ). Within NHEJ, there is a classical NHEJ (c-NHEJ) pathway and alternative NHEJ (a-NHEJ) pathway also known as microhomology mediated end joining. Suppressed by c-NHEJ under normal circumstances, the a-NHEJ pathway acts as a backup when c-NHEJ fails and HR is unavailable or also fails (20). Both HR and c-NHEJ have additional roles in normal cell processes. HR is responsible for generating genetic diversity during meiosis of gamete cells (21) while NHEJ factors are employed in V(D)J recombination integral to the adaptive immune system through T and B cell maturation (22). The prioritization and choice of different pathways for repair is related to their efficiency and availability (Figure 2). Despite the desirable quality of HR, which is able to maintain sequence fidelity by using a sister chromatid or homologous chromosome as a template for repair, HR does not constitute the major repair pathway in mammals because it is not always available to the cell. As the only time during the cell cycle that a sister chromatid is available, or homologous chromosomes is in close proximity to serve as a template is during the S/G2 phase, HR presents a desirable outcome for repair at only these phases of the cell cycle (23, 24). During the remainder of the cell cycle, c-NHEJ is the most prominent pathway, owing to its rapid activation and ability to re-join numerous types of DSBs (25).

The activation of DNA repair is only one component of the DNA damage response. Following initiation of repair, a composite signalling cascade is engaged in the cell directing its fate. DNA damage “sensors”, phosphatidylinositol 3-kinase-related kinases (PIKK) including ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3 related (ATR), interact with DSB DNA ends and mediates the repair pathways by interacting with a Mre11, Rad50 and Nbs1 (MRN) complex or DNA-PKcs (26–28). They also control cell cycle checkpoints by signalling through Chk1 and Chk2 (29, 30). Upon DNA damage, the histone H2AX is phosphorylated on serine 139 by ATM, allowing the chromosome to become less condensed, thereby facilitating access by DNA repair





**Figure 2. Overview of DNA repair pathway preference.** Ku70/80, with its high abundance and affinity for DNA ends, is typically the first to interact with the DNA ends, initiating the c-NHEJ pathway. Failing this, HR elements binds ends and the HR pathway proceeds. If NHEJ and/or HR are unable to repair the DNA damage in a timely manner, a-NHEJ pathway may be used as a final means to avoid chromosome loss or translocation.

factors. ATM also phosphorylates tumour suppressors, such as p53, whose activation occurs when DNA damage is beyond repair (31, 32). Under such conditions, p53 acts as a transcription factor for expression of proapoptotic proteins including B-cell lymphoma 2 associated X protein (Bax), and activates the mitochondrial pathway for programmed cell death (33). The elaborate process of DNA damage response is a field of active research, and a full description can be found in recent reviews (32, 34, 35).

### 1.2.1 Homologous Recombination

HR conserves genetic information by utilizing an undamaged DNA template, preferentially the sister chromatid, for repair (36). Because of the need for a sister chromatid or homologous chromosome to be in physical proximity to a DSB for complementary base pairing with the broken DNA strand, HR occurs more readily during the S and G2 phases of the cell cycle when homologous templates are accessible. During repair by HR, an MRN complex binds to the DNA at the double strand break and tethers the two ends in close proximity. C-terminus binding protein interacting protein (CtIP) binds the MRN complex and initiates DNA resection of the 5' strands, leaving 3' single-stranded DNA tails protected by Replication Protein A (RPA). Further resection is carried out by downstream factors. Radiation Sensitive 51 (RAD51) displaces RPA and forms filaments on the single-stranded DNA, promoting homologous pairing and strand invasion (37). Following DNA synthesis, the two DSB ends align with their respective homologous strand forming a double Holliday junction, which is then resolved by resolvase enzymes. In meiotic DSB repair, crossover products are generated allowing genetic exchange. In mitotic DSB repair, the DSB site is joined in a non-crossover manner to preserve the template sequence (38).

### 1.2.2 Classical Non-Homologous End Joining

Despite HR providing potential for high-fidelity repair during late S phase and G2 phase, NHEJ is the prevalent pathway employed to repair DSB in mammalian cells. Although NHEJ is error prone, the ability to rapidly ensure chromosomes remain intact supersedes the negative consequences of local mutations at the breakage site which are frequently acquired during repair. With a lack of homology based, “proofing template” for repair,

the process of NHEJ relies critically on its fast kinetics to minimize the consequences of DSB damage and mend the ends through ligation (39–41).

Two types of NHEJ exist, with c-NHEJ being the dominant pathway due to its rapid ability to re-join DNA ends thereby protecting from large resection or involvement in DNA rearrangement. As a vital repair mechanism for DNA DSB repair in mammalian cells, c-NHEJ employs a small set of core factors with impressive versatility in recognizing and processing DNA ends for efficient end joining. Depending on the complexity of the damaged DNA ends, the rate of c-NHEJ can be three to six times higher than HR (25). The core factors involved in c-NHEJ have been identified. Numerous auxiliary factors, including nucleases and polymerases, are also recruited to the damage site as needed to facilitate the repair process (42). While the primary sequence of NHEJ repair factors are poorly conserved, functional homologs of the core factors can be found widely across evolutionary descendants (Table 1).

The c-NHEJ pathway consists of three main steps: synapsis, end processing, and end joining (48) (Figure 3). When a DSB event occurs, the Ku70/80 heterodimer binds to each of the two DSB ends within seconds (39). Ku70/80 is constitutively expressed, and is one of the most abundant proteins in the cell. Its ring-shaped structure allows it to thread onto the broken DNA end in a sequence independent manner (49). In doing so, Ku also excludes HR from occurring by sterically blocking DNA resection factors required for HR (50). Following end binding, Ku70/80 translocates inward on the DNA for about one helical turn distal from the free end, allowing for recruitment of DNA-dependent protein kinase catalytic subunit (DNA-PKcs). DNA-PKcs directly interacts with DNA and occupies a region of ~10 bp proximal to the free end (51). Together, Ku70/80 and DNA-PKcs make up the DNA-dependent protein kinase holoenzyme (DNA-PK) which is able to synapse the two broken DNA ends, preventing them from long-range movements that would make re-joining of ends very challenging (52, 53). Interestingly, DNA-PKcs has only been identified in eukaryotes (54–56). NHEJ in prokaryotes require only a homodimer of Ku in complex with LigD, a DNA Ligase IV homolog (57, 58). This implies that Ku70/80 may be sufficient for simple end joining, and that DNA-PKcs represents a more elaborate, evolutionarily recent mechanism to manage complex aspects

of DNA repair such as cell cycle regulation. Phosphorylation by DNA-PK provides the means by which many interacting factors are recruited and regulated during repair. The dissociation of DNA-PK from the damage site after repair completion is suggested to depend on the auto-phosphorylation of DNA-PKcs, further indicating its function in coordinating the repair complex (59).

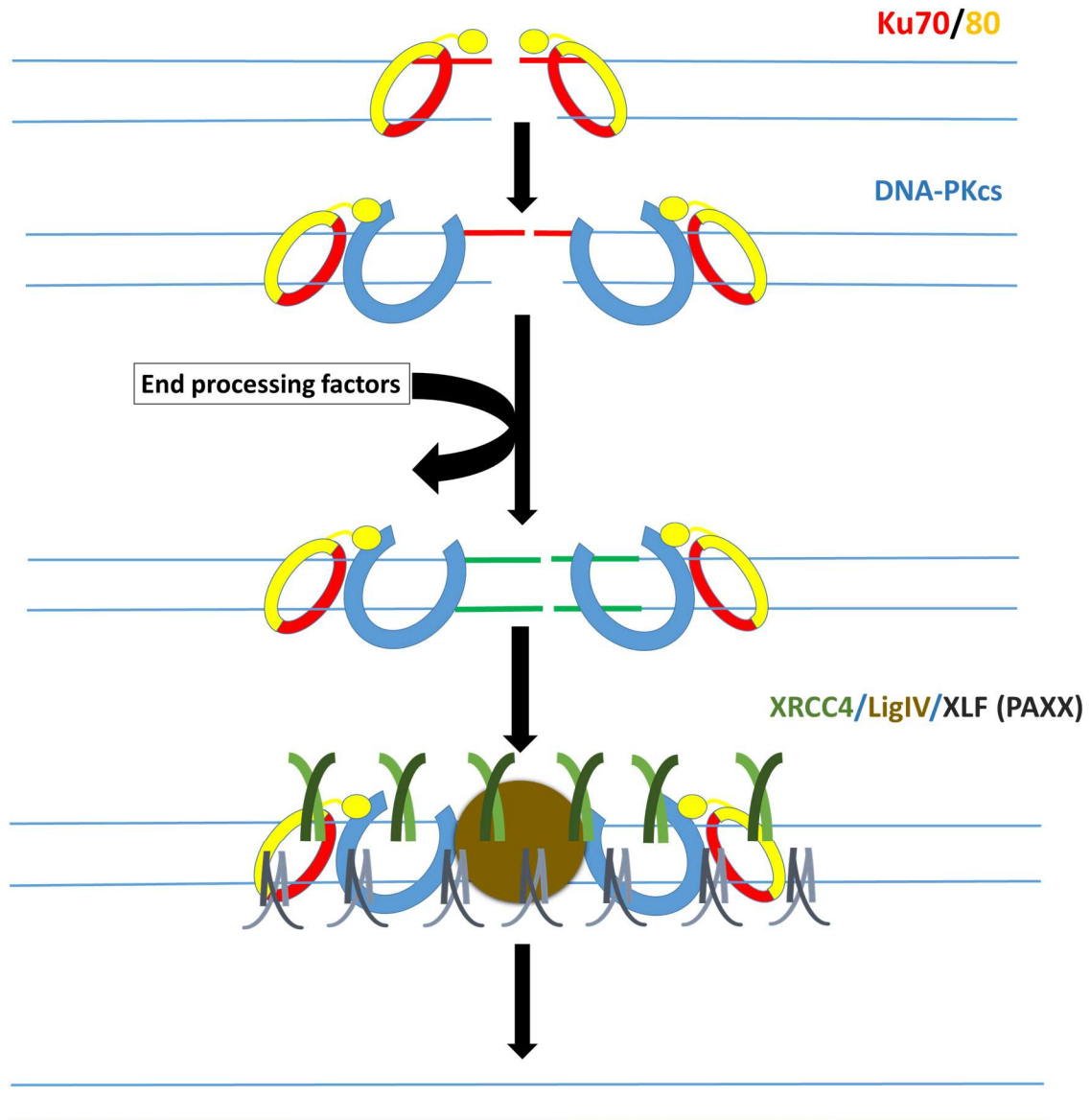
Due to the frequent presence of complex types of damage at a DSB, end processing represents an important part of DSB repair by NHEJ. The ability to accommodate diverse types of DNA ends underscores the multifunctionality and mechanistic flexibility of repair factors involved in end processing. In simple end joining, Ku70/80 has the capacity to coordinate ligation with downstream factors independent of DNA-PKcs (60). Cryo-electron microscopy (cryo-EM) and small angle X-ray scattering (SAXS) data have provided some insight into the assembly, auto-phosphorylation, and disassembly stages of DNA-PKcs activity (61–63). DNA-PK orchestrates an array of nucleases including Artemis, which is phosphorylated by DNA-PKcs, enabling its endonucleolytic activity (64, 65). Polymerases including pol $\mu$  and pol $\lambda$  are also recruited to adaptively modify DNA ends to make them chemically compatible for ligation (66, 67).

Finally, ligation of DNA is carried out by a ligase complex comprised of DNA Ligase IV, X-ray cross-complementation group 4 (XRCC4) and XRCC4-like factor/Cernunnos (XLF) (68–70). This ligation complex has been shown to be able to join DNA ends with blunt, microhomology, and non-compatible sequences. With coordinated positioning of DNA ends (termed DNA bridging) by filaments of XRCC4/XLF, DNA Ligase IV is stabilized at ends to form a covalent AMP-enzyme intermediate on a highly-conserved lysine residue (K273). This AMP is subsequently transferred to the 5' phosphate group of a DNA end. Following S<sub>N</sub>2 substitution by an activated 3' hydroxyl group from the opposing strand, with AMP as the leaving group, the sugar phosphate backbone of DNA is re-joined.

It is important to note that NHEJ does not proceed in an absolute step-wise fashion. Ku70/80 directly interacts with many factors including DNA-PKcs (71), XRCC4 (39), DNA Ligase IV (72), XLF (73), and a paralog of XRCC4 and XLF (PAXX) (74), and it is not yet clear how these various assemblies contribute at different stages of repair.

**Table 1. Comparison of prokaryotic and eukaryotic NHEJ homologs.** Core factors in NHEJ are functionally conserved.

	End recognition	End processing	Ligase
Prokaryote (43)	Ku (30-40kDa)	LigD	LigD
Eukaryote: Saccharomyces cerevisiae (44–46)	Ku70/80	pol4	Lif1 + Dnl4 + Nej1
Human (47)	Ku70/80	pol $\mu$ and pol $\lambda$ , Artemis	XRCC4 + DNA Ligase IV + XLF



**Figure 3. Overview of NHEJ general mechanism and complex assembly.** Ku70 (red) and Ku80 (yellow) form the heterodimer that detects an initial DSB break. Ku80 contains a C-terminal DNA-PKcs interaction domain that recruits DNA-PKcs (blue), forming a DNA-PK holoenzyme. The incompatible DNA ends (red) are modified by various end processing factors to make them compatible for ligation (green). XLF (grey), XRCC4 (olive) and DNA Ligase IV (brown) form a ligation complex at the DNA terminus in a filamentous arrangement that stabilize the ends for efficient ligation.

### 1.2.3 Alternative Non-Homologous End Joining

In the event that both c-NHEJ and HR fail, the a-NHEJ pathway may complete repair through a mechanism that requires end-to-end stabilization provided by the presence of micro-homology at opposing ends of the DSB. Given the serious nature of failure to repair a DSB, it is not surprising that cells have a backup repair mechanism to restore chromosome integrity and prevent chromosome loss (75, 76). Both c-NHEJ and HR factors have been suggested to participate in the a-NHEJ pathway (77, 78), thus recruitment of repair factors in a-NHEJ may be a function of their presence during the initial repair attempt. In cells with deficient Ku heterodimer, a-NHEJ is significantly elevated, which was initially termed Ku-independent pathway (20, 79). A-NHEJ is error-prone as it relies on only a few bases of homology at the broken DNA ends to promote stability. These sites of micro-homology typically form after small amounts of 5' strand resection. Repair by a-NHEJ occurs with slower kinetics and is implicated in chromosome translocation events contributing to cancer (75, 80). End joining in a-NHEJ is dependent on both Poly [ADP-ribose] polymerase 1 (PARP1) and DNA Ligase III (79). This pathway was only recently identified and has not been well studied. It will be interesting to see what redundancies, if any, exist between a-NHEJ and c-NHEJ.

## 1.3 NHEJ Core Factors

### 1.3.1 DNA Ligase IV

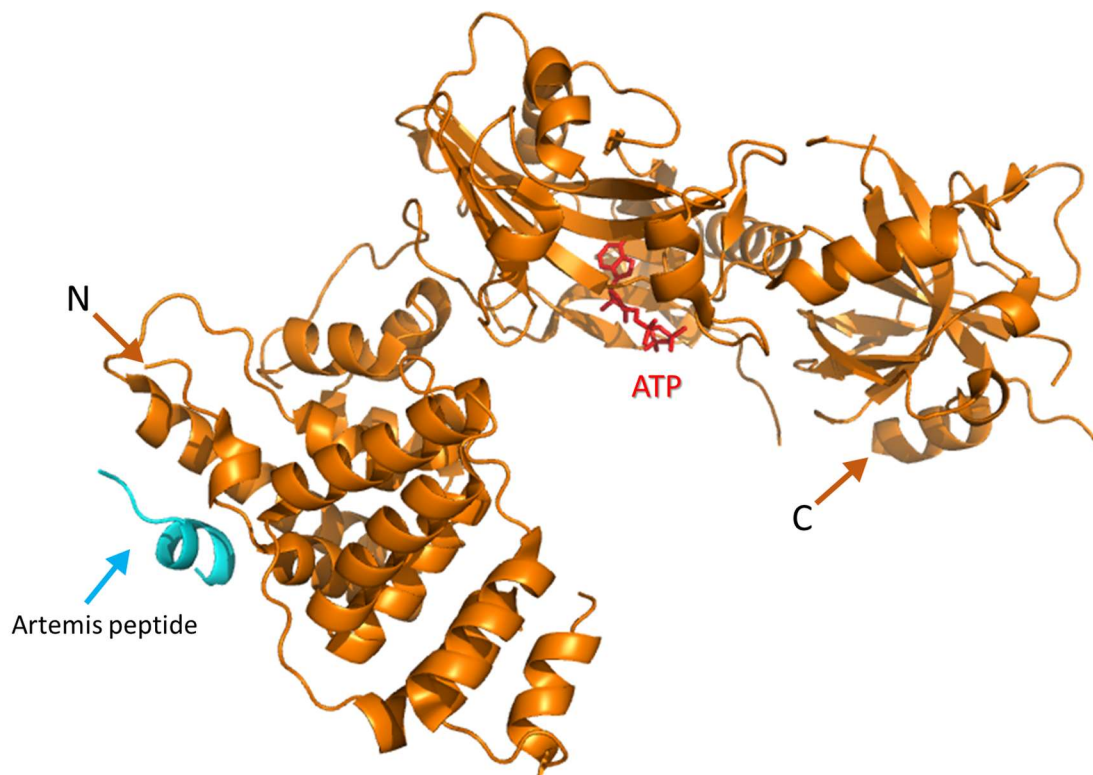
DNA Ligase IV has been identified as the crucial ligase component in joining ends at a DSB, and its function appears to be limited to c-NHEJ. Not surprisingly, a DNA Ligase IV knockout results in complete abolishment of c-NHEJ repair capacity, causing severe radiosensitivity and immunodeficiency in patients (81–84). In the absence of DNA Ligase IV, cells are forced to attempt repair through the less available HR or error-prone a-NHEJ. The error-prone repair of a-NHEJ may lead to chromosome translocation as utilizing micro-homology has potential to direct the damaged end to the end of other chromosomes. When DNA Ligase IV is utilized as part of the c-NHEJ pathway, translocation rates are shown to be suppressed (85).

On its own, DNA Ligase IV is very unstable and undetectable in cells. It is stabilized

through formation of a DNA Ligase IV-XRCC4 (LX4) complex, as shown in both mammal and yeast cells (86, 87). Deletion and/or mutation of genes for either XRCC4 or Ligase IV both result in similar phenotypes including severe combined immunodeficiency (SCID), hypersensitivity to DSB formation, neuronal apoptosis and, in the case of a complete gene deletion, embryonic lethality (88).

The structure of full length Ligase IV has not been determined, although the catalytic region of the protein (including the nucleotidyltransferase domain and an OB-fold domain) in complex with an Artemis polypeptide has been solved by X-ray crystallography (PDB 3W1G, Figure 4) (89, 90). The C-terminal domain of Ligase IV is required for interaction with XRCC4 and its recruitment to sites of DNA damage (91). The interaction of Ligase IV with XRCC4 maps to the two breast cancer associated 1 C Terminus (BRCT) domain at the C-terminus of Ligase IV, and the C-terminal tails (residues 160-200) of XRCC4 homodimer (89, 92). The binding results in a 1:2 stoichiometry and is therefore asymmetric in nature. Due to its similarity in domain organization with DNA Ligase I and III, its interaction with DNA is anticipated to occur in a similar manner. Ligase I and III have been shown to bind DNA by encircling DNA with the DNA binding domain, nucleotidyltransferase domain, and OB-fold domain (93, 94). However, this has yet to be validated through structural determination for Ligase IV. Current knowledge on DSB repair has focused on the first strand ligation event, and nothing is yet known about the interactions allowing a second end joining event required to fully complete repair. Current evidence suggests that a single complex of LX4 may seal both strands. This would require dramatic repositioning of the LX4 complex as well as recharging Ligase IV with ATP. XLF has been shown to promote re-adenylation of Ligase IV after the first joining event and may therefore enable the second strand to be joined (70). It is also possible that a symmetrical LX4 complex (2:4 ratio of Ligase IV to XRCC4) with opposing LX4 complexes on both strands, may perform ligation with each LX4 sealing just one strand. Interestingly, there is evidence to suggest that Ligase III, typically involved in single-strand break repair, may seal the second strand following LX4 sealing of the first strand (95). Improvement in the current understanding of the Ligase IV mechanism will require structures of Ligase IV complexes formed with DNA and other repair factors such as Ku70/80 and XLF.





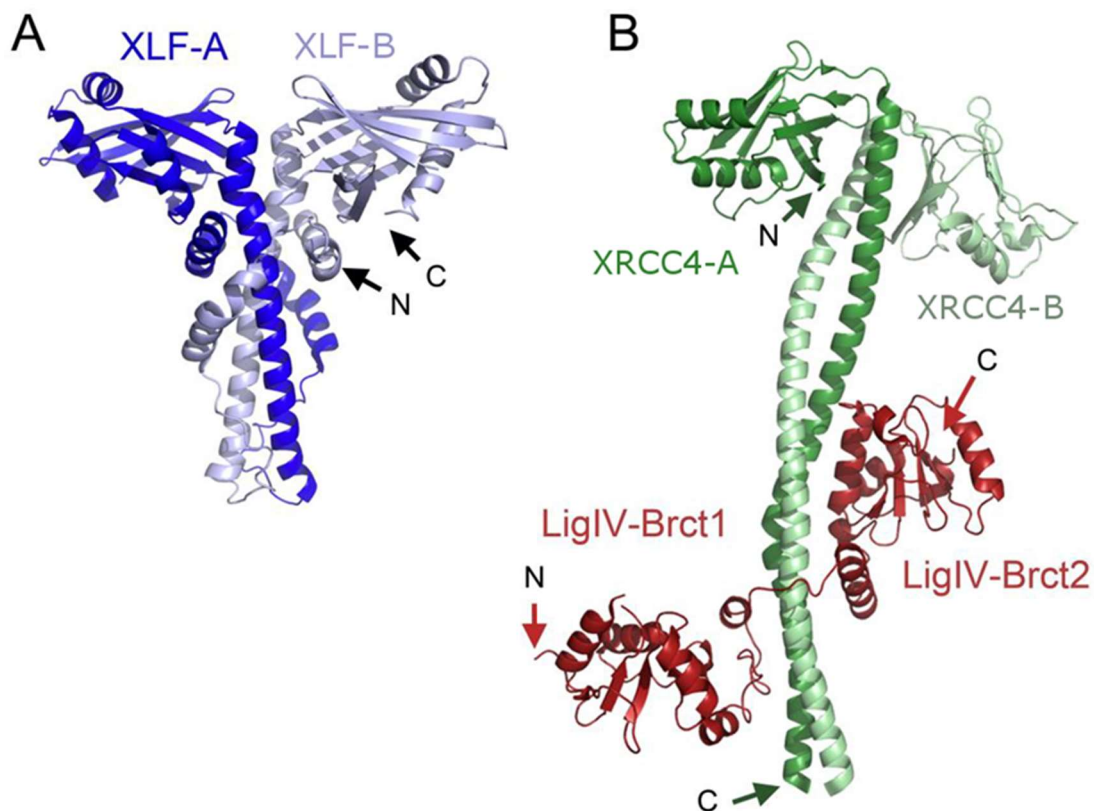
**Figure 4. Structure of DNA Ligase IV catalytic domain in complex with an Artemis peptide.** The DNA Ligase IV catalytic domain is shown in gold, and the Artemis peptide in cyan. The N- and C-terminus of the DNA Ligase IV catalytic domain, the Artemis peptide, and ATP are indicated by labelled arrows of corresponding colour. The ATP molecule (red) can be seen interacting within the catalytic domain at the location of the catalytic lysine K273. This structure shows an open conformation in which DNA is not bound. The structure is rendered from PDB 3W1G using PyMOL.

### 1.3.2 XRCC4

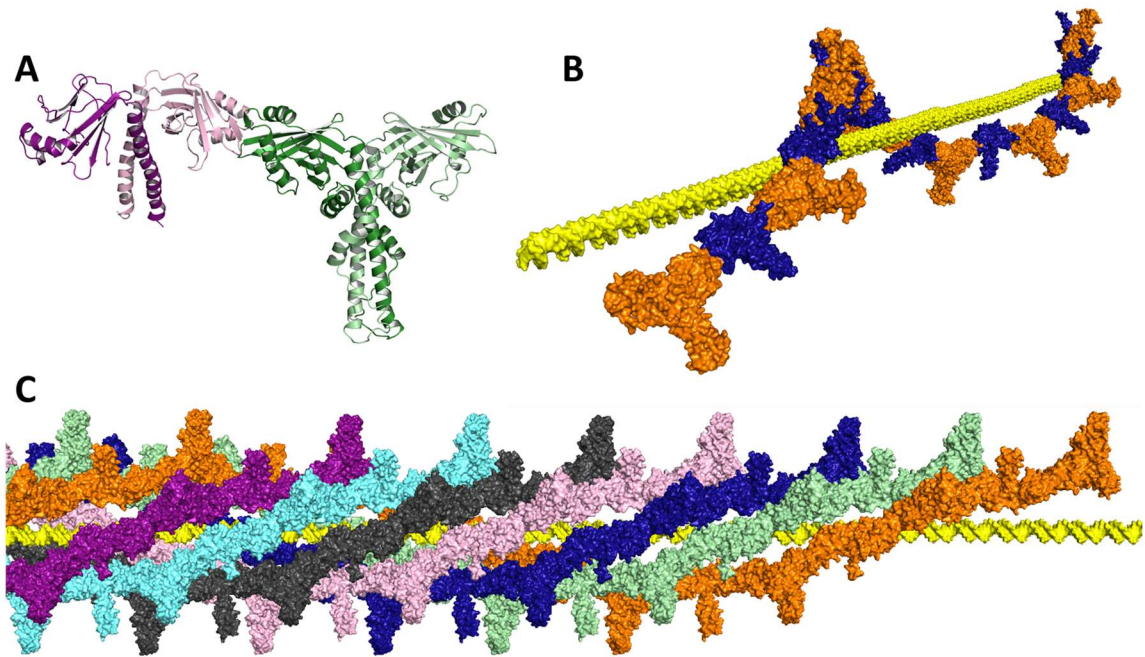
DNA Ligase IV stability absolutely requires XRCC4 binding, and the crystal structure of XRCC4 (residues 1-203) in complex with the tandem BRCT repeat region of Ligase IV has been determined (PDB 3II6, Figure 5B). Unfortunately, this complex lacked DNA and the catalytic core of Ligase IV (spanning the first ~600 residues). Nevertheless, insight into LX4 complex formation was obtained and new co-expression approaches were developed for production of the truncated complex. Utilizing a bacterial co-expression system for truncated XRCC4 and Ligase IV was essential to achieve sufficient material for structural studies (86). Based on this success, a similar approach may prove equally useful for structural studies of the full-length complex.

Since the reliance of DNA Ligase IV on XRCC4 for stabilization is not mutual, and there is an excess of XRCC4 in the cell relative to Ligase IV, it is thought that excess XRCC4 is used to form filaments with XLF. Long filaments of XRCC4-XLF have been observed in cells at DSB sites and there is evidence to suggest that Ligase IV is incorporated throughout the entire length of the filament. XRCC4-XLF filaments move along the DNA damaged ends, not only stably bridging ends but also positioning Ligase IV at DNA termini for successful ligation (96, 97).

Despite their structural similarity, XRCC4 and XLF carry out distinct functions during NHEJ repair. Functioning as a homodimer, XRCC4 associates with and is recruited to damage sites by a homodimer of XLF. This interaction can occur without coordination of DNA-PKcs (98), and stimulates the ligation and adenylation activity of Ligase IV by stabilizing the complex at the DSB site (99, 100). It is currently unclear how the repair complex manages to coordinate DNA Ligase IV to feed broken ends into its active site, but a model for XRCC4-XLF interaction suggests a sleeve of multiple XRCC4-XLF filaments bridges DNA ends and can accommodate necessary structural changes to permit access of Ligase IV to DNA ends. Filament structures of XRCC4 and XLF have been determined by X-ray crystallography (Figure 6, PDB 3SR2).



**Figure 5. Crystal structures of XLF and XRCC4-DNA Ligase IV<sup>BRCT</sup>.** (A) XLF (PDB 2R9A) and (B) XRCC4-DNA Ligase IV<sup>BRCT</sup> (PDB 3II6). XLF monomers (residues 1-224) within the homodimer are shown in different shades of blue. XRCC4 monomers (residues 1-203) are in different shades of green and the tandem BRCT domain of DNA Ligase IV are shown in red. N- and C-termini are indicated. The figure was generated using PyMOL.



**Figure 6. Filament assembly of XRCC4 and XLF.** (A) shows the N-terminal head-to-head interaction between XRCC4 (purple) and XLF (green). (B) illustrates a single filament formation along DNA (yellow). (C) illustrates multiple filament assembled into a bundle, each filament in a different colour.

### 1.3.3 XLF

Relative to Ku70/80, DNA-PKcs, and LX4, XLF was only recently identified as a necessary component for c-NHEJ and V(D)J recombination (101). XLF and its yeast homolog Nej1 are required for successful completion of ligation, although the precise mechanism of action remains unclear (69, 102–104). Patients with XLF-disrupting mutations suffer an inability to complete DNA ligation at similar levels compared to knockouts of Ligase IV and/or XRCC4, suggesting XLF interaction with these components is necessary for successful end joining (69). XLF and XRCC4 have different functions despite their structural homology. Functional studies show while XLF interaction is dispensable in joining of compatible ends, the efficiency of joining mismatched and blunt ends is largely dependent on the presence of XLF (104). Additionally, the DNA damage response protein ATM and histone remodelling protein H2AX share functional redundancy with XLF but not XRCC4, underscoring distinct roles for XLF and XRCC4 (105).

Although both XLF and XRCC4 form stable homodimers in solution, their ability to interact does not occur through heterodimer formation (106, 107). The crystal structure of the N-terminal domain of XLF (PDB 2R9A) illustrates a crucial difference between XLF and XRCC4. In XLF, the helical region that extends from the dimerization domain bends back towards the N-terminal globular region, instead of continuing to extend away as a helical coiled coil as shown in XRCC4 (Figure 5). As such, XLF exhibits a more compact structure and precludes Ligase IV from binding in the same manner as XRCC4. Currently there is no structure of the C-terminal region of either XLF or XRCC4. Although deletion of this region in XRCC4 appears to be dispensable for repair, similar deletions in XLF result in loss of DNA binding and ability to stimulate ligation of non-cohesive DNA ends (104). In addition, deletion of the last 10 amino acids in the C-terminus of XLF (XLF<sup>1-289</sup>) abrogates Ku70/80 interaction and results in accumulation of DSBs. This finding underscores the essential role of the flexible C-terminus of XLF for mediating DNA and protein-protein interaction (108). Interestingly, Ku70/80-XLF interaction appears to be mediated by the presence of DNA, making it clear that further structural knowledge of

the Ku70/80-XLF protein complex will be required to understand of the intricate interactions required for successful DSB repair by c-NHEJ (108).

Interaction of XLF-XRCC4 involves association of two homodimers and has been localized to their respective N-terminal head regions (XLF<sup>1-128</sup> and XRCC4<sup>1-119</sup>). The resulting interaction permits further formation of an extended filament composed of alternating XRCC4 and XLF homodimers. Single filaments further assemble into a multi-filament bundle (or sleeve) that is required for stable bridging of DNA ends (68, 109–111). Formation of multi-filament bundles is dependent on the extended tails of XLF. At this time, it is not yet clear how the tails govern sleeve assemble or interaction with DNA.

For joining of DSBs *in vitro*, a ligation complex consisting of DNA Ligase IV, XRCC4 and XLF is sufficient. In cells this is not the case as Ku70/80 is required (69, 112). At least part of this requirement is related to the ability of Ku70/80 to recruit XLF, and by extension LX4, at the DSB site. At this time, it would appear that Ku70/80 may play a more important role for DNA damage recognition and response, coordinating end processing factors and repair complex regulations, rather than directly participating in end joining.

### 1.3.4 Ku70/80

Ku70 and Ku80 were first discovered in 1981 as autoantigens in the serum of a scleromyositis patient (113). Ku70/80 heterodimer is an important element of maintaining genome stability. Its capacity to bind DNA ends is instrumental for its role in both DNA repair and telomere maintenance. As one of the most abundant proteins in mammalian cells, its expression is estimated at  $\sim 4 \times 10^5$  molecules per cell (48). As the initial DNA damage sensor, Ku70/80 heterodimer exhibits fast kinetics in binding and recruitment of the DNA dependent protein kinase catalytic subunit (DNA-PKcs) to form DNA-PK holoenzyme. This kinase complex is responsible for orchestrating the regulation of most if not all NHEJ core repair factors. Since c-NHEJ is the major DNA repair pathway, and DNA-PK is vital for c-NHEJ, it has become a primary target for cancer treatments aiming to knock down upregulated DNA repair in tumour cells (2).

When multiple types of damage occurred at a DSB, a multifunctional repair factor can provide a means for efficient repair without necessitating coordination of multiple repair pathways. In this regard, Ku70/80 contains 5' deoxyribosephosphate and apurinic/apyrimidinic (5'dRP/AP) lyase activity, preferentially processing AP sites within 1-2 nucleotides of a DSB (114). The end processing capacity of Ku70/80 resembles base repair activity observed in homodimeric Ku from prokaryotes. Prokaryotic Ku, with its ligation partner LigD, demonstrates not only the expected DNA end binding and joining activity (58, 115–117), but also nucleolytic activities that process ends for efficient ligation (118).

In addition to a role in NHEJ, Ku70/80 also participates in telomere maintenance as a part of the telomere binding complex, and is required for the perinuclear localization of the telomeres (58, 119, 120), where it paradoxically prevents end joining of chromosomes (121, 122). Since Ku70/80 binds DNA ends in a sequence independent manner, the choice of joining or preventing joining of ends is dependent upon its interacting partners. Studies in yeast and mammalian cells have shown that shelterin complex is crucial for modulating Ku70/80 function at telomeres. Telomere dysregulation by knockouts of shelterin proteins leads to chromosome-chromosome ligation, indicating that Ku70/80 lacks the ability to distinguish different types of DNA ends on its own (122, 123). The exact regulatory mechanism of Ku70/80 function at telomeres remains an area of active research (120, 121, 124, 125).

Ku70 and Ku80 share structural similarities in their core regions (Figure 7 and 8). Unlike the Ku homodimer from prokaryotes, Ku70/80 in eukaryotes contains a von Willebrand A (vWA) domain. The vWA domain serves as a regulator site and has been implicated in interaction with auxiliary factors. Prokaryotic Ku is a homodimer and only contains the core region, which resembles that of both Ku70 and Ku80 (126). Due to significant sequence homology with the Gam protein of bacteriophage Mu, prokaryotic Ku is proposed to have been acquired via lateral gene transfer. Similar to Ku from both prokaryotes and eukaryotes, Gam binds DNA ends as a dimer and protects DNA ends from nuclease digestion (127). Although both Ku70 and Ku80 possess additional domains compared to more primitive organisms, Ku80 may be particularly evolved for

function in NHEJ in eukaryotic organisms. This is likely as Ku80 is only found in eukaryotes (128, 129). Furthermore, the mitochondrial Ku80 homolog harbours a C-terminal truncation that makes its size (68 kDa) similar to Ku70 (130). The C-terminal region of Ku80 contains the DNA-PKcs interacting domain, suggesting the added complexity of concerted DNA end-modification is an evolutionarily recent adaptation. On the other hand, the C-terminus of Ku70 contains a SAP domain (a 35-residue motif, named after the three proteins containing it: SAF-A/B, Acinus and PIAS) that has been suggested to mediate at least some interaction with DNA. Indeed, similar SAP domains from unrelated proteins have a conserved DNA binding activity (58). The SAP domain of Ku70 is expected to play some role in modulating complex assembly on DNA ends (131), but the actual function remains unclear.

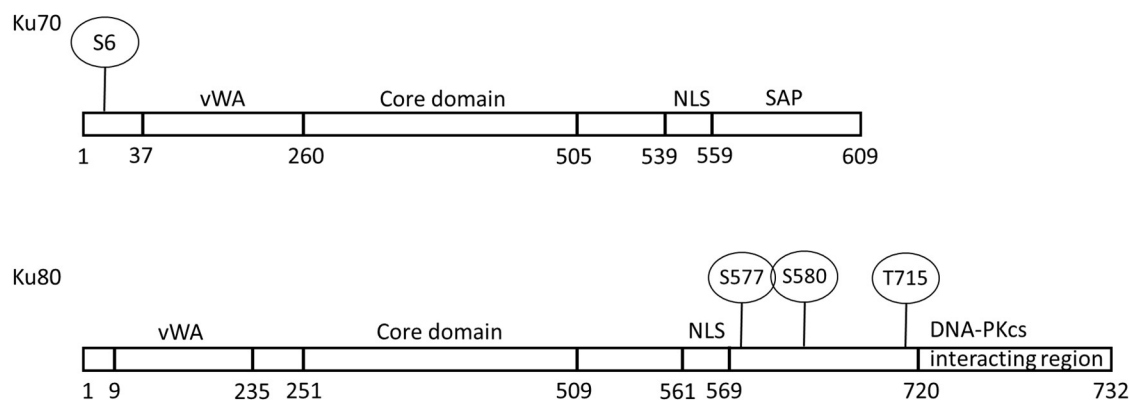
Although Ku70 and Ku80 are structural homologs, they do not appear to carry out redundant functions. Interestingly, investigation of the difference between Ku70/80 subunits suggests that at least one difference in function relates to DNA loading. Ku70/80 threads onto DNA ends with the Ku80 side first. Although the significance of this preference is not understood, it has been suggested that heterodimeric Ku (as opposed to homodimer Ku in prokaryotes) may provide a means for eukaryotic Ku70/80 to be regulated for different roles in NHEJ, telomere maintenance and other activities (132).

The function of Ku70 and Ku80 has been explored through several mouse-model knockout studies. Deletion of either Ku70 or Ku80 result in mice with SCID due to defective ligation of V(D)J recombination (133, 134),  $\gamma$ -radiation hypersensitivity (133, 135), smaller size (136), and shortened life span due to early aging (137–139). Given this shared phenotype and the intertwined nature of the Ku70/80 structure, the field has traditionally understood that Ku70 and Ku80 are an obligate heterodimer. The first evidence that this may not be the case came from knockout studies of individual Ku subunits that demonstrated depletion of a single subunit does not diminish expression of the other entirely, indicating that heterodimer formation is not obligatory for stabilizing Ku70 or Ku80 (133, 136). Furthermore, with deficient levels of p53, Ku80<sup>-/-</sup> mice demonstrated increased incidence of cancer, with nearly 100% of mice dying from pro-B-cell lymphoma (135), while p53<sup>-/-</sup> Ku70<sup>-/-</sup> mice exhibited longer life span with less

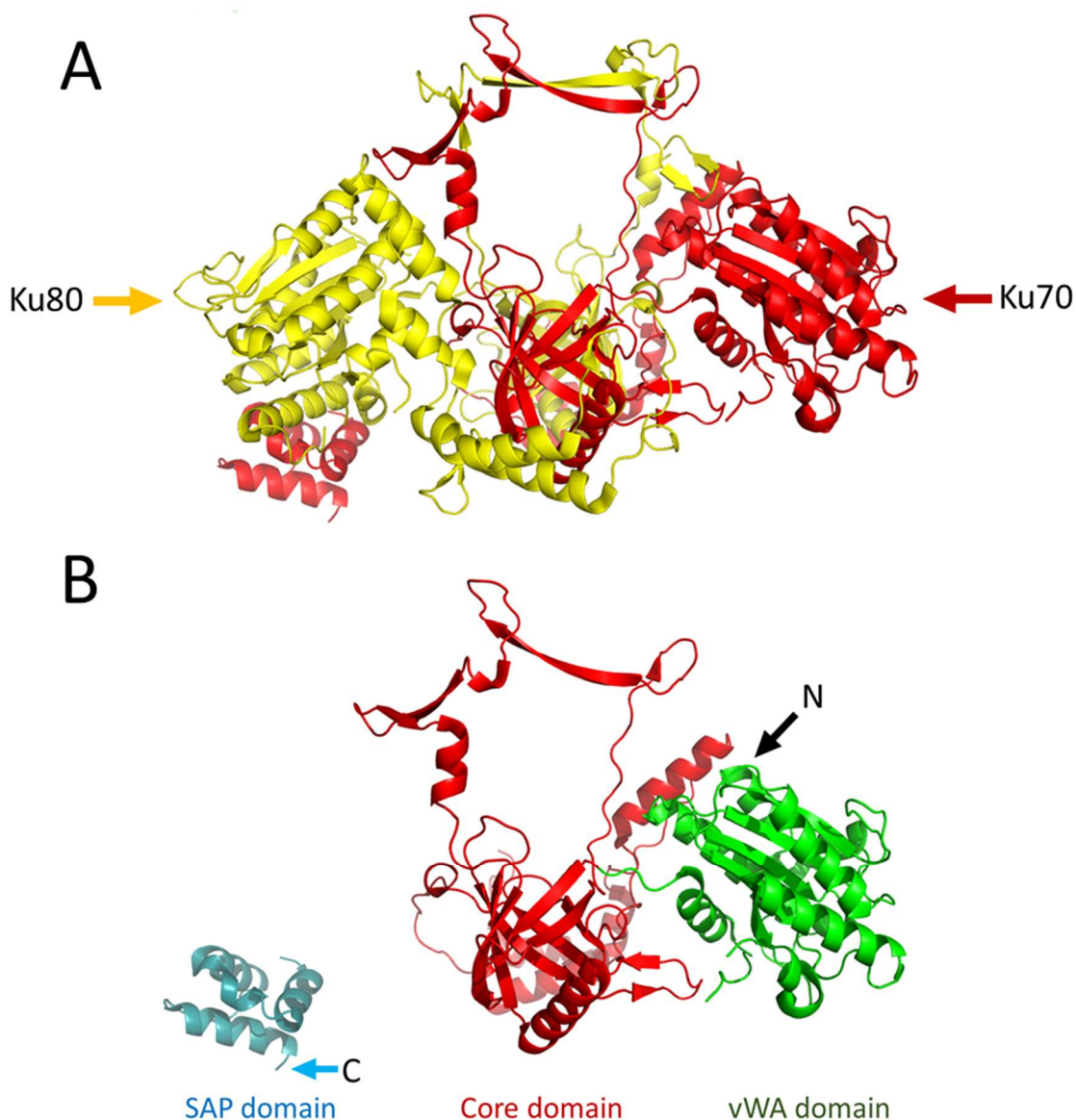


incidence of pro-B-cell lymphoma, with mortality largely due to medulloblastoma and T-cell lymphoma (140). A p53 knockout was important for separation of phenotypes between Ku80<sup>-/-</sup> and Ku70<sup>-/-</sup>, as p53-mediated responses led to premature replicative senescence and apoptosis, masking functional differences between Ku70 and Ku80 knockouts (135, 140, 141). This indicates that despite their well-known function as a heterodimer, Ku70 and Ku80 likely possess individual functions independent of heterodimer interactions.

Curiously, both Ku70 and Ku80 possess their own nuclear localization and are capable of entering the nucleus independently, suggesting they may have functions independent of a heterodimer complex (142–144). Moreover, Ku70 has been shown to interact directly with Bax, suppressing its function in apoptosis. Importantly, when Ku70 interacts with Bax it dissociates from Ku80, which suggests structural alterations depending on the interaction state of Ku70 (145–148). Thus, despite early claims, a Ku70/80 heterodimer may not be the exclusive means by which Ku70 and Ku80 function. Further structural and functional studies will be required to understand independent roles of Ku70 and Ku80.



**Figure 7. Domains of Ku70 and Ku80.** Phosphorylation sites are denoted by their residue numbers. vWA denotes the von Willebrand factor type A domains, NLS denotes the nuclear localization signal sequences. Ku70 and Ku80 shows structurally conserved vWA domains followed by Ku core domains. Their C-terminus differs as Ku70 possess a SAP domain whereas Ku80 possesses a DNA-PKcs interacting region. Length of functional domains is not to scale; linker regions <20 residues are omitted for clarity. Figure 6 is adapted and summarized from (149).



**Figure 8. Crystal structure of Ku70/80.** (A) The red polypeptide chain represents Ku70 and the yellow polypeptide chain represents Ku80. (B) Ku70 domains are highlighted, with the vWA domain in green, core domain in red, and SAP domain in cyan. The N- and C-terminus of the protein is labelled, with missing density between the core and SAP domain. Structure from PDB 1JEY is illustrated using PyMOL.

### 1.3.5 PAXX

Paralog of XRCC4 and XLF (PAXX) is a recently discovered factor that appears to play an auxiliary role in c-NHEJ (150–152). A partial structure of PAXX (residues 1–141) revealed significant structural similarity with the N-terminal head domains of XRCC4 and XLF (151). Like XRCC4 and XLF, PAXX is predicted to be disordered in ~30% of its C-terminal region and there is currently no structural information for the full-length version of any of these related proteins.

Recent studies using gene knockouts have uncovered overlapping functions between PAXX and XLF (151, 152). XLF but not PAXX knockout in mice shows dramatic decrease in DSB repair functions (153). Although a single knockout of PAXX shows no obvious defect in repair, a combined knockout of XLF and PAXX confers synthetic lethality in mice (154). Similar to XLF, PAXX interacts with the Ku70/80 complex; however, this interaction is mediated specifically through Ku70 and is dependent on Ku70 interaction with DNA. The Ku70-PAXX interaction has been mapped to the C-terminal region of PAXX<sup>177-204</sup> (151). On its own, PAXX does not appear to interact stably with DNA; however, when Ku70 is pre-bound to DNA containing 10–15 extra bases, PAXX binds with low micromolar affinity. The fact that Ku70 and extra DNA are required for stable interaction suggests that PAXX makes direct contact with both Ku70 and DNA, but neither interaction is sufficient for stable binding on its own.

In addition to its interaction with Ku70, PAXX has been shown by immunoprecipitation to interact with a number of other proteins. Surprisingly, these binding partners (PARP1 and DNA Ligase III) are involved in base excision repair and single strand break repair pathways (74). This may indicate PAXX participates in more than one DNA repair pathway, or perhaps is involved in recruitment of single-strand break repair factors for sealing of the second strand at a DSB. Although PAXX appears to serve a backup function for XLF during c-NHEJ, the mechanism of PAXX activity is only starting to be examined and significant questions remain unanswered. Structural studies of the Ku70-DNA-PAXX complex are required to provide the necessary information to begin answering these questions.

## 1.4 Thesis Objectives

The primary goal of work conducted in this thesis has been to gain structural knowledge of NHEJ factors in higher-order assemblies to address the knowledge gap that exists for understanding of their function in relation to structure and mechanism of action, particularly when in complex and in interaction with DNA. One of the key limitations in conducting such structural studies of NHEJ factors has been the inability to acquire high concentrations of pure protein required for crystallization experiments. Thus, in the first part of the thesis emphasis was placed on devising efficient methods for generating sufficient amounts of several NHEJ proteins from a bacterial expression system.

Following biochemical characterization of purified proteins to ensure activity, structural studies using X-ray crystallography were conducted for several different complexes including: LX4, LX4-DNA, LX4-XLF-DNA, Ku70, Ku70-DNA, Ku70-DNA-PAXX. Many crystallization parameters were investigated for each complex; however, as the most promising crystals were obtained from screens of Ku70-DNA and Ku70-DNA-PAXX, the focus of work was directed towards obtaining structural information for these complexes.

## Chapter 2

### 2 Methods

This chapter outlines the specific procedures and materials utilized to carry out experiments. The expression and purification procedures for all proteins underwent numerous optimization variations in conditions for maximal yield and reproducibility. Once sufficient protein was available, crystallization screening was performed using the hanging drop vapour diffusion method. Following broad screening of various protein complexes with DNA using commercial crystallization screen kits, promising leads were reproduced with laboratory reagents to ensure the pre-made kit conditions were not altered due to passage of time. Subsequently, crystallization conditions were varied to determine optimal crystal formation. Typical parameters that were varied included: pH of the buffer, protein concentration, precipitant concentration, temperature, and rate of dehydration of the crystallization drops, and other individual components unique to a given lead condition. Micro-crystals insufficient in size and/or quality for X-ray diffraction data collection were optimized for growth and crystal quality with additive screening, altered protein concentration and growth time. Once crystals of sufficient size were obtained, X-ray screening was performed. Those crystals that yielded some amount of diffraction were selected for further optimized using modified DNA substrates, altered ratios of protein and DNA, and additional crystallization agent screening. High quality, protein complex crystals were flash frozen in liquid nitrogen and data collected at the Advanced Photon Source (APS) synchrotron in Chicago.

#### 2.1 Plasmid Constructs

In order to improve levels of expression and prevent premature termination, recombinant LX4 and Ku70/80 co-expression constructs were commercially synthesized following codon optimization for bacteria (GenScript). The recombinant XLF construct was created by Dr Sara Andres (68). Ku70 and PAXX constructs were generous gifts from Dr Mauro Modesti (74). Details of each expression construct are described below.

## 2.2 Protein Expression and Purification

### 2.2.1 DNA Ligase IV/XRCC4 (LX4)

In order to improve yield and purity of LX4 complex, full-length DNA Ligase IV was cloned into the NdeI and XhoI sites of a pCOLADuet-1 co-expression vector with addition of a N-terminal His<sub>9</sub>-tag followed by a TEV protease cleavage site. Full-length XRCC4 was then cloned into the same vector using available NcoI and NotI sites. The coding regions for Ligase IV and XRCC4 were codon optimized by GenScript. *Rubidium chloride* (RbCl) competent BL21(DE3)-T1<sup>R</sup> cells were transformed with the expression vector and grown on LB agar plates supplemented with kanamycin (50µg/mL). A single transformant was selected and incubated in 10mL LB media supplemented with 50µg/mL kanamycin at 37°C until OD<sub>600</sub> reached 0.8, after which 1mL was subcultured into 1L terrific broth (TB) supplemented with 50µg/mL kanamycin. Autoinduction was then allowed to proceed at 20°C for ~60h.

Cells were harvested by centrifugation at 10,000g, washed with phosphate-buffered saline (PBS) and re-suspended in lysis buffer (50mM Tris pH8.0, 2M NaCl, 10mM imidazole, 0.4M ammonium acetate, 2mM β-mercaptoethanol, 5% v/v glycerol) prior to storage at -80°C. Four protease inhibitors (1µM pepstatin A, 1mM phenylmethylsulphonyl fluoride [PMSF], 1mM benzamidine, 10µM leupeptin) were added before samples were lysed by four sequential passages through a French press. Lysate was clarified by centrifugation at 50,000g, and applied to a 5ml Ni<sup>2+</sup> charged immobilized metal affinity column (Ni column; HisTrap HP, GE Healthcare). Bound LX4 was washed with 10 column volumes (cv) of lysis buffer, then with 5-10 cv of lysis buffer containing NaCl at a reduced concentration of 150mM. Further washes (10-20 cv) were carried out with the same lysis buffer containing 75mM imidazole. Final elution was performed by step-elution with lysis buffer supplemented with 500mM imidazole. Eluted protein was exchanged into Q buffer (20mM Tris pH8.0, 150mM KCl, 1mM EDTA, 10mM DTT, 10% v/v glycerol) and applied to a MonoQ 5/50 column (GE Healthcare) equilibrated with Q buffer, and eluted with a linear gradient up to 300mM KCl. Purified LX4 was buffer exchanged into storage buffer (20mM Tris pH8.0, 1mM DTT, 100mM KCl, 20% v/v glycerol), concentrated to 10mg/mL using a 30kDa MWCO

concentrator (Vivaspin 6, Sartorius stedim), flash frozen in liquid nitrogen, and then stored at -80°C.

### 2.2.2 XLF

A pET-Duet1 plasmid (68) containing a full length XLF open reading frame with a C-terminal His<sub>6</sub>-tag was used to transform RbCl competent *E. coli* Rosetta™(DE3)pLysS cells. Cell cultures were grown in LB supplemented with 100µg/mL ampicillin, and expression induced with 1mM IPTG when OD<sub>600</sub> reached 0.6. Incubation was continued for 4h at 37°C prior to harvesting by centrifugation at 10,000g.

Cell pellet from 1L of cell culture growth was washed with PBS and re-suspended in lysis buffer (20mM HEPES pH8, 2M NaCl, 10mM imidazole, 10mM β-mercaptoethanol). Four protease inhibitors (1µM pepstatin A, 1mM PMSF, 1mM benzamidine, 10µM leupeptin) were added before cells were lysed by four sequential passages through a French press. Following clarification by centrifugation, soluble protein was applied to a 5ml Ni column (HisTrap HP, GE Healthcare). After sequential washes with lysis buffer containing 40mM and 90mM imidazole, XLF was step eluted with buffer at 250mM imidazole. The eluent was exchanged into Hep buffer (20mM HEPES pH8.0, 150mM NaCl, 1mM EDTA, 10mM β-mercaptoethanol) and applied to a 5mL Heparin column (HiTrap Heparin HP, GE Healthcare). XLF was eluted using a linear gradient of Hep buffer from 0.15-1.0M NaCl at 1ml/min. Fractions containing pure XLF were pooled, buffer exchanged into XLF storage buffer (20mM Tris pH8.0, 150mM KCl, 1mM TCEP), concentrated to ~12mg/mL using 10kDa MWCO concentrator (Vivaspin 6, Sartorius stedim), flash froze in liquid nitrogen, and stored at -80°C.

### 2.2.3 Ku70/80

A pCOLADuet-1-based co-expression system was constructed to express Ku70/80 within a bacterial expression system. Ku80<sup>5-545</sup> with a N-terminus His<sub>9</sub>-tag was inserted into available NcoI and NotI sites, while Ku70<sup>33-609</sup> was cloned into NdeI and XhoI sites. The construct domains were chosen based on the current crystal structure of Ku70/80 (PDB 1JEY) in order to specifically aid crystallography studies. *E. coli* BL21(DE3)-T1<sup>R</sup> cells were transformed and colonies grown on LB agar plate supplemented with kanamycin



(50µg/mL). A single transformant was incubated in LB supplemented with 50µg/mL kanamycin until OD<sub>600</sub>~0.8. At this point, 10mL of culture was transferred to 1L of TB media supplemented with 50µg/mL kanamycin and allowed to autoinduce at 20°C for ~60h.

Cell pellets were washed with PBS and re-suspended in lysis buffer (50mM Tris pH8.0, 1M NaCl, 10mM imidazole, 0.4M ammonium acetate, 2mM β-mercaptoethanol, 5% v/v glycerol). Four protease inhibitors (1µM pepstatin A, 1mM PMSF, 1mM benzamidine, 10µM leupeptin) were added before cells were lysed by four sequential passage through a French press. Lysate was clarified via centrifugation at 10,000g, and soluble protein applied to a 5mL Ni column (HisTrap HP, GE Healthcare) equilibrated with lysis buffer. Bound Ku70/80 was sequentially washed (10 cv) with lysis buffer, after which salt concentration in the lysis buffer was reduced to a concentration of 50mM. A further wash (10 cv) was performed with the same lysis buffer containing 45mM imidazole prior to step-elution with 500mM imidazole. Eluted protein was directly applied to a Heparin column (HiTrap Heparin HP, GE Healthcare), equilibrated with Hep buffer (50mM Tris, pH8.0, 50mM NaCl, 0.5mM EDTA, 1mM DTT), and eluted using a linear gradient up to 1M NaCl. Purified Ku70/80 was buffer exchanged into storage buffer (50mM Tris pH8.0, 1mM EDTA, 1mM DTT, 100mM NaCl, 20% v/v glycerol), concentrated to ~11mg/mL using 30kDa MWCO concentrator (Vivaspin 6, Sartorius stedim), flash frozen in liquid nitrogen, and stored at -80°C.

## 2.2.4 Ku70

A pETDuet1 plasmid containing a full-length Ku70 open reading frame with a C-terminal His<sub>6</sub>-tag was expressed in *E. coli* Rossetta™2 (DE3) cells. Cells were grown at 37°C in LB media supplemented with ampicillin (100µg/mL) and expression induced with 0.5mM IPTG when cell density reached an OD<sub>600</sub> of ~0.6. The cells were further grown for 18h at 16°C prior to harvesting by centrifugation at 10,000g.

Cell pellets were washed with PBS and re-suspended in NiA buffer (50mM HEPES pH7.5, 1M NaCl, 15mM imidazole, 0.4M ammonium acetate, 5% v/v glycerol, 2mM β-mercaptoethanol). Four protease inhibitors (1µM pepstatin A, 1mM PMSF, 1mM

benzamidine, 10 $\mu$ M leupeptin) were added before cells were lysed cells by four sequential passages through a French press. Cell debris was removed by centrifugation and the clarified lysate was applied to a 5mL Ni column (HisTrap HP, GE Healthcare) equilibrated with Ni buffer. Protein which bound to the column was washed sequentially (10 cv per wash) with NiA buffer, NiB buffer (50mM HEPES pH7.5, 100mM NaCl, 15mM imidazole, 5% v/v glycerol, 2mM  $\beta$ ME), NiB buffer containing 50mM imidazole, and step-eluted with NiB buffer containing 500mM imidazole. Eluted protein was directly applied to a HiTrap Heparin HP column (GE Healthcare), equilibrated with Hep buffer (50mM HEPES pH7.5, 5% v/v glycerol, 100mM NaCl, 0.5mM EDTA, 1mM DTT). Bound protein was washed with 5 cv of Hep buffer containing 200mM NaCl, prior to elution with a linear gradient from 200 to 400mM NaCl at 1 ml/min. Purified Ku70 was buffer exchanged into Ku storage buffer (20mM HEPES pH7.5, 100mM NaCl, 1mM TCEP), concentrated to ~8mg/mL using 30kDa MWCO concentrator (Vivaspin 6, Sartorius stedim), flash frozen in liquid nitrogen, and stored at -80°C.

### 2.2.5 PAXX

Full-length PAXX with a C-terminal His<sub>6</sub>-tag was inserted into the NdeI-BamHI sites of a pHIS-parallel1 bacterial expression vector. *E. coli* Rossetta™ (DE3) pLysS cells were transformed with the vector for expression. Cells were grown in LB media supplemented with 100 $\mu$ g/mL ampicillin and 25 $\mu$ g/mL chloramphenicol until the density reached an OD<sub>600</sub> of ~0.6. The cell culture was chilled to 4°C by equilibrating the temperature slowly in a cold room before being induced with 0.5mM IPTG for 16h at 15°C.

Cell pellets were washed with standard PBS buffer and re-suspended with lysis buffer (20mM Tris pH8, 800mM NaCl, 10mM imidazole, 10% v/v glycerol, 2mM  $\beta$ -mercaptoethanol). Four protease inhibitors (1 $\mu$ M pepstatin A, 1mM PMSF, 1mM benzamidine, 10 $\mu$ M leupeptin) were added before cells were lysed cells by sonication with three sequential 1min cycles. Cell debris was removed by centrifugation at 10,000g and the clarified lysate was applied to a Ni column (HisTrap FF, GE Healthcare) equilibrated with lysis buffer. Protein that bound to the column was washed sequentially (10 cv) with lysis buffer containing 10mM and 30mM imidazole, and step-eluted with buffer containing 300mM imidazole. Eluted protein was dialyzed into Q buffer (20mM

HEPES pH8, 50mM NaCl, 1mM EDTA, 1mM DTT, 10% v/v glycerol), applied to a HiTrap Q HP column (GE Healthcare) equilibrated with Q buffer, and eluted with a linear gradient from 50 to 400mM KCl. Purified PAXX was dialyzed into storage buffer (20mM HEPES pH8, 150mM NaCl, 1mM TCEP, 10% v/v glycerol), concentrated to ~4.9mg/mL using 10kDa MWCO concentrator (Vivaspin 6, Sartorius stedim), flash frozen in liquid nitrogen, and stored at -80°C.

## 2.3 Protein Quantification

Bradford assays were performed for determination of protein concentration using Pierce™ Coomassie Plus (Bradford) Assay Kit (Thermo Fisher) according to the manufacturer's user instructions.

## 2.4 SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed by applying samples to a 9-15% w/v acrylamide gel supplemented with 0.1% v/v 2,2,2-Trichloroethanol. Electrophoresis was performed at 150V for ~60min using a Mini-PROTEAN® Tetra Vertical Electrophoresis Cell (Bio-Rad). Gels were directly visualized with UV photo-activation or conventional Coomassie Blue staining followed by imaging using a Gel Doc™ EZ Gel Documentation System (Bio-Rad).

## 2.5 Western Blot

Following separation by SDS-PAGE, the gel was stacked with a polyvinylidene difluoride membrane pre-wet with 100% methanol. The transfer was carried out in transfer buffer (50mM Tris pH8.0, 40mM glycine, 20% v/v methanol) using the Bio-Rad Trans-Blot Turbo transfer system (Bio-Rad), at 75 amperes for 115min. The membrane was then blocked in generic 5% w/v fat-free powdered milk dissolved in Tris buffered saline (TBS; 50mM Tris pH8, 150mM NaCl) supplemented with 0.05% v/v Tween-20 (TBS-T) for 1h at room temperature. After blocking, the membrane was incubated with Ku70 (N3H10) antibody (Santa Cruz) diluted in 3% w/v fat-free powdered milk in TBS-T at 4°C for 18h with slow shaking. Ku70 antibody was a generous gift from Dr Caroline Schild-Poulter. The blot was rinsed three times using TBS with 3% w/v milk before

being incubated with mouse-specific alkaline phosphatase-conjugated secondary antibody. The membrane was then washed (TBS-T with 3% w/v milk) three times for 20min, once with TBS, and finally developed using AP Conjugate Substrate Kit (Bio-Rad), incubated for 10min before the membrane was rinsed with water. Images of western blots were collected using a generic camera.

## 2.6 SEC-MALS Analysis

Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) experiments were performed at 4°C, with an in-line Superdex 200 GL size-exclusion column (GE Healthcare) at a flow rate of 0.5mL/min. Light-scattering experiments utilized 18 angles to determine the mass of protein in a given peak eluted during SEC. Prior to performing SEC-MALS, purified proteins were diluted to 2mg/mL using storage buffer (50mM HEPES pH7.5, 100mM NaCl, 1mM TCEP). Samples were centrifuged at 20,000g for 60sec to remove any large particulate, and injected onto a Superdex 200 10/300 GL size-exclusion column (GE Healthcare) equilibrated with the same storage buffer. Eluted sample was detected by a DAWN HELEOS light scattering detector at a wavelength of 662nm, connected in tandem to an Optilab refractive index concentration detector using a wavelength of 658nm (Wyatt Technology). A refractive index increment of 0.185 was used. The light scattering data was collected and displayed by ASTRA (version 6.1.5.22, Wyatt Technology).

## 2.7 DNA Ligation Assay

In order to ensure that recombinant LX4 sample was fully active, ligation capacity of purified LX4 was tested. A linear DNA substrate of ~2686bp with compatible ends (4 base 5' overhang) was generated by restriction digest of pUC19 using HindIII (NEB) by incubating at 37°C for 1h in Buffer R as per manufacturer's recommendation.

Subsequently, 100pmol of LX4 was incubated with 200ng of linearized pUC19 in T4 ligation buffer (NEB) in a final reaction volume of 20μL at 21°C for 1h. Subsequently, 1μL of Proteinase K (Thermo Fisher) was added and incubated at 50°C for 30min to terminate the reaction and remove all protein bound to DNA. Ligation reactions were

resolved on a 0.8% w/v 0.5x TBE agarose gel run in 0.5x TBE buffer at 50V for 145min and visualized by immersing the gel in sufficient volume of 0.5 $\mu$ g/mL ethidium bromide for 15min with agitation, followed by incubating in water for 15min with agitation. Gel imaging used a Gel Doc™ EZ Gel Documentation System (Bio-Rad) for qualitative analysis.

## 2.8 DNA Substrates for Crystallography

A critical feature of successful crystallization is ensuring a homogeneous, mono-disperse sample prior to crystallization. In this regard, selecting the correct type of DNA is essential. To stabilize a LX4 complex for optimal crystallization, two different types of DNA substrates were chosen for their ability to achieve interruption of the ligation reaction, effectively trapping the enzyme at discrete steps before ligation is complete. For ligation reaction to occur, first the active site lysine in DNA Ligase IV is loaded with an adenylyl group, which subsequently undergoes a S<sub>N</sub>2 reaction by the free electron pair on the 5' phosphate of the broken DNA end base (155). This AMPylation of the 5' phosphate allows it to become a better leaving group for a second S<sub>N</sub>2 reaction using attack by an activated 3' hydroxy group of the upstream strand, completing the ligation reaction. Hence there are two steps in this mechanism that can stall the ligation reaction. First, if the downstream strand does not contain a 5' phosphate group, there can be no AMPylation for a second S<sub>N</sub>2 reaction, blocking ligation. Alternatively, a dideoxyribonucleic acid on the 3' end of the upstream strand will also disable the reaction due to the lack of a 3' hydroxyl group. Utilizing these two strategies, we formulated different DNA substrate designs to screen for LX4-DNA crystal formation (Appendix 1 and 2). Various duplex and multiplex substrates were made with synthesized DNA oligonucleotides for the crystallization trials (Appendix 1). DNA oligonucleotides are synthesized by Bio Basic. Annealing of different DNA substrates was performed by mixing equimolar amounts of oligonucleotide to a final concentration of 100 $\mu$ M and heating to 95°C prior to slow cooling (1°C per minute) using a thermocycler (Thermo Fisher). Specific DNA substrates named and used for crystal screening are listed in Appendix 2.

## 2.9 Electrophoretic Mobility Shift Assay

In order to determine the ability and extent of protein-DNA interaction, electrophoretic mobility shift analysis (EMSA) was performed using recombinant proteins. In these studies, purified protein and DNA substrate were mixed in EMSA buffer (75mM KCl, 10mM Tris pH7.5, 0.5mM EDTA, 0.5mM DTT, 5% v/v glycerol) in a final volume of 10 $\mu$ L and allowed to incubate at room temperature for 30min. The amount of DNA added varied as indicated. Purified protein was diluted to varying concentrations prior to adding to EMSA reactions. DNA loading dye (30% v/v glycerol, 0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol FF) was added to each reaction sample, and DNA components (bound vs free) fractionated by 6% w/v polyacrylamide gel electrophoresis in 0.5x standard TBE buffer at 80V for 45min. The gel was stained with 0.2 $\mu$ g/ml ethidium bromide (EtBr) for 20min and de-stained in water for 15min before being visualized using a ChemDoc system set at a wavelength of 356nm (Bio-Rad).

## 2.10 Crystallography

All crystallization was performed using the hanging drop vapour diffusion method. Initial screening for suitable crystallization conditions to generate protein crystals was performed by mixing 1 $\mu$ L volumes of protein (at varying concentrations) with crystallization conditions obtained from commercial vendors including the MCSG Suites (Anatrace/Microlytic), Wizard series (Rigaku), PACT premier™ (Molecular Dimensions), NeXtal Classic Suite (QIAGEN). Four 96-condition kits were selected for investigating the suitable concentration of protein to yield informative screening results. Once a suitable concentration for screening was determined, additional sparse matrix screening was performed with additional 96-condition kits to obtain a broad range of variation of crystallization components for maximal screening success. Similar procedures were employed to generate protein-DNA crystals. The protein-DNA complex was first formed by mixing protein and DNA at an appropriate molar ratio (typically close to 1:1). Depending on the type of complex and desired buffer components and final concentration of complex, the sample was either concentrated through a Nanosep® Centrifugal Device (PALL) or dialyzed into desired buffer. Crystallization trials were performed by mixing 1 $\mu$ L volumes of protein-DNA complex and crystallization

conditions listed above. All crystallization mixtures are equilibrated over 1mL of 1.5M  $(\text{NH}_4)_2\text{SO}_4$  as reservoir solution, which functioned to slowly dehydrate crystallization sample. Depending on the general trend of solubility, reservoir solution was replaced with higher or lower concentration of  $(\text{NH}_4)_2\text{SO}_4$  for manipulating a desirable level of sample precipitation. Crystallization was performed at 20 and 4°C.

The progress of protein crystallizations was assessed and imaged using a Nikon SMZ1500 zoom stereomicroscope mounted with a Nikon Coolpix 4500 digital camera.

### 2.10.1 Crystallization of LX4, LX4-DNA, LX4-XLF-DNA

Purified LX4, LX4-DNA, and LX4-DNA-XLF underwent crystallization screens using sparse matrix screening kit conditions MCSG I and II (Anatrace) and Wizard I and II (Rigaku). In all cases, the starting protein concentration was ~10mg/mL and the same 500 conditions were used sparse matrix screening. LX4 and the DNA substrate were mixed to a molar ratio of 1:1. Unless otherwise specified, crystallization trials were screened at 4°C. For the protein mixture to reach metastable phase, crystallization mixtures were equilibrated against 800μL of 2M  $(\text{NH}_4)_2\text{SO}_4$  for dehydration.

### 2.10.2 Crystallization of Ku70, Ku70-DNA, Ku70-DNA-PAXX

Initial crystallization screens of Ku70, Ku70-DNA, and Ku70-DNA-PAXX employed condition kits MCSG I - IV (Anatrace), Wizard I and II (Rigaku), and NeXtal Classic Suite (QIAGEN). For Ku70-DNA complex, Ku70 and the DNA substrate were initially mixed to a molar ratio of 1:1.2, which later increased to 1:1.5 after optimization. For Ku70-DNA-PAXX complex, Ku70 and the DNA substrate were first mixed to a molar ratio of 1:1.5, after which PAXX was added in equimolar ratio to Ku70. Unless otherwise specified, crystallization trials were performed at 20°C.

Initial microcrystals of Ku70-DNA complex were obtained in two crystallization conditions: MCSG I #95 (0.2M NaCl, 0.1M HEPES pH7.5, 25% w/v PEG3350; Anatrace) and MCSG II #80 (0.1M Tris pH7, 20% w/v PEG1000; Anatrace). These conditions were reproduced using lab reagent NaCl (BioShop), HEPES (Bio Basic), PEG3350 (Bio Basic), Tris (BioShop), and PEG1000 (Bio Basic). Crystal morphology

was improved through successive variation of parameters including: pH, salt identity and concentration, PEG molecular weight and concentration, additive screens (Silver Bullets, Additive Screen; Hampton), as well as drop volume and ratio of protein mixed with crystallization condition. Chapter 3.5 details the sequential improvement of crystal morphology, which correlated with enhanced diffraction data resolution, and the principle factors influencing the crystallization process.

### 2.10.3 Optimization of Ku70-DNA and Ku70-DNA-PAXX Crystals

DNA substrates were designed in three structural groups to help improve crystallization success by increasing potential crystal contacts. While a duplex region is required for Ku70 binding, one end of the DNA was modified with a stem-loop structure, Y-shaped fork, or polythymine tail to obstruct potential movements once Ku70 dimer was bound. Duplex length was varied to optimize the stability of protein-DNA interaction (Appendix 2). For initial screening of crystallization conditions, Ku70 dimer was mixed with loop14, the DNA substrate, at a 1:1.2 molar ratio with DNA in excess. The final concentration of Ku70 was maintained at 4mg/mL. Later on, other variables were explored including: DNA length/structure, DNA concentration, incubation temperature, small molecule additives, drop ratio, concentration of dehydration solution in reservoir, and individual components of the crystallization conditions that yielded initial hits.

The diffraction data quality of crystals containing Ku70 complexed with DNA were enhanced by varying the structure of DNA substrates. Crystallization trials were performed with conditions from the following screens: MCSG I - IV (Anatrace), Wizard I and II (Rigaku), NeXtal Classic Suite (QIAGEN), The PEGs Suite (QIAGEN), PEGRx (Hampton), PEG smear conditions (made with lab reagents, composition adapted from Chaikuad et al.). Later optimization screens employed included the Silver Bullets (Hampton) screen and Additive Screen (Hampton).

## 2.11 X-ray Diffraction Data Collection

Initial screening of crystals was performed using a MicroMax-007 HF X-ray generator mounted with VariMax optics and Raxis 4<sup>++</sup> detector (Rigaku). Crystals were directly mounted on pins and flash frozen in the cryo-jet stream held at 100 Kelvin.



High resolution diffraction data were collected at the APS, a USA national synchrotron-radiation light source research facility (Argonne, IL, USA).

## 2.12 Structural Determination and Refinement

Diffraction data were indexed, integrated, and scaled with iMOSFLM of the CCP4 suite (157). The structure of PAXX was solved using PDB 3WTF as a search model for molecular replacement performed using Phenix (158). The initial model was refined using iterative cycles of phenix.refine (159) and manual model adjustment using coot (160) until R values converged.

## Chapter 3

### 3 Results

In this chapter, advances made towards structural determination of NHEJ factors in higher order assemblies are described. These include: modification of expression and purification methods to substantially increase NHEJ protein availability, development of novel LX4 and Ku70/80 bacterial co-expression constructs, expression and characterization of a novel Ku70 homodimer from bacteria, and crystallization of Ku70-DNA and PAXX.

#### 3.1 Improvements in Production of NHEJ Factors

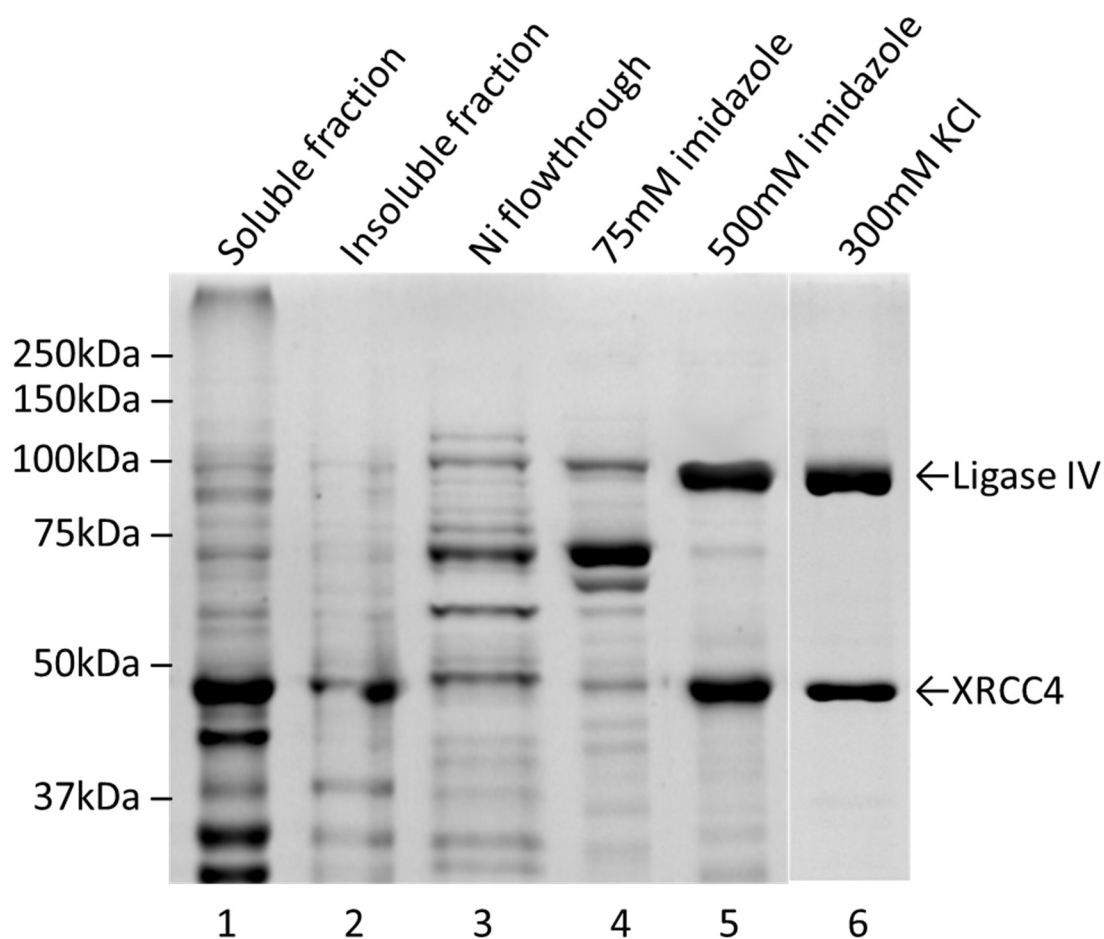
The first step toward enabling crystallographic studies of NHEJ assemblies is achieving methods for easy acquisition of milligram amounts of soluble and purified protein. To this end, the general strategy taken (for LX4, Ku70/80 and Ku70) involved developing novel expression methods (optimizing domain boundaries, codon usage, construction of co-expression vectors, choice of bacterial strain for expression, etc.). Previous systems for expression of XLF (161) and PAXX (74) were also adapted from the literature and modified to maximize yield and reproducibility. Collectively, this work resulted in desired improvements for expression and purification with the highest reported yields to date for LX4, XLF, Ku70/80, Ku70 and PAXX (Table 2).

##### 3.1.1 Purification of LX4, XLF and Ku70/80

Using existing methods for production of LX4 were prohibitive for structural studies due to low yield and purity. Consequently, a new LX4 construct was engineered to explore improvements in expression and purification quality. Since XRCC4 is typically expressed at higher levels than Ligase IV, and Ligase IV is unstable when unbound to XRCC4, a His9-tag was added to Ligase IV instead of XRCC4 so that LX4 will co-elute as a complex during purification. Additionally, both human open reading frames were codon optimized for bacterial expression. This co-expression vector led to more than 10-fold increase in production efficiency compared to previous vectors (161) (Table 2). The purification procedure was also able to be simplified by reducing the number of columns

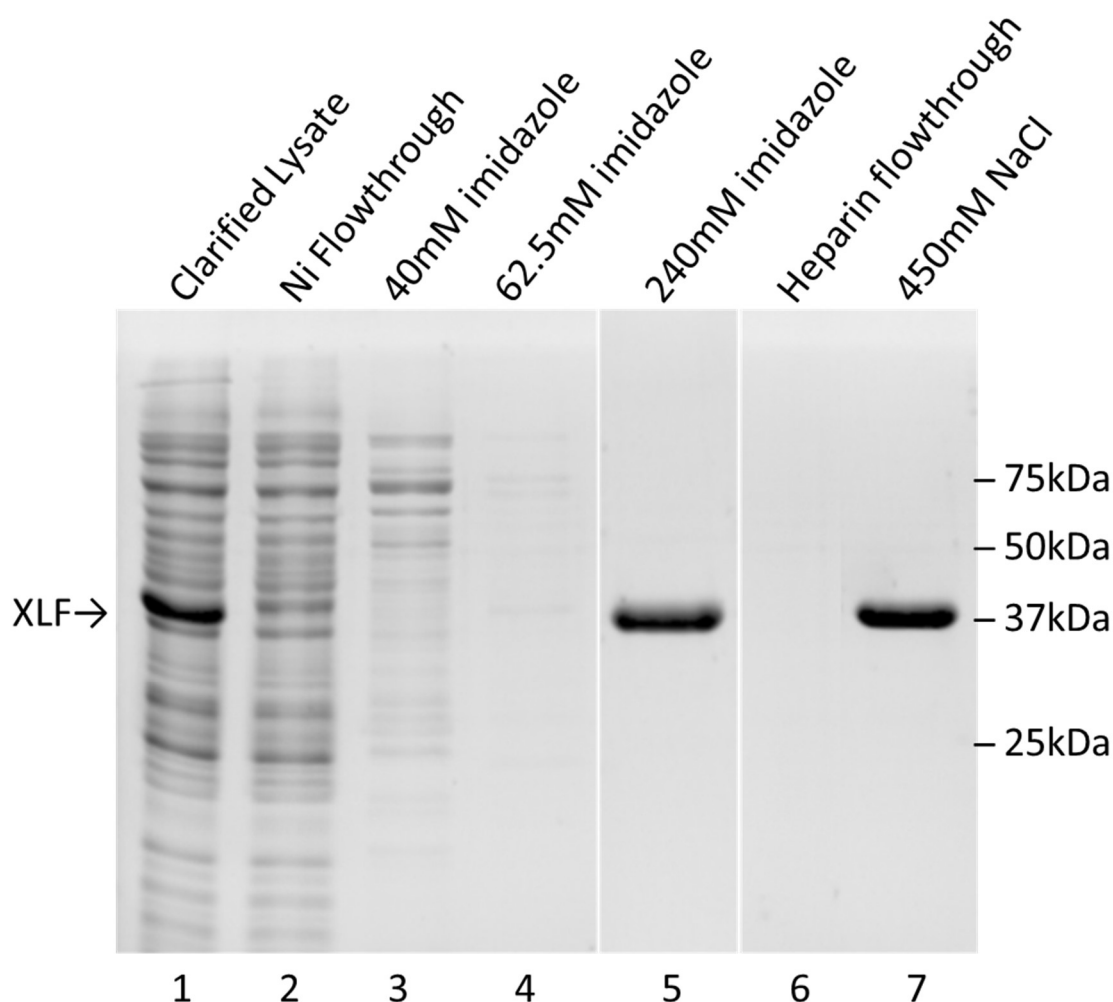
from three (IMAC followed by two types of ion exchange) to two (IMAC and ion exchange) (Figure 9). A significant amount of LX4 complex was required for crystallization screening due to the relatively high degree of LX4 solubility, which warranted screening at higher than usual protein concentrations. Using the method for LX4 expression and purification developed here, LX4 could be maintained in a soluble, stable state at >15mg/mL making further structural studies feasible.

Two different XLF expression and purification protocols were previously established (162, 163), with the main differences being choice of bacterial strain, purification buffer components and chromatography column choices. After comparing purification efficiency and replicability between these two methods, one was chosen for further optimization (Figure 10). Use of *E. coli* Rosetta™(DE3)pLysS cells for expression led to a reduction in XLF degradation, among other contaminants (Appendix 3), presumably due to an enhanced ability of these cells to express eukaryotic proteins that contain codons rarely used in *E. coli*. Although XLF is not highly toxic to bacteria, use of Rosetta™(DE3)pLysS cells permitted better control over leaky expression that may also have helped to improve protein yield and quality. Another important alteration to XLF production came from changing the type of ion exchange column from a MonoQ to Heparin, which resulted in better resolution of protein elution. Heparin not only acts as a general ion exchange resin, but also mimics some aspects of DNA and is therefore often more suitable for purification of DNA binding proteins (164). Further details of the final protocol developed for XLF expression and purification are presented in Chapter 2.2.2.

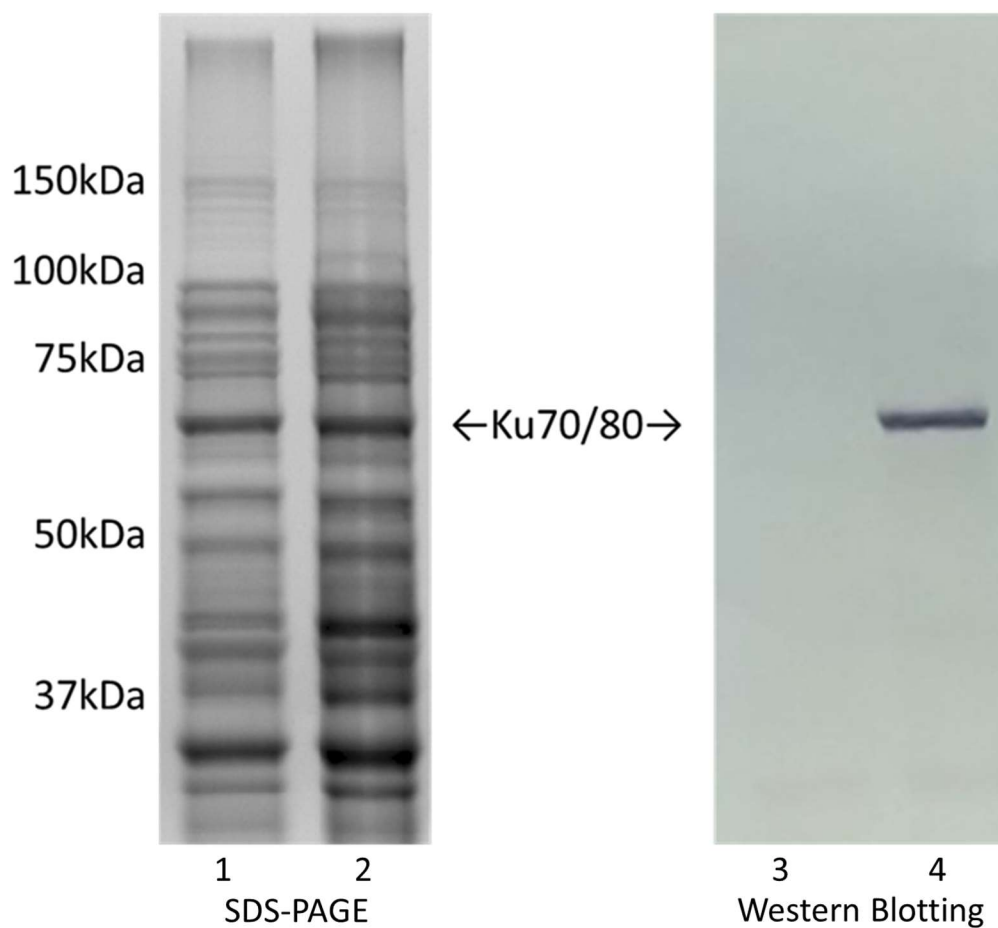


**Figure 9. Representative purification of LX4 complex.** Lane 1 shows the soluble fraction of proteins applied to Ni column; lane 2 shows protein extracted from insoluble sample of cell lysate; lane 3 shows unbound protein; lane 4 shows protein eluted by buffer solution with elevated imidazole concentration; lane 5 shows sample eluted from Ni column; lane 6 shows the purified LX4 eluted from MonoQ column.

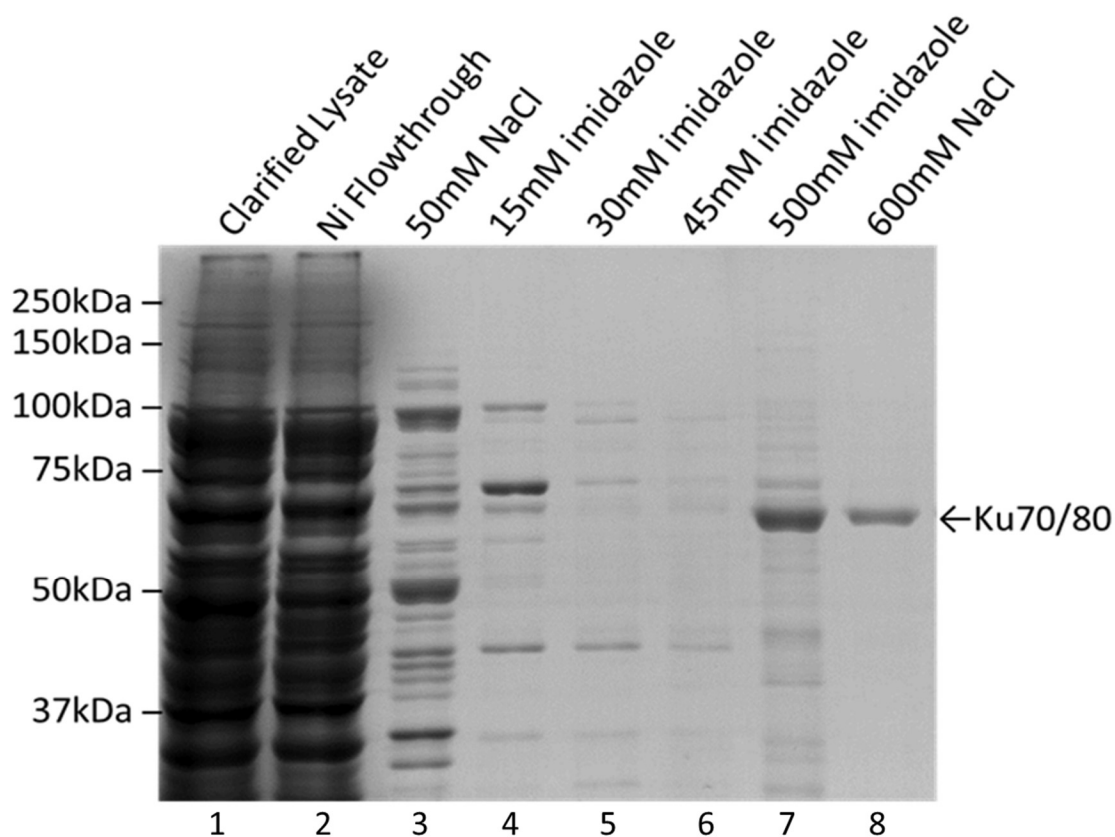
Previous methods for generating Ku70/80 required expression in insect cells, which is both time consuming and expensive. A new Ku70/80 bacterial co-expression construct was designed based on domain boundaries from the current crystal structure. Following codon optimization, this co-expression vector showed Ku70/80 expression using either IPTG induction in *E. coli* Rosetta™ 2(DE3), or auto-induction in BL21(DE3)-T1<sup>R</sup>. Because of increased levels of expression, auto-induction was chosen for large scale production of Ku70/80. Unfortunately, both Ku70 and Ku80 migrated at the same position when resolved by SDS-PAGE. Since Ku80 was tagged with a His<sub>9</sub>-fusion for purification by IMAC, a Ku70-specific antibody was used to confirm its presence in the co-migrating bands (Figure 11). A previous purification protocol for Ku70/80 expressed from bacterial cells (165) was used as a starting point for further modification based on empirical evidence of protein yield and purity. The final optimized protocol for bacterial co-expression of Ku70/80 produced ~1.2mg of pure protein from 1L of cell culture growth and therefore provides a significant improvement to the ease and cost in preparation of Ku70/80 for future structural studies involving complexes formed by Ku70/80 (Figure 12).



**Figure 10. Purification of XLF.** A representative purification profile of XLF is shown, where the arrow indicates the expected size of XLF. Lane 1 shows the soluble fraction of proteins applied to Ni column; lane 2 shows unbound protein; lane 3 and 4 shows protein eluted with elevated imidazole concentration; lane 5 shows the eluted protein from Ni column; lane 6 shows unbound protein from the Heparin column; lane 7 shows the final purified XLF eluted from Heparin column. The lanes are rearranged from the original gel for better comparison.



**Figure 11. Ku70/80 Autoinduction profile.** Lane 1, cell growth after 3h mirrors uninduced state of typical LB+IPTG induction system. lane 2, after 60h of cell growth, the level of Ku70 induction is similar to that of IPTG induced cells. lane 3, western blotting was performed with Ku70-specific antibodies showing no induction. lane 4, purified Ku70/80 sample (from Lane 8, Figure 12) allows much clearer identification of Ku70 induction.



**Figure 12. Purification of Ku70/80.** Lane 1 shows the soluble fraction of proteins applied to Ni column; lane 2 shows unbound protein; lane 3 shows protein eluted with buffer of decreased ionic strength; lane 4, 5 and 6 shows contaminants from elevated imidazole washes; lane 7 shows the eluted protein from Ni column; lane 8 shows the final purified Ku70/80 eluted from a Heparin column.

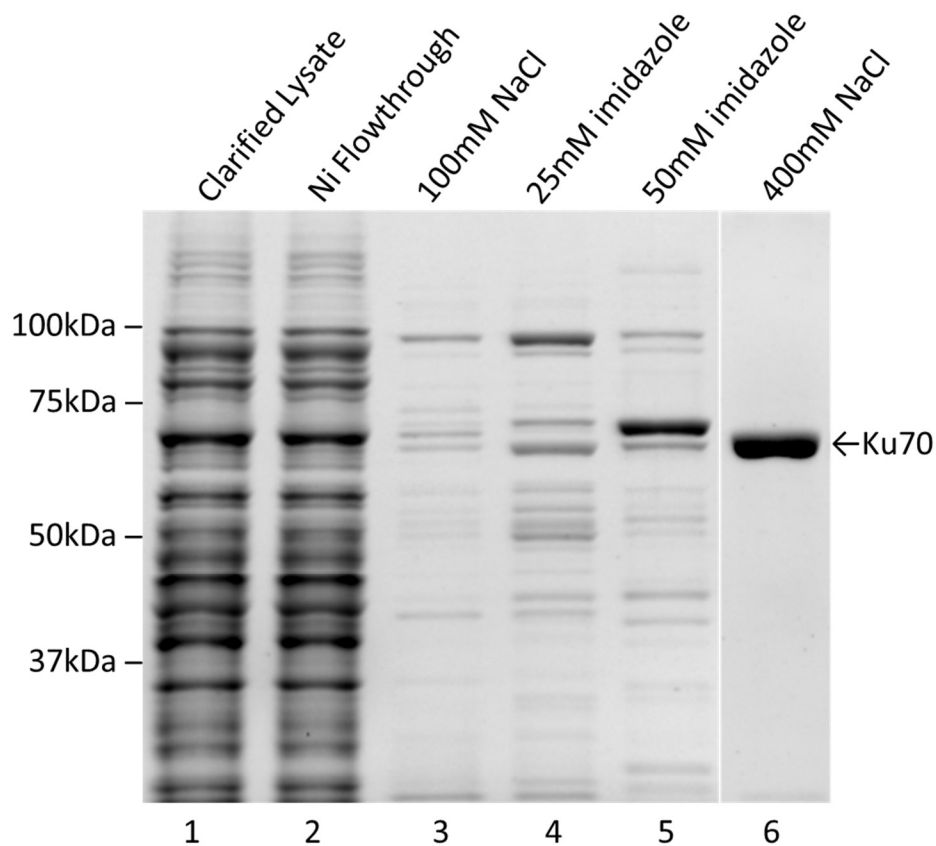


### 3.1.2 Purification of Ku70 and PAXX

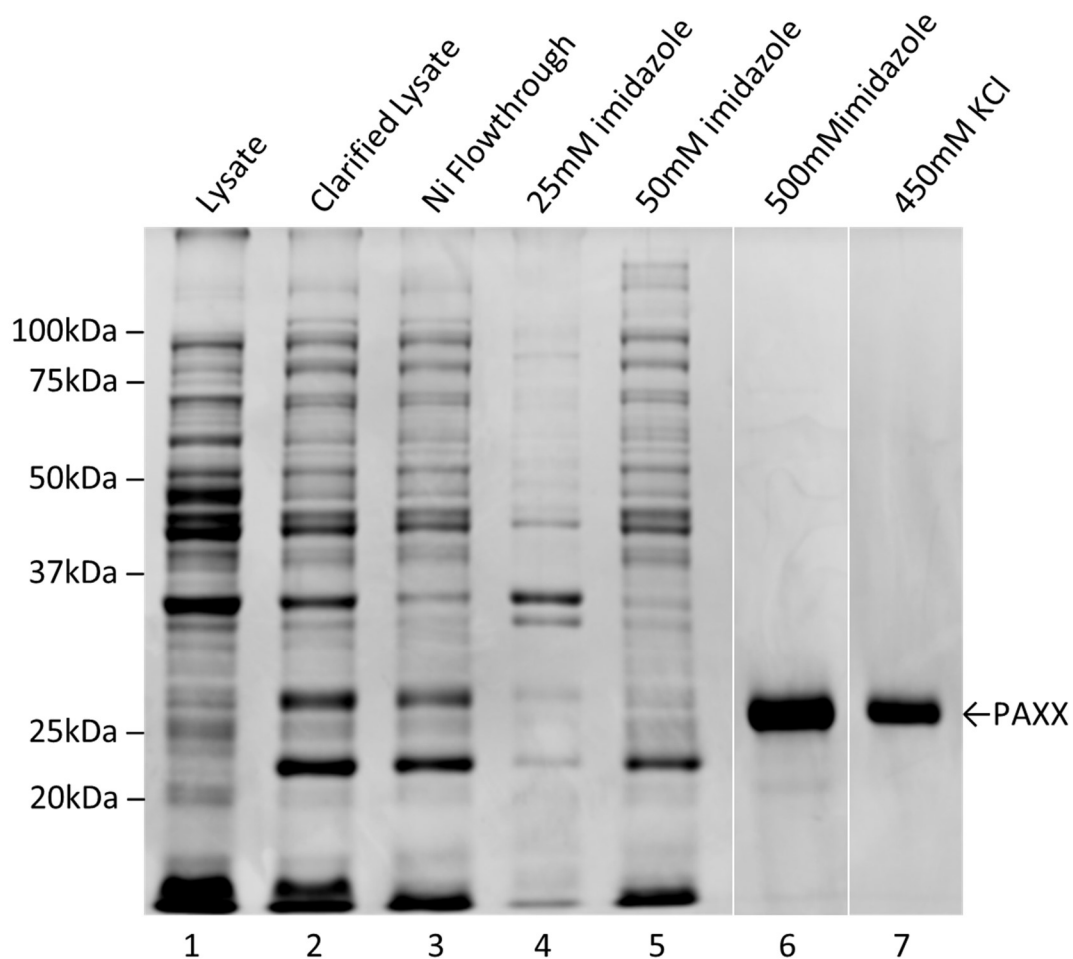
Recent reports in the literature suggest that Ku70 and Ku80 may not function entirely as a heterodimer. This and the fact that Ku70 interacts specifically with other factors such as PAXX, suggested that structural studies of Ku70 alone would be of interest. Contrary to what the field would expect, expression and purification of Ku70 was readily achieved in the absence of Ku80 from bacterial cells. Purification of full-length Ku70 utilized a nickel affinity chromatography step in which elution from the IMAC column was performed in tandem with an inline Heparin column to minimize time and sample loss. This method resulted in high yield and very desirable purity for Ku70 (Figure 13). While the protein remained in solution during concentration up to ~8mg/mL with no significant precipitation, Ku70 appeared to be more stable in the presence of DNA. Qualitative assessment showed that a sample of Ku70 became cloudy after prolonged storage at 4°C, but regained transparency once DNA was added.

High level expression and methods for efficient purification have been reported for PAXX, and therefore no significant modifications were required for PAXX production used in structural studies with Ku70-DNA (162; Figure 14). In order to minimize component variability when producing Ku70-PAXX-DNA complex for crystallization, the final PAXX storage buffer was modified to resemble that of Ku70 storage buffer.

A comparison of protein production efficiency for all NHEJ factors used for screening of crystallization conditions is presented in Table 2. In all cases, the amount of pure protein produced from a single litre of bacterial cell growth exceeded 1mg making these proteins amenable to further structural studies using X-ray crystallography.



**Figure 13. Purification of Ku70.** Lane 1 shows the soluble fraction of proteins applied to Ni column; lane 2 shows unbound protein; lane 3 shows protein eluted with buffer of decreased ionic strength; lane 4 and 5 shows contaminants from elevated imidazole washes; eluted protein was directly applied to a Heparin column, with lane 6 showing the final eluted Ku70 sample. Gel lanes are contracted for clarity.



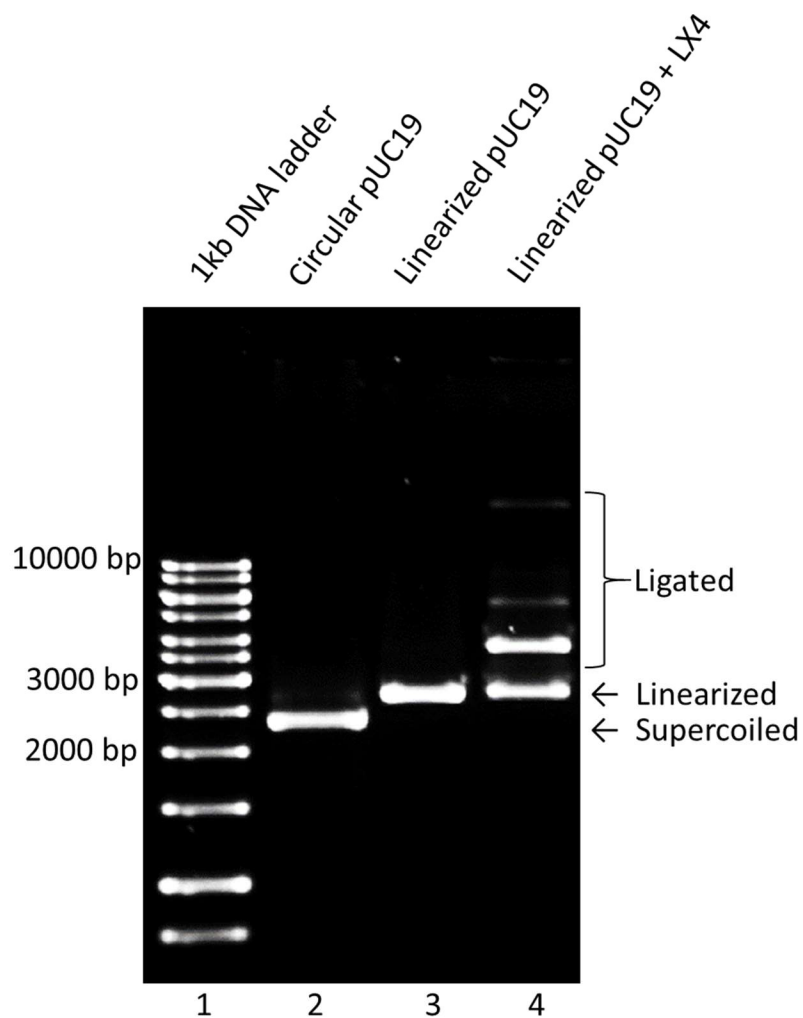
**Figure 14. Purification of PAXX.** Lane 1 shows the total protein from induced cells; lane 2 shows the soluble fraction of proteins applied to Ni column; lane 3 shows unbound protein; lane 4 and 5 shows contaminants from elevated imidazole washes; lane 6 shows the eluted protein from Ni column; lane 7 shows the final eluted PAXX sample. The lanes are rearranged for clarity.

**Table 2. Summary of purification yield of NHEJ factors.** Extraction ratio is in relation to the total *E. coli* protein from clarified lysate of 1L of cell growth.

LX4	Total protein (mg)	Extract ratio (%)
Soluble lysate	205.2	100.0%
Ni elution	14.3	7.0%
MonoQ elution	10.9	5.3%
XLF	Total protein (mg)	Extract ratio (%)
Soluble lysate	156.5	100.0%
Ni elution	18.6	11.9%
Heparin elution	12.8	8.2%
Ku70/80	Total protein (mg)	Extract ratio (%)
Soluble lysate	318.1	100.0%
Ni elution	3.5	1.1%
Heparin elution	1.2	0.4%
Ku70	Total protein (mg)	Extract ratio (%)
Soluble lysate	209.1	100.0%
Ni elute	8.0	3.8%
Heparin elution	3.7	1.8%
PAXX	Total protein (mg)	Extract ratio (%)
Soluble lysate	92.7	100.0%
Ni elution	29.8	32.1%
MonoQ elution	15.8	17.0%

### 3.2 Purified LX4 Retains Ligation Activity

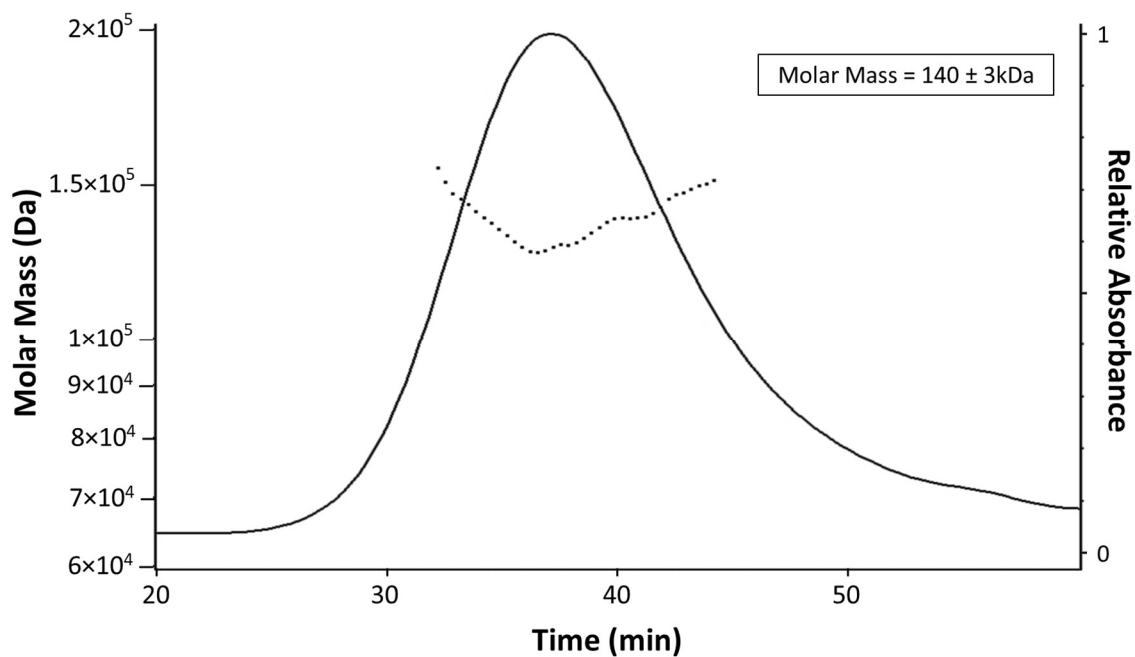
The functional fidelity of LX4 generated from bacterial co-expression was assessed using a standard DNA ligation assay. Purified LX4 sample exhibited robust ligation activity with a HindIII linearized pUC19 DNA substrate containing compatible ends (Figure 15). Typical of LX4 from insect and mammalian cell expression, bacterial LX4 protein produced intermolecular ligation products corresponding to two, three, and four pUC19 linear plasmids ligated together.



**Figure 15. DNA ligation assay showing LX4 activity.** Lane 1 shows a 1kb DNA Ladder (Geneaid); lane 2 shows supercoiled pUC19 plasmid; lane 3 shows pUC19 linearized with HindIII; lane 4 shows DNA ligation products generated by purified LX4 complex after 1h incubation at 21°C using identical amount of DNA as lane 3.

### 3.3 Ku70 Forms a Stable Dimer in Solution

Although Ku70 and Ku80 may function independent of heterodimer formation in cells, whether this involves self-oligomerization is not known. Using SEC-MALS, the molecular mass of species in a solution can be accurately determined. Purified Ku70 was applied to a size-exclusion column, where a single species of protein complex eluted. By placing the MALS detector in tandem with the size exclusion column, a light scattering profile of the eluted protein was found to be ~140kDa, indicating that Ku70 forms a homodimer in solution (Figure 16). Complex formation with DNA did not alter this oligomeric state (Appendix 4). Importantly, results from this analysis indicate that Ku70 exists in a defined (monodisperse) state in solution, suggesting that it is well suited for X-ray crystallographic studies. As well, the fact that Ku70 forms a homodimer in solution further suggests that this quaternary structure may represent the Ku70 structure responsible for Ku80-independent functions (167, 168). To examine and compare the Ku70 homodimer with that of Ku70/80, structural studies of Ku70 were performed for the DNA-bound and unbound states (see section 3.6).

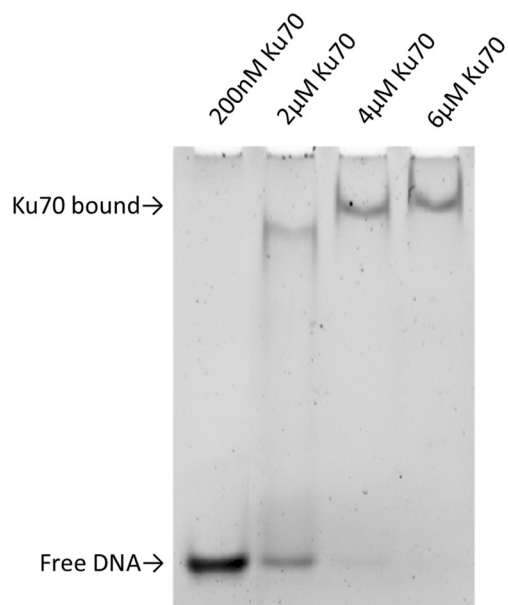


**Figure 16. SEC-MALS data showing stable dimer of Ku70 in solution.** The left y-axis indicates the estimated molecular mass of the sample, corresponding to the dotted line; right y-axis indicates the UV absorbance scale of the eluted protein, corresponding to the solid line; x-axis indicates the elution time of sample from the size exclusion column (flow rate  $\sim 0.5$  mL/min).



### 3.4 Ku70 Homodimer Exhibits DNA Binding Activity

Although Ku70 was shown to exist as a stable homodimer in solution, it remained possible that some misfolding had occurred. In order to verify the functional integrity of Ku70, its DNA binding capacity was determined using EMSA analysis. As shown in Figure 17, Ku70 homodimer readily shifted 16bp DNA with an apparent affinity in the low micromolar range. Thus, Ku70 DNA binding capacity was found to be similar to that previously reported for Ku70/80 produced in insect cells. Taken together, results from the characterization of bacterial expressed Ku70 suggest that Ku70 is well suited for structural studies in complex with DNA or other binding partners such as PAXX.



**Figure 17. Analysis of Ku70 DNA binding activity with loop16 DNA substrate.** The arrows indicate the species of Ku70-bound DNA. Samples were separated on an 8% w/v polyacrylamide gel, and DNA visualized by staining with ethidium bromide.

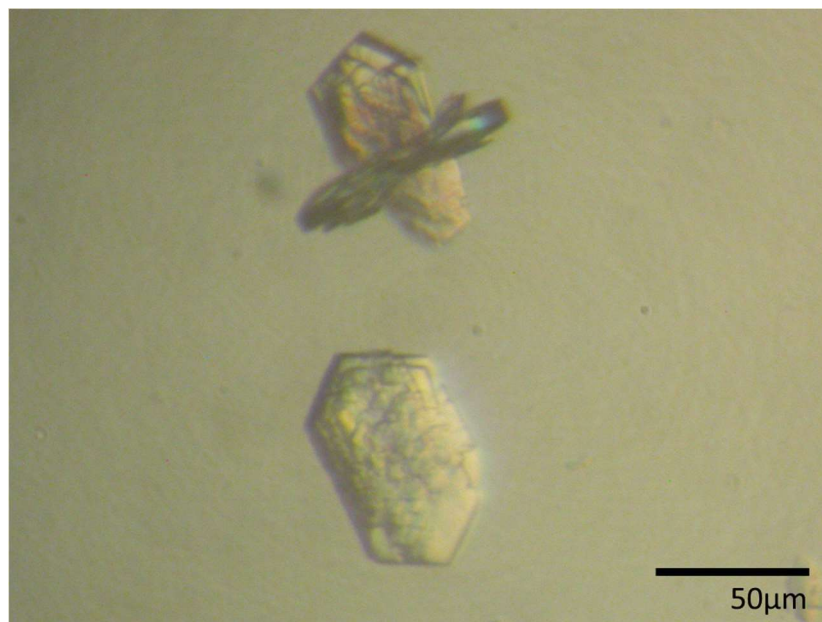
### 3.5 Crystallization Screening of LX4, LX4-DNA and LX4-XLF-DNA

Building on success in obtaining large amount of soluble, functional NHEJ factors, sparse matrix screening trials were performed to determine crystallization conditions to enable structure determination of several LX4-based complexes including: LX4, LX4 with DNA, LX4 with XLF and DNA. Due to the dynamic nature of filament formation that is proposed to coordinate the ligation complex, crystallization strategies needed to consider DNA substrates that were designed to simulate ligation, but at the same time selectively limit the flexibility of the filamentous interactions of the complex. Here we produced three types of rationally designed substrates that focused on facilitating complex assembly without ability to complete ligation, effectively “stalling” the complex and capturing a monodisperse sample on DNA ends (Appendix 2). The first substrate used (18stall) contained a dideoxyribonucleotide at the 3’ end of the nicked strand nicked, disabling its ability to attack the AMPylated 5’ phosphate on the opposite strand, effectively stalling the reaction before the final ligation could be carried out. The second substrate (stack9) contained a 5’-OH at the nick site and therefore prevented initial AMPylation. This substrate contained cohesive ends that allowed elongated DNA structures to assemble in tandem. This design was chosen with the hope that such a substrate might assist in formation of an orderly lattice. Similar DNA substrates have been used successfully for crystallization of XRCC4/XLF complexes (Chris Brown, unpublished data). Since prior crystallization trials (161) were attempted without success at 20°C, and a lower temperature often increases protein stability, crystal screening for LX4 was initially conducted at 4°C.

LX4 was found to be highly soluble. Even when a starting concentration of 10mg/mL was used for sparse matrix crystal screening, most of the drops remained clear. This suggests a higher protein concentration may be optimal for screening. Nevertheless, under these conditions, two different crystal morphologies were identified (Figure 18). X-ray diffraction was carried out on both crystals; however, neither crystal showed any diffraction. Thus, significant additional optimization of these crystallization conditions would be required to yield diffraction quality crystals.

Overall, LX4 complex remained soluble in ~65% of sparse matrix conditions tested and precipitated in the remaining 35% of conditions. As an ideal ratio of conditions resulting in soluble or precipitated protein is about 1:3 so that the protein complex would reach supersaturated state in most conditions, subsequent screening of LX4-DNA complex was performed with the final LX4 complex concentration increased to 11mg/mL. Despite the higher concentration of protein complex used, the screens still gave results with about 50% soluble protein. This finding suggests that, despite a higher concentration of the final complex, the addition of DNA led to less precipitation of the protein complex compared to LX4 crystallization performed in the absence of DNA. Although such an effect is often associated with increased stabilization that can lead to greater success in crystallization, few crystallization hits were obtained under these conditions. Further screening could therefore be performed using higher concentrations of LX4-DNA complex. With this in mind, screening for crystallization conditions of LX4-XLF-DNA was carried out at concentrations of 50 $\mu$ M LX4, 100 $\mu$ M XLF and 50 $\mu$ M DNA, reflecting a 1:2:1 molar ratio as predicted by structural modelling (163). This resulted in better outcomes with only ~25% of conditions showing a lack of precipitation. After two months of vapour equilibration, no crystals were observed and therefore the temperature was adjusted from 4°C to 20°C for further evaluation.

Since the results of initial crystal screening for LX4 complexes was less promising than those of Ku70 and PAXX (see section 3.6), and a decision was made to pursue an M.Sc. instead of a Ph.D., subsequent structural pursuits were focused towards Ku70, Ku70-DNA and Ku70-DNA-PAXX complexes.



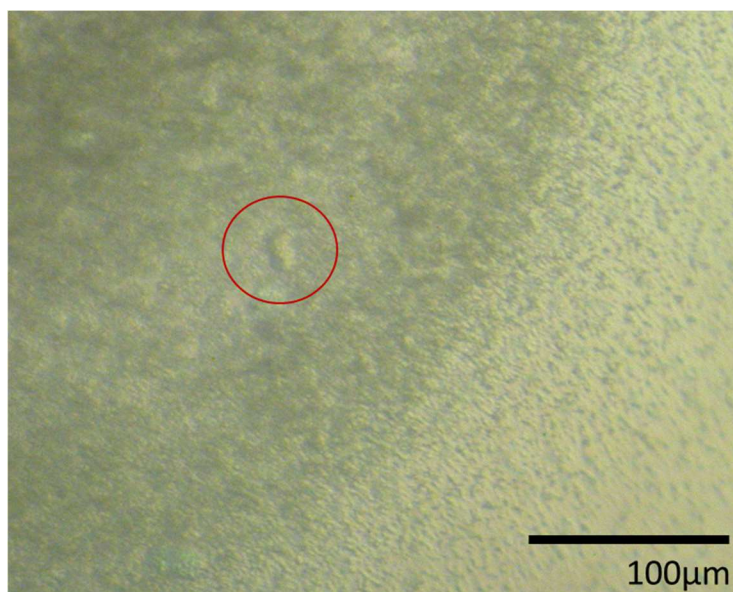
**Figure 18. Unoptimized LX4 crystals.** Crystals appeared after 3–4 weeks with 11mg/mL LX4 complex in 0.2M ammonium acetate, 0.1M Tris8.5, 25% (w/v) PEG3350. The drop was set at 1:1 protein to condition volume ratio, and incubated at 4°C. Two distinct crystal morphologies formed (polygonal with radiating needles, and hexagonal) with the most dominant form being hexagonal.

### 3.6 Crystallization of Ku70 and Ku70-DNA

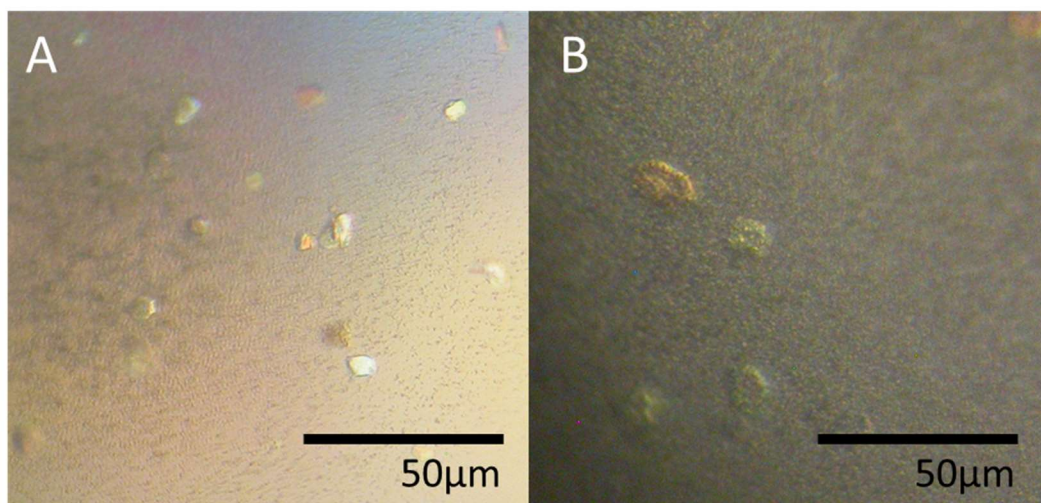
Several concentrations of Ku70 were initially tested for overall suitability for broad sparse matrix crystal screening. A concentration of 4mg/mL was determined to yield the most informative results from crystallization trials using commercially available kits. At this concentration, approximately 75% of the screened conditions resulted in drops with significant protein precipitation. Despite extensive screening using sparse matrix sampling, no promising leads were found. Several conditions generated microcrystals that did not diffract. Further optimization through additive screen, volume ratio and dehydration variance did not improve their crystal growth nor X-ray diffraction quality (Figure 19, Appendix 5). Ku70 alone may have too much structural flexibility in the absence of DNA to readily accommodate ordered crystal lattice interactions.

In contrast to Ku70 alone, promising leads were obtained upon screening of Ku70-DNA complex at similar concentrations (Figure 20, Appendix 6). Microcrystals resulting from Ku70-loop14 complex screening were tested by X-diffraction. Diffraction at low resolution was observed for microcrystals from several crystallization conditions, suggesting that crystals were formed from protein and not salt. This interpretation was further supported by lack of crystal formation in negative controls that contained all crystallization components except Ku70 and/or loop14 DNA. The fact that crystals with similar morphology were grown in several different crystallization conditions suggested that the crystal would likely be amenable to optimization of improved growth and diffraction quality. Because crystal quality of protein-DNA complexes is notoriously dependent on DNA length and end structure, subsequent screening efforts were focused on exploring the effect of different DNA substrates. DNA substrate that was 14bp in length with an additional single base overhang on each strand showed propensity to crystallize on its own, reflected in their small unit cell sizes of 26.48, 26.48, 64.71 and 27.01, 27.01, 58.92, respectively (Figure 21). This likely occurred since the substrate corresponded to ~1.5 helical turns of DNA, allowing easier packing into the structured lattices of a crystal. The results from the overall screening of Ku70 with various types of DNA substrate are summarized in Appendix 6. Out of 40 crystals identified from sparse matrix screens, 24 were composed of salt; 2 were DNA alone; 11 were non-diffracting

due to small size or poor crystal packing and could not be improved by further optimizations; and 3 represented promising leads. Promising initial crystallization conditions were obtained from the following kits: Wizard I #10, MCSG II #80, and MCSG I #95 (Figure 22, Table 3).

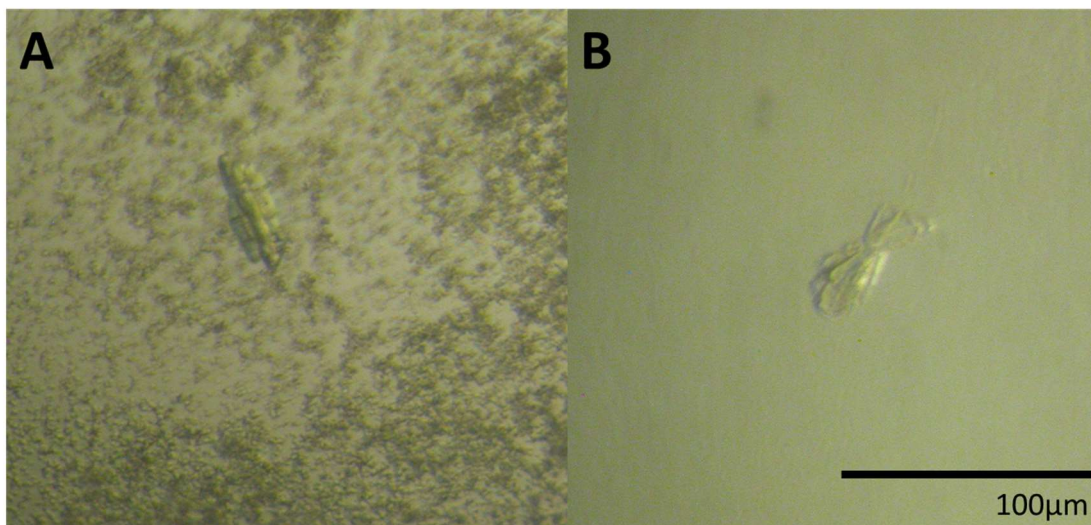


**Figure 19. Typical Ku70 non-diffracting microcrystals.** The red circle highlights one of the rectangular shaped microcrystals among precipitation. The crystal shown was formed in 0.2M calcium acetate, 0.1M imidazole pH8, 20% w/v PEG1000. The drop was set with 1:1 protein to condition volume ratio, and incubated at 20°C.

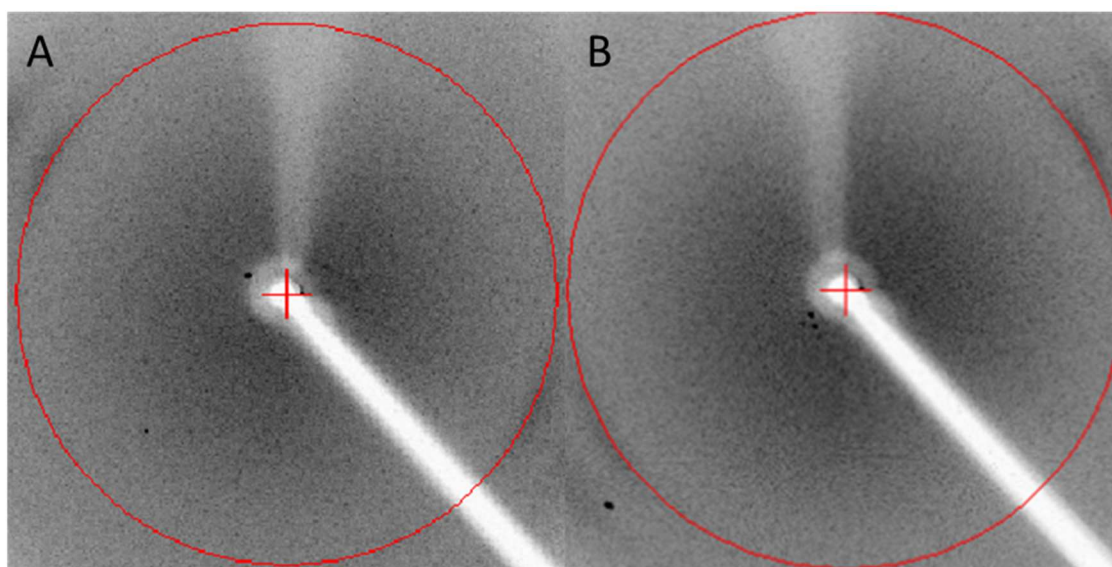


**Figure 20. Initial hits of Ku70-DNA crystals.** These crystals were detected in mixtures with Ku70-loop16, under (A) 0.2M sodium chloride, 0.1M HEPES pH7.5, 25% w/v PEG3350 and (B) 0.1M Tris pH7, 20% w/v PEG1000. The drops were set with 1:1 protein to condition volume ratio, and incubated at 20°C.





**Figure 21. DNA crystals obtained from Ku70-DNA complex crystallization.** In a different DNA substrate, the 14bp duplex DNA with thymine overhangs crystallized on its own. (A) shows crystals of 14b3T and (B) that of 14b1T.



**Figure 22. Diffraction patterns of the initial hits of Ku70-DNA.** The diffraction pattern shows low resolution diffraction spots close to the direct X-ray beam. The red ring represents the resolution range for  $\sim 8\text{\AA}$  diffraction. Both crystals exhibited low resolution diffraction  $\sim 15\text{\AA}$ . (A) Ku70-DNA was crystallized in 0.2M sodium chloride, 0.1M HEPES pH7.5, 25% w/v PEG3350, and (B) 0.1M Tris pH7, 20% w/v PEG1000.

**Table 3. Components of promising Ku70-DNA crystallization conditions.**

<b>Name</b>	<b>Salt</b>	<b>Buffer</b>	<b>Precipitant</b>
MCSG I #95	0.2M sodium chloride	0.1M HEPES pH7.5	25% w/v PEG3350
MCSG II #80		0.1M Tris pH7	20% w/v PEG1000
Wizard I #10		0.1M Tris pH7	20% w/v PEG2000 MME

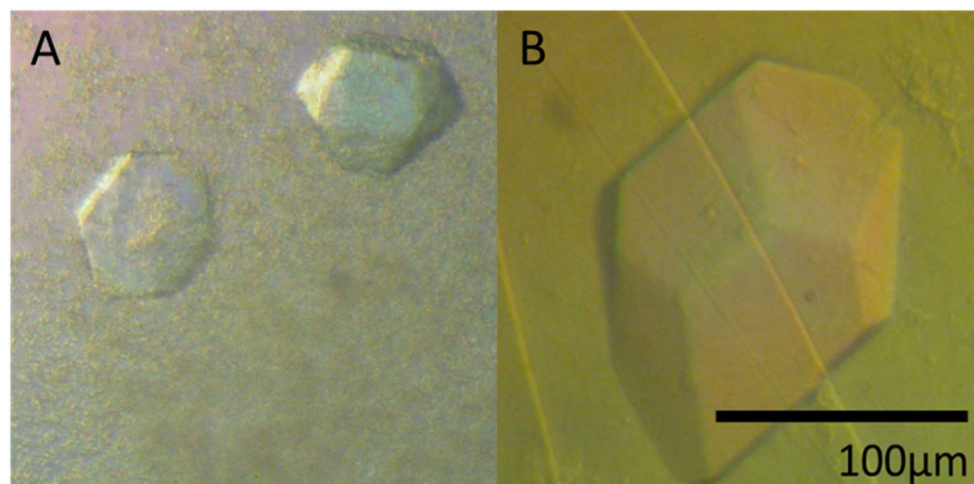
### 3.7 Optimization of Ku70-DNA Crystals Improved Diffraction Data Quality

After identifying several promising lead crystallization conditions for Ku70-DNA complex, further optimization was performed using additive screening and volume ratio variation. From this analysis, MCSG II #80 and MCSG I #95 were shown to be the most promising conditions and both conditions were optimized further in parallel.

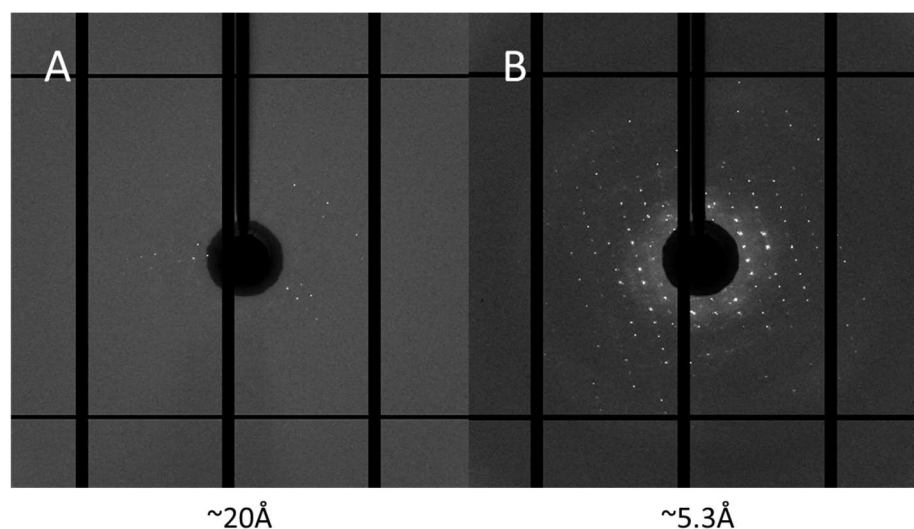
Preliminary optimization yielded inconsistent nucleation events and the crystal morphology remained poorly defined. By varying the structures and lengths of the DNA duplex employed in crystallization trials, we found that DNA containing a defined stem-loop structure and varying duplex length resulted in improved crystal growth reproducibility. In particular, loop16 DNA with its increased length of duplex, compared to loop14, resulted in better crystal morphology with reliable reproducibility. Subsequent iterations of crystallization trials focused on additional crystallization parameters including volume ratio, drop dehydration, temperature, small molecule additives, etc. Wizard I #10 and MCSG II #80 conditions contained similar components, and the final best condition was obtained by combining shared components into a single new condition. The optimal crystal condition was selected by varying molecular weight and concentration of PEG. Subsequent additive screening identified sodium iodide as a suitable ionic species able to positively modulate crystal packing and growth. With sodium iodide incorporated into the crystallization condition, X-ray diffraction data quality improved to  $<8\text{\AA}$  resolution (Appendix 7). Unfortunately, X-ray diffraction data collected from this crystal could not be properly scaled and/or integrated, indicating that the crystal lattice was not perfectly aligned. Therefore, further modifications and optimizations were required to overcome this limitation.

Initial attempts to replicate crystals obtained from MCSG I #95 using the composition tables provided by the manufacturer were unsuccessful likely due to alteration in chemical composition as a result of prolonged storage. By varying individual components within the initial condition, crystals were reproduced; albeit, with lower concentration of lower molecular mass PEG. This observation spawned further rounds of component optimization involving individual component variation, DNA ratio, and further additive

screening. Ultimately, crystals were generated that diffracted  $\sim 5.3\text{\AA}$  resolution. Thus, by careful, systematic optimization two initial lead crystallization conditions were substantially improved into a single condition that resulted in full X-ray diffraction data collection (Figure 23 and 24). The optimized Ku70-DNA crystal occupied a similar sized unit cell compared to prior crystals of Ku70/80-DNA complex (Table 5), but crystallized in a different space group. Although X-ray diffraction data obtained from optimized crystals was able to be scaled and integrated, attempts at solving the structure by molecular replacement failed, perhaps due to conformational changes within the Ku70 structure. Further attempts to obtain phase information could incorporate selenomethionine SAD phasing strategies.



**Figure 23. Progression of Ku70-loop16 crystal morphology through iterations of optimization.** The initial hit from MCSG I #95 (left) was optimized through lowering the concentration and molecular weight of the PEG in the crystallization condition. This was followed by raising the DNA-to-protein ratio, and performing additive screening, which ultimately resulted in the optimized crystal (right). Each stage of crystal optimization yielded improvement in crystal growth shape, size, and X-ray diffraction data quality.



**Figure 24. Progression of Ku70-DNA crystal X-ray diffraction data quality.** Comparison of improvement in both diffraction intensity and resolution of data correlated with the improvement in crystal size and morphology, with (A) the initial hit and (B) the optimized crystal.

**Table 4. Optimization of MCSG I #95 for Ku70-loop16.** The list is sectioned by the order of optimization focus, each varying one aspect of the crystallization environment. Resolution is the average of multiple crystals examined by X-ray diffraction. The bolded condition is carried forward as the base condition for the next round of optimization.

Condition variance	X-ray screening result (Resolution)
Additive with 0.1M Praseodymium(III) acetate hydrate	Non-diffracting crystal/Salt
Additive with 0.1M Praseodymium(III) acetate hydrate and 0.05M ammonium sulphate	Salt
<b>Lower PEG to 12% PEG2000</b>	<b>Protein (13Å)</b>
Lower PEG to 12% PEG3350	Protein (15Å)
Lower PEG to 15% PEG1000	Protein (19Å)
<b>DNA ratio increase from 1:1.2 to 1:1.5</b>	<b>Protein (8Å)</b>
DNA ratio increase from 1:1.2 to 1:3	Protein (20Å)
5% v/v Ethyl acetate as additive	<b>Protein (~5.3Å)</b>

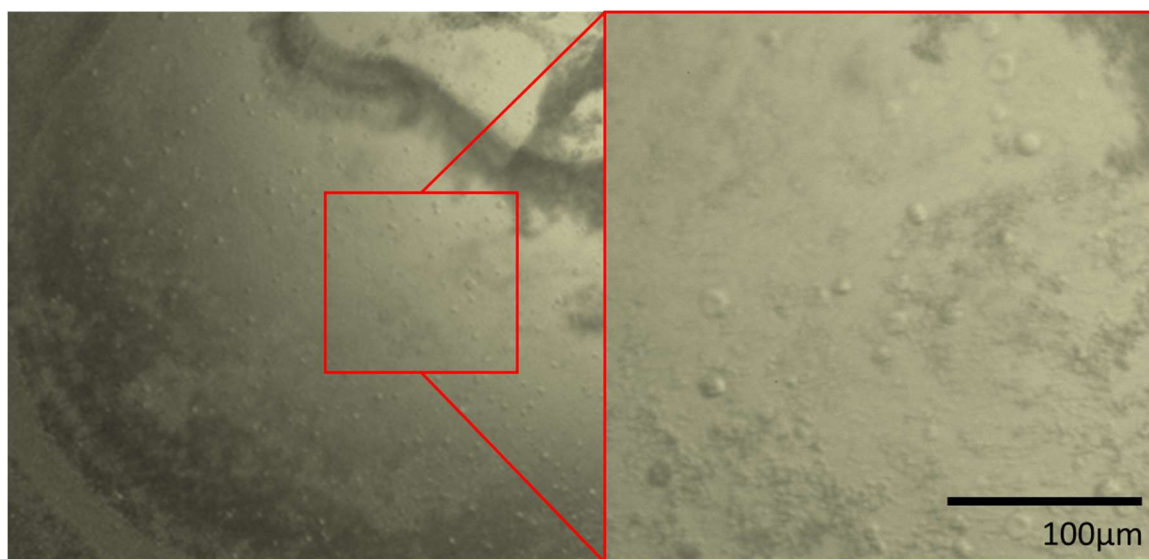
**Table 5. Comparison of Ku70-DNA and Ku70/80-DNA crystal parameters.**

Crystal	Unit cell lengths ( <i>a</i> , <i>b</i> , <i>c</i> ) (Å)	Unit cell angles ( $\alpha$ , $\beta$ , $\gamma$ ) (°)	Space group
Ku70-DNA (5.3Å)	115.92, 115.92, 141.07	90, 90, 120	P6 <sub>1</sub> 22
1JEY (2.5Å)	91.07, 91.07, 152.84	90, 90, 90	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>

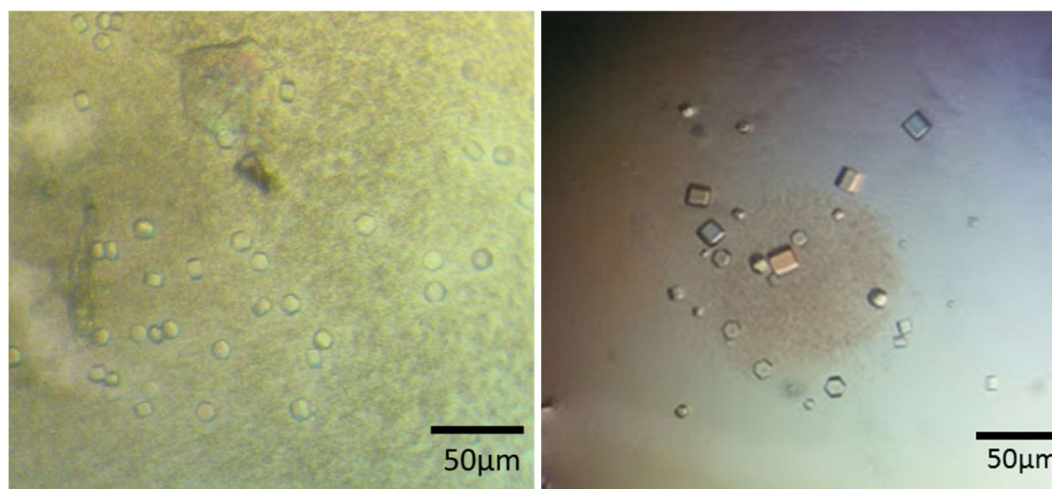
### 3.8 Crystallization of Ku70-DNA-PAXX Generated PAXX Crystals

Initial crystallization hits obtained from sparse matrix screening of Ku70-DNA-PAXX complex resulted in microcrystals that were determined to be composed of protein rather than salt using X-ray diffraction scanning (Figure 25, Appendix 8). Initial hits were optimized for increased crystal size using additive screening and volume ratio variation. Optimization generated numerous, well-shaped crystals (Figure 26 and 27, Table 6) that diffracted to relatively high resolution ( $\sim 5.3\text{\AA}$ ). Unexpectedly, when X-ray diffraction data was collected and analysed for these crystals, the unit cell parameters were found to be too small to contain a complete Ku70-DNA-PAXX complex within the asymmetric unit. Furthermore, the space group was found to be very similar to that of truncated PAXX from PDB 3WTF (Table 7), suggesting that PAXX alone had crystallized under the conditions used. This suggests that while EMSA analysis showed that PAXX could supershift a Ku-DNA complex (74), in crystallization trials, PAXX was dissociated from Ku70-DNA complex. In the supersaturated state, PAXX crystallization was apparently favoured over its interaction with the Ku70-DNA complex in less concentrated conditions. Nevertheless, since the conditions used for full length PAXX crystallization were different from those used to determine the structure of the N-terminal PAXX domain (PDB 3WTF), X-ray data was collected and the structure determined by molecular replacement. The resulting model of PAXX was refined and compared to the deposited structure (see section 3.9). Due to the high propensity of PAXX for crystallization, later trials to obtain crystals of Ku70-DNA-PAXX were performed using longer DNA substrates that enhance interaction of Ku70-DNA and PAXX. At the point of completing this thesis no promising conditions had been identified for Ku70-DNA-PAXX complex using any of the DNA substrates tested.





**Figure 25. Initial hit of Ku70-DNA-PAXX complex.** Similar microcrystals were obtained from several conditions. All crystals exhibited a 3D hexagonal shape with well-defined edges and faces. Image quality was limited by the magnification power of the microscope and mounted imaging system.

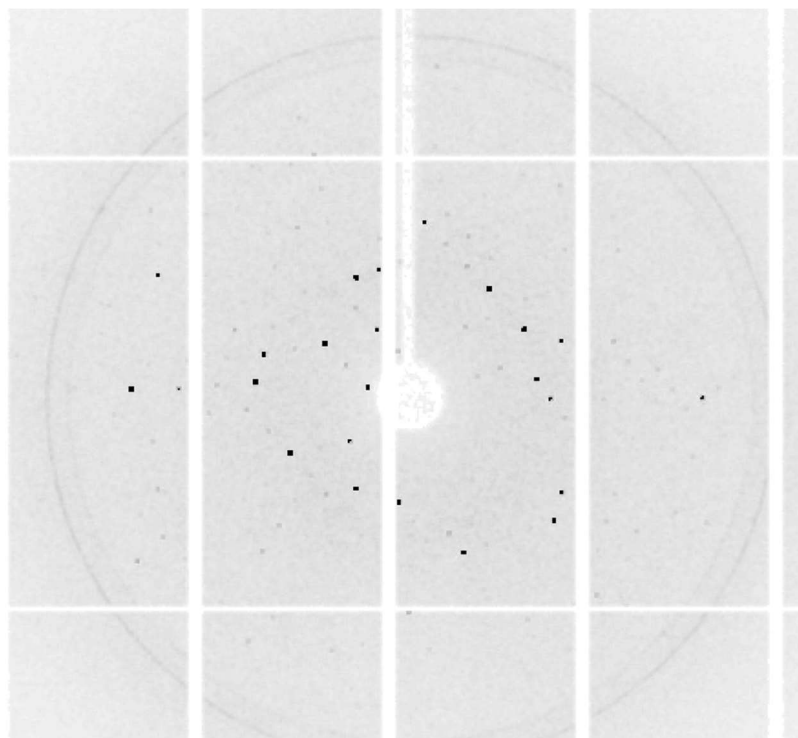


**Figure 26. Images of PAXX crystals in two different crystallization conditions.** For complex mixtures of Ku70-DNA-PAXX, Ku70 and the DNA substrate were mixed in a molar ratio of 1:1.5, then PAXX was added in equimolar ratio to Ku70. (A) shows the typical hexagonal crystal morphology from Ku70-DNA-PAXX sample with mother liquor containing 0.1M Bis-Tris Propane pH7, 1M ammonium citrate tribasic pH7. (B) Optimized PAXX crystals used for data collection diffracted to 3-3.5Å resolution.



**Table 6. Summary of PAXX Crystallization and X-ray diffraction.** Crystals of sufficient size were screened using X-ray diffraction and the crystal contents were identified. The bolded condition yielded the best initial hit for subsequent optimization by equilibrating with 1.1M ammonium acetate instead of the usual 1.5M.

Crystallization condition	DNA substrate	Crystal identity (Resolution)
0.02M calcium chloride, 0.1M sodium acetate pH4.6, 30% v/v MPD	Loop16	Salt
0.1M sodium chloride, 0.1M Bis-Tris pH6.5, 1.5M Ammonium Sulphate, 0.1M Strontium chloride hexahydrate	Loop16	Salt
0.2M magnesium chloride, 0.1M HEPES pH7.5, 30% v/v PPG P400	Loop16	Salt
0.2M magnesium chloride, 20% w/v PEG3350	Loop16	Salt
0.2M lithium acetate, 20% w/v PEG3350	Loop16	Protein (25Å)
0.1M HEPES pH7.5, 10% w/v PEG8000, 8% v/v ethylene glycol	Loop16	Protein (19Å)
1M lithium chloride, 30% w/v PEG6000, 0.1M sodium acetate	Loop16	Protein (15Å)
0.1M Tris pH8.5, 1.4M ammonium tartrate dibasic	Loop16	Protein (10Å)
<b>0.1M Bis-Tris Propane pH7, 1M ammonium citrate tribasic pH7</b>	<b>Loop16</b>	<b>Protein (5.5Å)</b>
Optimized with lower dehydration and longer time	Loop16	Protein (3-3.5Å)
0.2M potassium acetate pH4.5, PEG 400, 550 MME, 600, 1K, 2K, 3350, 4K, 5K MME, 6K, 8K and 10K	14b10T	Non-diffracting crystal
20% v/v 2-Propanol, 0.1 M MES monohydrate pH 6.0, 20% w/v PEG2000	14b10T	Non-diffracting crystal



**Figure 27. Single image of X-ray diffraction pattern from PAXX data collection.** X-ray diffraction data reached 3.5Å resolution. The experimentally indexed space group and unit cell closely resembles that of PDB 3WTF (Table 7), suggesting the crystal unit cell contained a single PAXX homodimer.

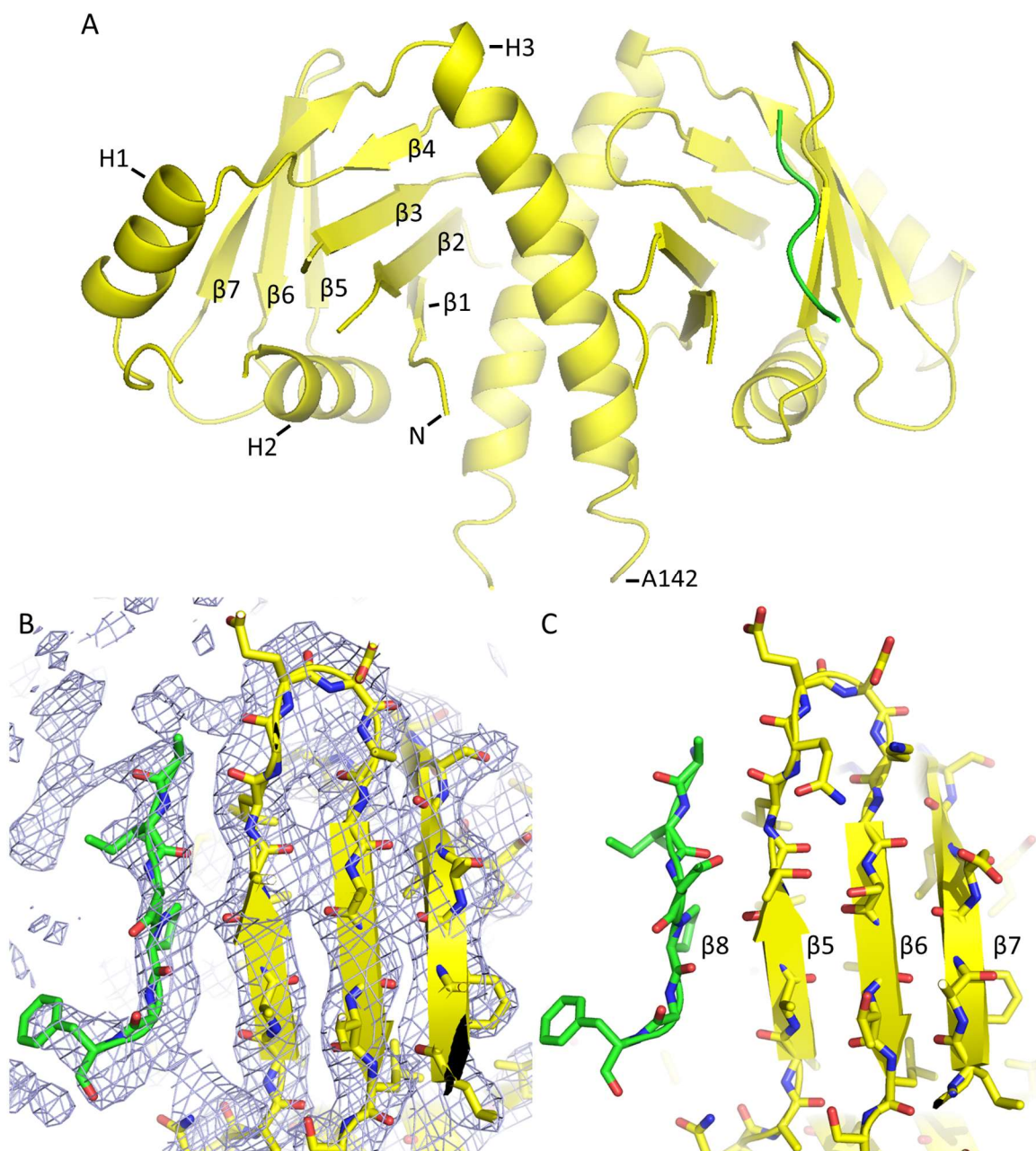
**Table 7. Comparison of experimental and referenced PAXX crystal parameters.** The unit cell parameters and space group determined are listed.

Crystal	Unit cell lengths ( <i>a</i> , <i>b</i> , <i>c</i> ) (Å)	Unit cell angles ( $\alpha$ , $\beta$ , $\gamma$ ) (°)	Space group
PAXX (3Å)	90.99, 90.99, 152.82	90, 90, 120	P6 <sub>1</sub> 22
PAXX (3.5Å)	91.07, 91.07, 152.84	90, 90, 120	P6 <sub>1</sub> 22
3WTF (3.45Å)	91.95, 91.95, 153.2	$\beta=120$	P6 <sub>5</sub> 22

### 3.9 PAXX Structure Determined from Full-length PAXX

X-ray diffraction data from crystals of full-length PAXX were collected to 3.5 Å at the Advanced Photo Source Synchrotron, Argonne National Labs. The structure of full-length PAXX was determined by molecular replacement using the N-terminal domain of PAXX as the search model (PDB 3WTF). The final model was refined to  $R_{\text{work}}$  and  $R_{\text{free}}$  values of ~0.26 and ~0.30, respectively (Figure 28, Table 7). Comparison of the full-length structure of PAXX to the N-terminal domain of PAXX (3WTF) showed very little difference with both containing a 7-stranded antiparallel beta-sandwich with a helix-turn-helix motif inserted between strands 4 and 5. As well, both had almost identical amounts of coverage for the extended C-terminal helices that forms the dimerization domain and extends away from the N-terminal head domain. The fact that both structures were similar despite the current structure containing full-length protein, suggests that the C-terminal region of PAXX remains highly flexible in the absence of DNA and/or Ku70.

The full-length structure did however, contain some additional coverage compared to the N-terminal structure of PAXX. Interestingly, the determined PAXX structure contains an additional  $\beta$ -strand (residues L185 – F190) that could be modelled into electron density (Figure 28B). This  $\beta$ -strand stacked adjacent to the outer  $\beta$ -sheet in an anti-parallel arrangement extending the sheet from 3 to 4 strands (Figure 28C). The presence of the C-terminal strand not only indicates that our structure contained full-length protein, but also represents the first evidence that PAXX C-terminal tails fold back to interact with the head domain. This finding is of particular interest since the C-terminal region is known to mediate both DNA and Ku70 interaction. As well, the C-terminus strand folding back to the N-terminal head domain mirrors the structure of XLF, further extending the similarity between these two related NHEJ repair proteins. The presence of additional density in our structure also excludes the possibility that the C-terminus of PAXX was truncated during the crystallization process. Overall, the structure of full-length PAXX recapitulated the features of a previously deposited PAXX N-terminal domain structure, and unveiled additional information about the general structure of PAXX.



**Figure 28.** Highlighted Structural Information of PAXX. (A) An additional  $\beta$ -strand was solved using the experimental data, highlighted in green. The N-terminus, secondary structures, and the C-terminal amino acids are denoted. (B) The electron density map is illustrated showing the modelled  $\beta$ -strand structure fitted in an anti-parallel orientation stacking with the outer  $\beta$ -sheet at the N-terminal head domain of PAXX. (C) The same region without the electron density map. The figures are rendered using PyMOL.

## Chapter 4

### 4 Discussion

Through the efforts in this study, protein yield no longer limits structural studies of NHEJ assemblies. The methods developed here, therefore, provide the grounds upon which future functional characterization and complementary structural analysis can proceed. Such investigation will greatly increase insight into the molecular mechanisms of NHEJ DNA repair and permit effective targeting of this pathway for cancer treatment. Using advances made in protein production, several NHEJ assemblies were chosen for further structural studies.

LX4-XLF-DNA represents the ‘holy grail’ for structural characterization of NHEJ assemblies due to its essential role in mammalian DNA DSB repair. This assembly represents the final ligation complex that completes repair by coordinating re-joining of DNA ends at a DSB site. As a DSB site includes two broken strands, current models speculate that two LX4 complexes participate in repair; however, how LX4 interacts with DNA and carries out ligation is currently unknown (70, 95). Understanding the molecular mechanism driving the ligation complex assembly is vital for clarifying how the complex coordinates DSB repair. Importantly, knowledge of LX4 and related NHEJ complexes may lead to new avenues for chemotherapeutic intervention.

Given recent reports that Ku70 and Ku80 seem to have discrete functions independent of their heterodimeric assembly, understanding how these proteins form complexes on their own with DNA, and binding factors such as PAXX, becomes of great interest. In particular, the Ku70-DNA-PAXX complex is poorly characterized and warrants further investigation. Crystallization of such complexes provides an obvious route for elucidating structure required for understanding molecular mechanism; however, large multi-component assemblies are challenging targets for crystallization given the multitude of factors influencing crystal growth and quality. Work in this thesis focused on pursuit of both LX4- and Ku70-based complexes. Numerous conditions were identified that gave rise to crystals that were of no value (i.e. non-diffracting, unable to be optimized, and/or salt). Nevertheless, by exploring many conditions, microcrystals were identified that

could be optimized. Through careful observation and systematic optimization, a PAXX homodimer structure was solved and refined at 3Å resolution. This structure validated the idea that the C-terminal region of PAXX is highly flexible in the absence of Ku70 and/or DNA, and that at least a small portion of the C-terminal tail folds back to interact with the N-terminal head domain as has been observed for XLF. In addition, two promising conditions were identified for crystallization of Ku70-DNA, with one that permitted X-ray diffraction data to be collected up to 5.3Å resolution. Ultimately, significant progress has been made towards elucidating higher-order structures of NHEJ factors. Based on these findings, further optimization, combined with complementary structural techniques, are expected to permit structural determination of several new NHEJ assemblies.

## 4.1 Overcoming Practical Limits to Studying NHEJ Complex Structures

The preferred expression system for generating protein for X-ray crystallographic studies is indisputably bacterial. This reflects the ease of use, low cost and potential for high yield. Not surprisingly, the vast majority of structures deposited in the protein data bank were determined from protein generated with bacterial expression systems.

Unfortunately, many proteins are not able to be easily expressed in bacteria due to problems associated with low expression, degradation and/or solubility. Heterologous gene expression often fails due to codon-bias, lack of necessary post-translational modifications, missing prosthetic groups, unavailable chaperones, etc. Some of these limitations have been successfully overcome by codon-optimization, use of bacterial strains with altered properties, variation in cell growth temperature and levels of induction, changes in promoters, addition of fusions and altering domain boundaries of target proteins. Currently, determining which (if any) of these approaches result in desirable outcomes must be determined empirically.

In this thesis, the expression and purification of several core NHEJ factors was significantly improved. Guided by empirical evidence, numerous alterations were explored to improve protein yield and purity. Although not discussed in depth, a variety of bacterial strains, induction times, IPTG concentrations and temperatures were evaluated for their impact on expression efficiency and solubility for each protein of

interest. In addition to these considerations, other parameters (buffer components, chromatographic techniques) were explored to maximize yield and purity of the final product. Perhaps the largest gains came from a combination of codon optimization and re-engineering of construct design. For LX4, the use of a dual co-expression vector was particularly important for increasing the overall levels of Ligase IV expression. Ligase IV is not stable in the absence of XRCC4 and therefore having both proteins co-expressed at similar levels ensured maximal recovery of complex. LX4 purity was improved dramatically, simply by extending the His tag from 6 to 9 residues and moving the placement of the tag to the N-terminus of Ligase IV (the limiting component of LX4). Overall, LX4 expression and purification was significantly improved using the newly synthesized construct, with a >10-fold increase in final yield compared to earlier expression strategies. The resulting protein was stable to >15 mg/ml and highly active for intermolecular ligation.

Whether the approaches used to improve LX4 yield will be generally transferrable for heterologous expression of other complexes remains unclear. All of the parameters that were modified for LX4 expression and purification have been reported for other systems. Unfortunately, at this point it is not possible to say with certainty if any particular parameter will result in desired outcomes. However, based on the findings for LX4, and also Ku70/80, it seems that codon optimization, selection of domain boundaries, and placement of purification/solubility fusions are particularly important factors for improving protein yield.

In order to understand the molecular mechanism of NHEJ it is essential to determine structures of higher-order assemblies in the presence of DNA. Thus, an important consideration in crystallographic studies involving NHEJ factors is not just protein, but also DNA. The challenge with nucleoprotein complexes is that the DNA component plays a critical role in crystal lattice packing that can be very difficult to optimize. Frequently, crystals of protein-DNA complexes diffract poorly until an optimal DNA substrate can be determined. For NHEJ studies the consideration is further complicated by the need to stabilize large complexes in a monodisperse state that is essential for crystallization to occur. NHEJ is not DNA sequence-dependent. As such, designing DNA

substrates that remain bound to protein in only one state is challenging. For complexes of Ku70, limiting diffusion along DNA was addressed by use of a stem-loop structure on one end of the substrate. Although the stem-loop with an optimal length of 16bp seemed to work reasonably well for Ku70, more work needs to be done to explore variation in the stem-loop structure itself and perhaps the type of ends (blunt, overlap, etc.). For complexes of LX4, optimization of DNA substrates is critical. Although several attempts were made here to use substrates that would ‘stall’ LX4 activity at a given step in ligation, the overall success was very limited. It will be important to try different DNA substrates that employ variation in length and end-compatibility. Now that ample amounts of LX4, XLF, PAXX, Ku70/80, and Ku70 are readily available, it will be feasible to screen a large variety of DNA substrates for structure determination of various NHEJ assemblies.

## 4.2 Crystallographic Strategies to Facilitate High Quality Diffraction Data

Crystallography represents the most successful method for determining structures of biological macromolecules, with ~90% of all PDB deposited structures determined by X-ray crystallography. Nevertheless, challenges in obtaining well-ordered crystals are frequently encountered. To improve success in obtaining suitable crystals for high resolution data collection, it is important to understand and fully exploit the basic principles of crystal growth.

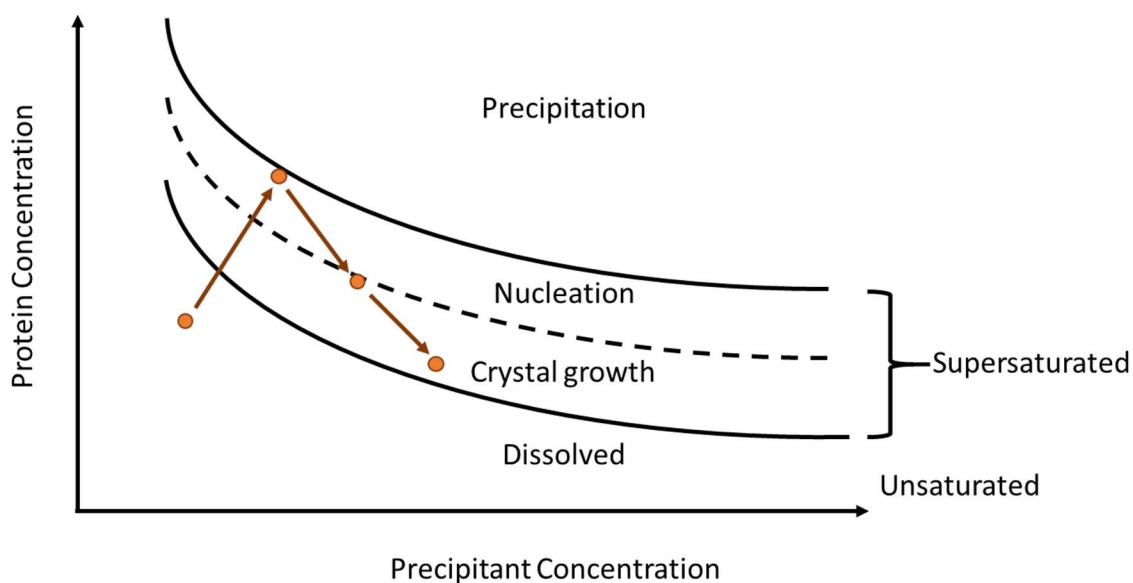
Two major events are required for obtaining crystals: supersaturation and nucleation. Under supersaturating conditions, solutes (proteins) in solution start to self-associate. When a cluster of well-ordered solute molecules reaches a critical size, nucleation of the crystal can occur. Nucleation is energetically unfavourable and considered the limiting, black-box step in crystal formation. Once nucleation occurs, sustained supersaturation of protein is sufficient to drive crystal growth (Figure 29). Despite the principles of crystallization being the same for all solutes, macromolecules present significant challenges compared to their smaller molecule counterparts. In particular, macromolecular crystallography must ensure that the protein is able to remain stable (i.e. toward degradation, oxidation, unfolding and aggregation) over the conditions (time,



temperature, concentration) required to achieve supersaturation. Practically, this requires empirically determining proper conditions (ionic strength, pH, temperature, etc.). During sparse matrix screening, which samples reagent formulations that have previously crystallized a protein and permits rapid and economical coverage of successful crystallization space, many of these parameters are systematically explored and therefore, it is unlikely that simply screening more crystallization conditions would be fruitful once several thousand conditions have been evaluated.

As in the case of Ku70-DNA, once conditions are found that promote stability during supersaturation, crystal hits tend to occur frequently. Instead of further screening for new crystallization conditions, which is unlikely to improve crystal quality, small alterations of individual components in a crystallization condition, small changes to domain boundaries to remove flexible regions and/or changes in the rate and extent of supersaturation and nucleation through micro-seeding, etc. are likely to result in better crystal growth.

Moving forward, it would seem most appropriate to investigate altered DNA structures in combination with micro-seeding and perhaps different methods of crystallization (anaerobic conditions, microbatch and free interface diffusion) other than hanging drop vapour diffusion to help improve diffraction quality of Ku70-DNA crystals.



**Figure 29. Idealized states of solutes.** The curves separating the unsaturated, supersaturated and precipitation zones are idealized for illustration of concept. In the hanging drop crystallization method, the volume of the drop continues to decrease, leading the state of the protein and precipitant to rise, crossing into the supersaturated nucleation zone. Once nucleation occurs, the effective protein concentration will decrease as more protein arranges into the crystal rather than remaining dispersed in the solvent. Meanwhile, precipitant concentration continues to increase as the drop becomes dehydrated by equilibration with the reservoir solution (illustrated by the state shift of the orange dot). Hence, careful control of the rate and extent of dehydration allows the state of the solute to linger within the supersaturated crystal growth zone, facilitating protein crystallization to suitable size for X-ray diffraction.

### 4.3 Methods to Complement Crystal Structure Determination

Current technical limitations prevent visualization of protein interactions at the atomic level in real-time, resulting in a disconnect between functional and mechanistic understanding. Although X-ray crystallography permits visualization of snapshots of proteins and other macromolecules, this is only achieved within a static, crystalline state.

There are numerous techniques that can be employed for structural characterization and interaction of proteins. Nuclear magnetic resonance spectroscopy (NMR) has been extensively used for structure determination of small-to-medium sized proteins in solution. Unfortunately, the sample criteria for NMR is relatively stringent and its use for further studies of larger NHEJ assemblies with DNA is likely to be limited. Similarly, hydrogen-deuterium exchange (HDX) can be a great complement to X-ray crystallography and NMR studies (169). However, this requires prior assignment of the spectrum for the protein which may not be available and, considering the size of the core components in a LX4 complex (DNA Ligase IV, ~104kDa) or Ku70-DNA-PAXX (~70kDa), obtaining the spectrum may not be feasible. Mass spectrometry provides an alternative detection method with much fewer restraints and higher capacity to analyse complexes (170); however, spatial resolution is very poor and data interpretation is frequently very difficult.

SAXS is another emerging, complementary method for characterizing complexes. This method enables investigation of large complexes in solution, bypassing the experimental bottleneck of crystallization; however, the degree of resolution is still very poor. SAXS is incredibly informative for determining the size, shape and distribution of proteins within a sample. As well, shape differences due to complex formation can be determined to infer conformational changes within proteins. Although on its own, SAXS is not overly well-suited to understand molecular structures, it can be useful when combined with other techniques such as X-ray crystallography. SAXS information would greatly assist in the overall understanding of protein-protein and protein-DNA interaction pertaining to LX4 and PAXX complexes, in particular, revealing the overall assembly of LX4 complexes

around the DNA ends, and the conformational changes that allows PAXX to bind Ku-DNA complexes.

The use of cryo-EM has recently received renewed attention due to improved detectors that have resulted in the ability to determine structures of large macromolecular assemblies at near atomic resolution. Unlike NMR, cryo-EM is not limited by sample size. In addition, unlike crystallography, there is no requirement for crystal growth and samples can be imaged in solution in a variety of dynamic states. Cryo-EM allows studies of protein complexes that resist crystallization due to surface flexibility and/or sample heterogeneity due to dynamic intermolecular interactions (171). Cryo-EM imaging can resolve the general structure of protein complexes of over 100kDa, which is the case for both LX4-XLF-DNA and Ku70-DNA-PAXX. Importantly, cryo-EM can provide spatial details about protein complexes, and has the potential to resolve structure at resolutions that rival X-ray crystallography (172). With crystal structures available for domains and truncated complexes of NHEJ proteins, one could use available structures to fit density maps generated through cryo-EM and thereby provide structural insight into larger more dynamic complexes. Given these advantages and the development of modern detectors and image processing capacities, cryo-EM seems perfectly suited for future use in studying large NHEJ assemblies.

## 4.4 Implications for NHEJ Complex Interactions

Although some structural information is available for domains of DNA Ligase IV and segments of XRCC4 and XLF, their molecular mechanism of action in NHEJ DNA repair has yet to be determined. Models of the ligation complex have been suggested based on functional studies, however, these remain highly speculative and further structural studies involving higher-order assemblies with DNA are required.

Work reported here for Ku70 suggests that it forms a stable homodimer in the absence of Ku80. This is particularly interesting since recent studies suggest that Ku70 and Ku80 may function independent of a heterodimer. Our finding that Ku70 forms a stable homodimer agrees with these reports and suggests that Ku80 may also be able to form homodimers. As well, analysis of Ku70 demonstrates its ability to bind DNA at similar

levels to Ku70/80 heterodimer and also interact with PAXX (74). These findings clearly suggest that Ku70 alone can function in the absence of Ku80 and supports findings from prior reports.

Previous studies of Ku70 in Bax-mediated apoptosis have shown a Ku70-dependent competitive interaction with Bax and cyclin E under genotoxic stress (173). Further studies demonstrated that phosphorylation of Ku70, through interaction with cyclin B1-Cdk1 and A2-Cdk2, caused dissociation of the Ku70/80 dimer. This type of structural alteration may reduce participation of Ku70 in DSB repair during S and M phase and allow HR to occur in a cell cycle-dependent manner (174, 175). The finding that Ku70-Bax interaction results in regulated interaction outside the nucleus, potentially uncovering a role of Ku70 in controlling cellular processes related to genome stability, raises many questions (176–178). Importantly, post-translation modification observed in this study demonstrates a means of regulating participation of Ku70 in hetero- and homodimer formation.

The lack of robust evidence for Ku70 functioning outside of a Ku70/80 heterodimer may be due to the strong phenotype Ku70/80 have in NHEJ, as well as limitations in methodological approaches. For instance, recent studies have shown that previous studies overlooked the Ku70/80 RNA binding effect, leading to potential bias in interpreted results since a sizable amount of Ku70 and Ku80 would not have been released to the soluble fraction during sample extraction (50). Further studies are needed to address whether Ku70 can carry out functions, such as its interaction with PAXX, in the absence of Ku80 *in vivo*.

Although our structure of the full-length PAXX only provided new information for a portion of the C-terminal region, observing that a part of the C-terminal region interacts with the head domain strengthens the parallel between PAXX and XLF. The fact that most of the C-terminal region of PAXX remains disordered, further suggests that it may only become ordered during interaction with Ku70 and DNA. PAXX has been shown to have some redundancy with XLF, which is not surprising given their structural similarity. XLF likely occupies the space required for PAXX interaction with Ku. Since XLF has

additional interactions with XRCC4 and LX4 through its C-terminal tail, that do not appear to be shared with PAXX, it may be that PAXX is only required as a backup when XLF fails to function (179). Structural studies of Ku-DNA-PAXX and Ku-DNA-XLF will be essential for addressing these possibilities.

## 4.5 Outstanding Questions for NHEJ in Cancer Treatment

One of the best-established hallmarks of cancer is genomic instability due to loss of DNA repair. In many types of cancer, HR is inactivated through mutation of repair factors such as BRCA1 and BRCA2 (180, 181). In this case, tumour cells become particularly reliant on NHEJ, making NHEJ an attractive pathway for therapeutic intervention. Simultaneous inhibition of NHEJ and treatment with clastogenic agents can greatly favour elimination of tumour cells. Similar strategies (involving PARP1 inhibition in BRCA2 deficient tumour cells) have already shown great promise (180, 182, 183). Therefore, molecular understanding of the NHEJ pathway is critical for further efforts to develop adjuvant cancer therapy targeting the NHEJ repair pathway.

Previous efforts to determine structures of proteins involved in NHEJ have been largely limited to domains or individual proteins. Moving forward, what will be required for both mechanistic understanding and the ability to effectively target the repair pathway is high-resolution structural information for larger NHEJ assemblies. In particular, the general assembly of LX4 around DSB ends in coordination with Ku and other factors has never been observed, hindering our progress of understanding the molecular mechanisms of NHEJ. With the discovery of PAXX, structural complex of PAXX and Ku can reveal important characteristics that allows Ku to interact with multiple targets and coordinate repair efforts. Much of the functional outcome of altering factor participation in repair has been explored, but how does these factors achieve those specific outcomes? What is the molecular mechanisms driving their interactions? And perhaps more importantly for understanding pathological states, how does the absence of a factor impact the overall repair? Given the recent advances in cryo-EM, such efforts are likely to benefit greatly from a combined approach incorporating both X-ray crystallography and cryo-EM.

## References

1. Li W, Xie C, Yang Z, Chen J, Lu NH. Abnormal DNA-PKcs and Ku 70/80 expression may promote malignant pathological processes in gastric carcinoma. *World J Gastroenterol* [Internet]. 2013 Oct 28 [cited 2016 Apr 8];19(40):6894–901. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84887010596&partnerID=tZOtx3y1>
2. Pastwa E, Malinowski M. Non-Homologous DNA End Joining in Anticancer Therapy. *Curr Cancer Drug Targets* [Internet]. 2007;7(3):243–50. Available from: <http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1568-0096&volume=7&issue=3&spage=243>
3. Benzie IF. Evolution of antioxidant defence mechanisms. *Eur J Nutr*. 2000;39(2):53–61.
4. BOREK C, SACHS L. In Vitro Cell Transformation by X-Irradiation. *Nature* [Internet]. 1966 Apr 16;210(5033):276–8. Available from: <http://www.nature.com/doi/10.1038/210276a0>
5. Gomez-Mendoza M, Banyasz A, Douki T, Markovitsi D, Ravanat JL. Direct Oxidative Damage of Naked DNA Generated upon Absorption of UV Radiation by Nucleobases. *J Phys Chem Lett*. 2016;7(19):3945–8.
6. Ward JF. Some Biochemical Consequences of the Spatial Distribution of Ionizing Radiation-Produced Free Radicals. *Radiat Res*. 1981;86(2):185–95.
7. Ward JF. Dna Damage Produced By Ionizing-Radiation in Mammalian-Cells - Identities, Mechanisms of Formation, and Reparability. *Prog Nucleic Acid Res Mol Biol*. 1988;35:95–125.
8. Ward JF, Kuo I. Deoxynucleotides—models for Studying Mechanisms of Strand Breakage in DNA. *Int J Radiat Biol Relat Stud Physics, Chem Med* [Internet]. 1973 Jan 3;23(6):543–57. Available from: <http://www.tandfonline.com/doi/full/10.1080/09553007314550651>
9. Barker S, Weinfeld M, Zheng J, Li L, Murray D. Identification of mammalian proteins cross-linked to DNA by ionizing radiation. *J Biol Chem*. 2005;280(40):33826–38.
10. Henner WD, Rodriguez LO, Hecht SM, Haseltine WA. gamma Ray induced deoxyribonucleic acid strand breaks. 3' Glycolate termini. *J Biol Chem*. 1983;258(2):711–3.
11. Murphy MP. How mitochondria produce reactive oxygen species. *Biochem J* [Internet]. 2009;417(1):1–13. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19061483%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2605959%5Cnhttp://biochemj.org/lookup/doi/10.1042/BJ20081386>
12. Richter C, Park JW, Ames BN. Normal oxidative damage to mitochondrial and

- nuclear DNA is extensive. *Proc Natl Acad Sci U S A* [Internet]. 1988;85(17):6465–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/3413108>
13. Lieber MR, Karanjawala ZE. Ageing, repetitive genomes and DNA damage. *Nat Rev Mol Cell Biol* [Internet]. 2004;5(1):69–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14708011> %5CnC:%5CUsers%5CEugene%5CDocuments%5CAcademic%5CLibrary%5CLieber, M.R. and Karanjawala, Z.E. (2004) *Nat Rev Mol Cell Biol* 5, 69-75.pdf
  14. Lieber MR, Ma YM, Pannicke U, Schwarz K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol*. 2003;4(9):712–20.
  15. Saintigny Y, Delac??te F, Var??s G, Petitot F, Lambert S, Aeverbeck D, et al. Characterization of homologous recombination induced by replication inhibition in mammalian cells. *EMBO J*. 2001;20(14):3861–70.
  16. Hoeijmakers JHJ. DNA damage, aging, and cancer. *N Engl J Med* [Internet]. 2009;361(15):1475–85. Available from: <http://www.nejm.org/doi/full/10.1056/NEJMra0804615>
  17. Gensler HL, Bernstein H. DNA damage as the primary cause of aging. *Q Rev Biol* [Internet]. 1981;56(3):279–303. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7031747>
  18. Li H, Mitchell JR, Hasty P. DNA double-strand breaks: A potential causative factor for mammalian aging? *Mech Ageing Dev* [Internet]. 2008 Jan [cited 2016 Apr 6];129(7–8):416–24. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-45449083227&partnerID=tZOtx3yl>
  19. Hart RW, Setlow RB. Correlation between deoxyribonucleic acid excision-repair and life-span in a number of mammalian species. *Proc Natl Acad Sci USA* [Internet]. 1974;71(6):2169–73. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=388412&tool=pmcentrez&rendertype=abstract> %5Cn<http://www.pnas.org/cgi/doi/10.1073/pnas.71.6.2169>
  20. Wang HH, Perrault AR, Takeda Y, Qin W, Wang HH, Iliakis G. Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res*. 2003;31(18):5377–88.
  21. Sun H, Treco D, Schultes NP, Szostak JW. Double-strand breaks at an initiation site for meiotic gene conversion. Vol. 338, *Nature*. 1989. p. 87–90.
  22. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. *Cell*. 2002;109(2 SUPPL. 1):45–55.
  23. Escribano-Díaz C, Orthwein A, Fradet-Turcotte A, Xing M, Young JTF, Tkáč J, et al. A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. *Mol Cell*. 2013;49(5):872–83.



24. Grzegorz I, Pelliccioli A, Balijja A, Wang X, Fiorani S, Carotenuto W, et al. DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature*. 2004;431(October):1011–7.
25. Mao Z, Bozzella M, Seluanov A, Gorbunova V. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)* [Internet]. 2008 Oct 1;7(10):1765–71. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3158983&tool=pmcentrez&rendertype=abstract>
26. Lavin M, Kozlov S, Gatei M, Kijas A. ATM-Dependent Phosphorylation of All Three Members of the MRN Complex: From Sensor to Adaptor. *Biomolecules* [Internet]. 2015;5(4):2877–902. Available from: <http://www.mdpi.com/2218-273X/5/4/2877/>
27. Caron P, Choudjaye J, Clouaire T, Bugler B, Daburon V, Aguirrebengoa M, et al. Non-redundant Functions of ATM and DNA-PKcs in Response to DNA Double-Strand Breaks. *Cell Rep*. 2015;13(8):1598–609.
28. Jazayeri A, Falck J, Lukas C, Bartek J, Smith G, Lukas J, et al. ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol* [Internet]. 2006;8(1):37–45. Available from: [papers3://publication/uuid/935B758D-D4E0-412A-BE9E-D36108C332DC](http://www.ncbi.nlm.nih.gov/pubmed/16181111)
29. Beucher A, Birraux J, Tchouandong L, Barton O, Shibata A, Conrad S, et al. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *EMBO J* [Internet]. 2009;28(21):3413–27. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2752027&tool=pmcentrez&rendertype=abstract>
30. Gatei M, Sloper K, Sørensen C, Syljuäsen R, Falck J, Hobson K, et al. Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent phosphorylation of Chk1 on Ser-317 in response to ionizing radiation. *J Biol Chem*. 2003;278(17):14806–11.
31. Banin S, Moyal L, Shieh S-Y, Taya Y, Anderson CW, Chessa L, et al. Enhanced Phosphorylation of p53 by ATM in Response to DNA Damage. *Science* (80- ). 1998;281(5383):1674–7.
32. Bolderson E, Richard DJ, Zhou B-BS, Khanna KK. Recent advances in cancer therapy targeting proteins involved in DNA double-strand break repair. *Clin Cancer Res* [Internet]. 2009;15(20):6314–20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19808869>
33. Toshiyuki M, Reed JC. Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell*. 1995;80(2):293–9.
34. Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell* [Internet]. 2017;66(6):801–17. Available from: <http://dx.doi.org/10.1016/j.molcel.2017.05.015>
35. Luo Q, Beaver JM, Liu Y, Zhang Z. Dynamics of p53: A master decider of cell fate. *Genes (Basel)*. 2017;8(2).

36. Johnson RD. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J* [Internet]. 2000;19(13):3398–407. Available from: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/19.13.3398>
37. Krogh BO, Symington LS. Recombination proteins in yeast. *Annu Rev Genet*. 2004;38:233–71.
38. Heyer W-D, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* [Internet]. 2010;44:113–39. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20690856>  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4114321>
39. Mari P-O, Florea BI, Persengiev SP, Verkaik NS, Bruggenwirth HT, Modesti M, et al. Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proc Natl Acad Sci* [Internet]. 2006 Dec 5 [cited 2016 Apr 8];103(49):18597–602. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1693708&tool=pmcentrez&rendertype=abstract>
40. Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. Elseviers [Internet]. 2011;711(1–2):61–72. Available from: <http://dx.doi.org/10.1016/j.mrfmmm.2011.02.005>
41. Fattah F, Lee E, Weisensel N, Wang Y, Lichter N, Hendrickson E. Ku Regulates the Non-Homologous End Joining Pathway Choice of DNA Double-Strand Break Repair in Human Somatic Cells. *PLoS Genet* [Internet]. 2010 Feb [cited 2016 Apr 8];6(2):e1000855. Available from: [citeulike-article-id:8450638%5Chttp://dx.doi.org/10.1371/journal.pgen.1000855](http://dx.doi.org/10.1371/journal.pgen.1000855)
42. Bertocci B, De Smet A, Weill JC, Reynaud CA. Nonoverlapping Functions of DNA Polymerases Mu, Lambda, and Terminal Deoxynucleotidyltransferase during Immunoglobulin V(D)J Recombination In Vivo. *Immunity*. 2006;25(1):31–41.
43. Doherty AJ, Jackson SP, Weller GR. Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett*. 2001;500(3):186–8.
44. Zhang Y, Hefferin ML, Chen L, Shim EY, Tseng H-M, Kwon Y, et al. Role of Dnl4-Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination. *Nat Struct Mol Biol* [Internet]. 2007;14(7):639–46. Available from: <http://dx.doi.org/10.1038/nsmb1261>
45. Chen X, Tomkinson AE. Yeast Nej1 is a key participant in the initial end binding and final ligation steps of nonhomologous end joining. *J Biol Chem*. 2011;286(6):4931–40.
46. Yang H, Matsumoto Y, Trujillo KM, Lees-Miller SP, Osley MA, Tomkinson AE. Role of the yeast DNA repair protein Nej1 in end processing during the repair of DNA double strand breaks by non-homologous end joining. *DNA Repair (Amst)* [Internet]. 2015;31:1–10. Available from: <http://dx.doi.org/10.1016/j.dnarep.2015.04.003>

47. Lieber MR, Gu J, Lu H, Shimazaki N, Tsai AG. Nonhomologous DNA end joining (NHEJ) and chromosomal translocations in humans. *Subcell Biochem.* 2010;50:279–96.
48. Lieber MR. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu Rev Biochem* [Internet]. 2010;79(1):181–211. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3079308&tool=pmcentrez&rendertype=abstract>
49. Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* [Internet]. 2001 Aug 9;412(6847):607–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11493912>
50. Britton S, Coates J, Jackson SP. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J Cell Biol.* 2013;202(3):579–95.
51. Yoo S, Dynan WS. Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic Acids Res* [Internet]. 1999;27(24):4679–86. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=148766&tool=pmcentrez&rendertype=abstract>
52. DeFazio LG, Stansel RM, Griffith JD, Chu G. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J.* 2002;21(12):3192–200.
53. Dynan WS, Yoo S. Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.* 1998;26(7):1551–9.
54. Hudson JJR, Hsu DW, Guo K, Zhukovskaya N, Liu PH, Williams JG, et al. DNA-PKcs-dependent signaling of DNA damage in *Dictyostelium discoideum*. *Curr Biol.* 2005;15(20):1880–5.
55. Kirchgessner C, Patil C, Evans J, Cuomo C, Fried L, Carter T, et al. DNA-dependent kinase (p350) as a candidate gene for the murine SCID defect. *Science* (80- ) [Internet]. 1995;267(5201):1178–83. Available from: <http://science.sciencemag.org/content/267/5201/1178.abstract>
56. Shao L, Goronzy JJ, Weyand CM. DNA-dependent protein kinase catalytic subunit mediates T-cell loss in rheumatoid arthritis. *EMBO Mol Med* [Internet]. 2010 Oct [cited 2016 Apr 8];2(10):415–27. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-79952115866&partnerID=tZOtx3y1>
57. Pitcher RS, Wilson TE, Doherty AJ. New insights into NHEJ repair processes in prokaryotes. *Cell Cycle.* 2005;4(5):675–8.
58. Aravind L, Koonin E V. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res.* 2001;11(8):1365–74.

59. Chan DW, Lees-Miller SP. The DNA-dependent protein kinase is inactivated by autophosphorylation of the catalytic subunit. *J Biol Chem*. 1996;271(15):8936–41.
60. Reynolds P, Anderson JA, Harper J V., Hill MA, Botchway SW, Parker AW, et al. The dynamics of Ku70/80 and DNA-PKcs at DSBs induced by ionizing radiation is dependent on the complexity of damage. *Nucleic Acids Res* [Internet]. 2012 Nov [cited 2016 Apr 8];40(21):10821–31. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84870587195&partnerID=tZOtx3y1>
61. Sharif H, Li Y, Dong Y, Dong L, Wang WL, Mao Y, et al. Cryo-EM structure of the DNA-PK holoenzyme. *Proc Natl Acad Sci* [Internet]. 2017;(10):201707386. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1707386114>
62. Davis AJ, Chen BPC, Chen DJ. DNA-PK: A dynamic enzyme in a versatile DSB repair pathway. *DNA Repair (Amst)* [Internet]. 2014 May [cited 2016 Apr 8];17:21–9. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84900458715&partnerID=tZOtx3y1>
63. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, et al. Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J Biol Chem* [Internet]. 2010 Jan 8 [cited 2016 Mar 1];285(2):1414–23. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-74049134920&partnerID=tZOtx3y1>
64. Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* [Internet]. 2002;108(6):781–94. Available from: <http://www.sciencedirect.com/science/article/pii/S0092867402006712>
65. Ma Y, Schwarz K, Lieber MR. The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst)*. 2005;4(7):845–51.
66. Lee JW, Blanco L, Zhou T, Garcia-Diaz M, Bebenek K, Kunkel TA, et al. Implication of DNA Polymerase  $\lambda$  in Alignment-based Gap Filling for Nonhomologous DNA End Joining in Human Nuclear Extracts. *J Biol Chem*. 2004;279(1):805–11.
67. Mahajan KN, Nick McElhinny S a, Mitchell BS, Ramsden D a. Association of DNA polymerase  $\mu$  (pol  $\mu$ ) with Ku and ligase IV: role for pol  $\mu$  in end-joining double-strand break repair. *Mol Cell Biol* [Internet]. 2002;22(14):5194–202. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12077346> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC139779>
68. Andres SN, Modesti M, Tsai CJ, Chu G, Junop MS. Crystal Structure of Human XLF: A Twist in Nonhomologous DNA End-Joining. *Mol Cell* [Internet]. 2007;28(6):1093–101. Available from: <http://dx.doi.org/10.1016/j.molcel.2007.10.024>

69. Buck D, Malivert L, De Chasseval R, Barraud A, Fondanèche MC, Sanal O, et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell*. 2006;124(2):287–99.
70. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA. XLF Regulates Filament Architecture of the XRCC4-Ligase IV Complex. *Structure* [Internet]. 2010;18(11):1431–42. Available from: <http://dx.doi.org/10.1016/j.str.2010.09.009>
71. Gell D, Jackson SP. Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic Acids Res* [Internet]. 1999;27(17):3494–502. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=148593&tool=pmcentrez&rendertype=abstract>
72. Costantini S, Woodbine L, Andreoli L, Jeggo PA, Vindigni A. Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA Repair (Amst)*. 2007;6(6):712–22.
73. Yano K, Morotomi-Yano K, Wang S-Y, Uematsu N, Lee K-J, Asaithamby A, et al. Ku recruits XLF to DNA double-strand breaks. *EMBO Rep* [Internet]. 2008;9(1):91–6. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2246615&tool=pmcentrez&rendertype=abstract>
74. Tadi SK, Tellier-Lebègue C, Nemoz C, Drevet P, Audebert S, Roy S, et al. PAXX Is an Accessory c-NHEJ Factor that Associates with Ku70 and Has Overlapping Functions with XLF. *Cell Rep* [Internet]. 2016;17(2):541–55. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2211124716312487>
75. Simsek D, Jasin M. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4/ligase IV during chromosomal translocation formation. *Nat Struct Mol Biol* [Internet]. 2010;17(4):410–416. Available from: <http://dx.doi.org/10.1038/nsmb.1773>
76. Simsek D, Brunet E, Wong SYW, Katyal S, Gao Y, McKinnon PJ, et al. DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet*. 2011;7(6):1–11.
77. Della-Maria J, Zhou Y, Tsai MS, Kuhnlein J, Carney JP, Paull TT, et al. Human Mre11/human Rad50/Nbs1 and DNA ligase III $\alpha$ /XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway. *J Biol Chem*. 2011;286(39):33845–53.
78. Dibiase SJ, Zeng Z, Chen R, Hyslop T, Curran WJ, Iliakis G. DNA-dependent Protein Kinase Stimulates an Independently DNA-dependent Protein Kinase Stimulates an Independently Active ,. *Cancer Res*. 2000;60(1):1245–53.
79. Oh S, Harvey A, Zimbric J, Wang Y, Nguyen T, Jackson PJ, et al. DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells. *DNA Repair (Amst)* [Internet]. 2014;21:97–110. Available from: <http://dx.doi.org/10.1016/j.dnarep.2014.04.015>
80. Bentley J, Diggle CP, Harnden P, Knowles MA, Kiltie AE. DNA double strand

- break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Res.* 2004;32(17):5249–59.
81. Riballo E, Doherty AJ, Dai Y, Stiff T, Oettinger MA, Jeggo PA, et al. Cellular and Biochemical Impact of a Mutation in DNA Ligase IV Conferring Clinical Radiosensitivity. *J Biol Chem.* 2001;276(33):31124–32.
  82. Riballo E, Critchlow SE, Teo SH, Doherty AJ, Priestley A, Broughton B, et al. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Curr Biol.* 1999;9(13):699–702.
  83. Buck D, Moshous D, de Chasseval R, Ma Y, le Deist F, Cavazzana-Calvo M, et al. Severe combined immunodeficiency and microcephaly in siblings with hypomorphic mutations in DNA ligase IV. *Eur J Immunol [Internet].* 2006 Jan [cited 2016 Apr 8];36(1):224–35. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-30944455282&partnerID=tZOTx3y1>
  84. O'Driscoll M, Cerosaletti KM, Girard PM, Dai Y, Stumm M, Kysela B, et al. DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. *Mol Cell.* 2001;8(6):1175–85.
  85. Kurosawa A, Saito S, So S, Hashimoto M, Iwabuchi K, Watabe H, et al. DNA Ligase IV and Artemis Act Cooperatively to Suppress Homologous Recombination in Human Cells: Implications for DNA Double-Strand Break Repair. *PLoS One.* 2013;8(8).
  86. Bryans M, Valenzano MC, Stamato TD. Absence of DNA ligase IV protein in XR-1 cells: Evidence for stabilization by XRCC4. *Mutat Res - DNA Repair.* 1999;433(1):53–8.
  87. Herrmann G, Lindahl T, Schär P. *Saccharomyces cerevisiae* LIF1: a function involved in DNA double-strand break repair related to mammalian XRCC4. *EMBO J.* 1998;17(14):4188–98.
  88. Gao Y, Sun Y, Frank KM, Dikkes P, Fujiwara Y, Seidl KJ, et al. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell.* 1998;95(7):891–902.
  89. Ochi T, Gu X, Blundell TL. Structure of the catalytic region of DNA ligase IV in complex with an artemis fragment sheds light on double-strand break repair. *Structure [Internet].* 2013;21(4):672–9. Available from: <http://dx.doi.org/10.1016/j.str.2013.02.014>
  90. Ochi T, Wu Q, Chirgadze DY, Grossmann JG, Bolanos-Garcia VM, Blundell TL. Structural insights into the role of domain flexibility in human DNA ligase IV. *Structure.* 2012;20(7):1212–22.
  91. Liu S, Liu X, Kamdar RP, Wanotayan R, Sharma MK, Adachi N, et al. C-terminal region of DNA ligase IV drives XRCC4/DNA ligase IV complex to chromatin. *Biochem Biophys Res Commun [Internet].* 2013;439(2):173–8. Available from: <http://dx.doi.org/10.1016/j.bbrc.2013.08.068>

92. Sibanda BL, Critchlow SE, Begun J, Pei XY, Jackson SP, Blundell TL, et al. Crystal structure of an Xrcc4-DNA ligase IV complex. *Nat Struct Biol*. 2001;8(12):1015–9.
93. Cotner-Gohara E, Kim IK, Hammel M, Tainer JA, Tomkinson AE, Ellenberger T. Human DNA ligase III recognizes DNA ends by dynamic switching between two DNA-bound states. *Biochemistry*. 2010;49(29):6165–76.
94. Pascal JM, Tsodikov O V., Hura GL, Song W, Cotner EA, Classen S, et al. A Flexible Interface between DNA Ligase and PCNA Supports Conformational Switching and Efficient Ligation of DNA. *Mol Cell*. 2006;24(2):279–91.
95. Waters CA, Strande NT, Wyatt DW, Pryor JM, Ramsden DA. Nonhomologous end joining: A good solution for bad ends. *DNA Repair (Amst)* [Internet]. 2014;17:39–51. Available from: <http://dx.doi.org/10.1016/j.dnarep.2014.02.008>
96. Jayaram S, Ketner G, Adachi N, Hanakahi LA. Loss of DNA ligase IV prevents recognition of DNA by double-strand break repair proteins XRCC4 and XLF. *Nucleic Acids Res*. 2008;36(18):5773–86.
97. Brouwer I, Sitters G, Candelli A, Heerema SJ, Heller I, Melo de AJ, et al. Sliding sleeves of XRCC4–XLF bridge DNA and connect fragments of broken DNA. *Nature* [Internet]. 2016;535(7613):566–9. Available from: <http://www.nature.com/doifinder/10.1038/nature18643>
98. Yano KI, Chen DJ. Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end-joining pathway. *Cell Cycle*. 2008;7(10):1321–5.
99. Grawunder U, Wilm M, Wu X, Kulesza P, Wilson TE, Mann M, et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* [Internet]. 1997;388(6641):492–5. Available from: <http://dx.doi.org/10.1038/41358>
100. Modesti M, Hesse JE, Gellert M. DNA binding of Xrcc4 protein is associated with V(D)J recombination but not with stimulation of DNA ligase IV activity. *EMBO J* [Internet]. 1999;18(7):2008–18. Available from: <http://dx.doi.org/10.1093/emboj/18.7.2008>
101. Dai Y, Kysela B, Hanakahi L a, Manolis K, Riballo E, Stumm M, et al. Nonhomologous end joining and V(D)J recombination require an additional factor. *Proc Natl Acad Sci U S A*. 2003;100(5):2462–7.
102. Callebaut I, Malivert L, Fischer A, Mornon JP, Revy P, De Villartay JP. Cernunnos interacts with the XRCC4??DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. *J Biol Chem* [Internet]. 2006;281(20):13857–60. Available from: <http://www.jbc.org/content/281/20/13857.short>
103. Hentges P, Ahnesorg P, Pitcher RS, Bruce CK, Kysela B, Green AJ, et al. Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos. *J Biol Chem* [Internet]. 2006 Dec 8 [cited 2016 Apr 8];281(49):37517–26. Available from:

<http://www.scopus.com/inward/record.url?eid=2-s2.0-33846002032&partnerID=tZOtx3y1>

104. Tsai CJ, Kim S a, Chu G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proc Natl Acad Sci U S A* [Internet]. 2007;104(19):7851–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17470781>  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1859989>
105. Zha S, Guo C, Boboila C, Oksenysh V, Cheng H-L, Zhang Y, et al. ATM damage response and XLF repair factor are functionally redundant in joining DNA breaks. *Nature* [Internet]. 2011;469(7329):250–4. Available from: <http://dx.doi.org/10.1038/nature09604>
106. Deshpande RA, Wilson TE. Modes of interaction among yeast Nej1, Lif1 and Dnl4 proteins and comparison to human XLF, XRCC4 and Lig4. *DNA Repair (Amst)*. 2007;6(10):1507–16.
107. Lu H, Pannicke U, Schwarz K, Lieber MR. Length-dependent binding of human XLF to DNA and stimulation of XRCC4·DNA ligase IV activity. *J Biol Chem* [Internet]. 2007;282(15):11155–62. Available from: <http://dx.doi.org/10.1074/jbc.M609904200>
108. Yano KI, Morotomi-Yano K, Lee KJ, Chen DJ. Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF. *FEBS Lett* [Internet]. 2011;585(6):841–6. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3066473&tool=pmc.ncbi&rendertype=abstract>
109. Mahaney BL, Hammel M, Meek K, Tainer JA, Lees-Miller SP. XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair. *Natl Institutes Heal*. 2012;29(6):997–1003.
110. Andres SN, Junop MS. Crystallization and preliminary X-ray diffraction analysis of the human XRCC4-XLF complex. *Acta Crystallogr Sect F Struct Biol Cryst Commun* [Internet]. 2011;67(Pt 11):1399–402. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22102241>
111. Hammel M, Rey M, Yu Y, Mani RS, Classen S, Liu M, et al. XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair. *J Biol Chem* [Internet]. 2011;286(37):32638–50. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3173232&tool=pmc.ncbi&rendertype=abstract>
112. Ahnesorg P, Smith P, Jackson SP. XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining. *Cell* [Internet]. 2006;124(2):301–13. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0092867406000031>
113. Mimori T, Akizuki M, Yamagata H, Inada S, Yoshida S, Homma M.



- Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *J Clin Invest*. 1981;68(3):611–20.
114. Strande N, Roberts SA, Oh S, Hendrickson EA, Ramsden DA. Specificity of the dRP/AP lyase of ku promotes nonhomologous end joining (NHEJ) fidelity at damaged ends. *J Biol Chem*. 2012;287(17):13686–93.
  115. Weller GR, Kysela B, Roy R, Tonkin LM, Scanlan E, Della M, et al. Identification of a DNA Nonhomologous End-Joining Complex in Bacteria. *Science* (80- ) [Internet]. 2002;297(2002):1686–9. Available from: [papers3://publication/uuid/6F87A620-FE73-4F92-84B1-3EE8A866DCAF](http://papers3://publication/uuid/6F87A620-FE73-4F92-84B1-3EE8A866DCAF)
  116. Gong C, Bongiorno P, Martins A, Stephanou NC, Zhu H, Shuman S, et al. Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat Struct Mol Biol* [Internet]. 2005;12(4):304–12. Available from: <http://www.nature.com/doi/10.1038/nsmb915><http://www.ncbi.nlm.nih.gov/pubmed/15778718><http://www.nature.com/nsmb/journal/v12/n4/pdf/nsmb915.pdf>
  117. Della M. Mycobacterial Ku and Ligase Proteins Constitute a Two-Component NHEJ Repair Machine. *Science* (80- ) [Internet]. 2004;306(5696):683–5. Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.1099824>
  118. De Ory A, Zafra O, De Vega M. Efficient processing of abasic sites by bacterial nonhomologous end-joining Ku proteins. *Nucleic Acids Res*. 2014;42(21):13082–95.
  119. Mishra K, Shore D. Yeast Ku protein plays a direct role in telomeric silencing and counteracts inhibition by Rif proteins. *Curr Biol*. 1999;9(19):1123–6.
  120. Hsu HL, Gilley D, Blackburn EH, Chen DJ. Ku is associated with the telomere in mammals. *Proc Natl Acad Sci U S A* [Internet]. 1999;96(22):12454–8. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=22947&tool=pmcentrez&rendertype=abstract>
  121. Bertuch AA, Lundblad V. The Ku heterodimer performs separable activities at double-strand breaks and chromosome termini. *Mol Cell Biol* [Internet]. 2003;23(22):8202–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14585978>
  122. Celli GB, Denchi EL, de Lange T. Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. *Nat Cell Biol* [Internet]. 2006 Aug [cited 2016 Apr 8];8(8):885–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16845382>
  123. Ribes-Zamora A, Indiviglio SM, Mihalek I, Williams CL, Bertuch AA. TRF2 Interaction with Ku Heterotetramerization Interface Gives Insight into c-NHEJ Prevention at Human Telomeres. *Cell Rep* [Internet]. 2013 Oct 17 [cited 2016 Apr 8];5(1):194–206. Available from:

<http://www.scopus.com/inward/record.url?eid=2-s2.0-84885872275&partnerID=tZOtx3y1>

124. Roy R, Meier B, McAinsh AD, Feldmann HM, Jackson SP. Separation-of-function Mutants of Yeast Ku80 Reveal a Yku80p-Sir4p Interaction Involved in Telomeric Silencing. *J Biol Chem*. 2004;279(1):86–94.
125. Larcher M V., Pasquier E, MacDonald RS, Wellinger RJ. Ku Binding on Telomeres Occurs at Sites Distal from the Physical Chromosome Ends. *PLOS Genet* [Internet]. 2016;12(12):e1006479. Available from: <http://dx.plos.org/10.1371/journal.pgen.1006479>
126. Pitcher RS, Brissett NC, Doherty AJ. Nonhomologous End-Joining in Bacteria: A Microbial Perspective. *Annu Rev Microbiol* [Internet]. 2007;61(1):259–82. Available from: <http://www.annualreviews.org/doi/10.1146/annurev.micro.61.080706.093354>
127. d’Adda di Fagagna F, Weller GR, Doherty AJ, Jackson SP. The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep* [Internet]. 2003;4(1):47–52. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1315809&tool=pmcentrez&rendertype=abstract>
128. Myung K, Braastad C, He DM, Hendrickson EA. KARP-1 is induced by DNA damage in a p53- and ataxia telangiectasia mutated-dependent fashion. *Proc Natl Acad Sci U S A*. 1998;95(13):7664–9.
129. Myung K, He DM, Lee SE, Hendrickson EA. KARP-1: A novel leucine zipper protein expressed from the Ku86 autoantigen locus is implicated in the control of DNA-dependent protein kinase activity. *EMBO J*. 1997;16(11):3172–84.
130. Coffey G, Campbell C. An alternate form of Ku80 is required for DNA end-binding activity in mammalian mitochondria. *Nucleic Acids Res*. 2000;28(19):3793–800.
131. Aravind L, Koonin E V. SAP - A putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci* [Internet]. 2000;25(3):112–4. Available from: [http://dx.doi.org/10.1016/S0968-0004\(99\)01537-6](http://dx.doi.org/10.1016/S0968-0004(99)01537-6)
132. Ribes-Zamora A, Mihalek I, Lichtarge O, Bertuch A a. Distinct faces of the Ku heterodimer mediate DNA repair and telomeric functions. *Nat Struct Mol Biol* [Internet]. 2007 Apr [cited 2016 Apr 8];14(4):301–7. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-34247194102&partnerID=tZOtx3y1>
133. Gu Y, Jin S, Gao Y, Weaver DT, Alt FW. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Immunology* [Internet]. 1997;94(15):8076–81. Available from: <http://pubget.com/paper/9223317?institution=cam.ac.uk%5Cnpapers3://publication/uuid/0CA73E75-5580-4F9A-9EEA-6A7D7FEDBC3D>
134. Zhu C, Bogue MA, Lim DS, Hasty P, Roth DB. Ku86-deficient mice exhibit

- severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell* [Internet]. 1996 Aug [cited 2016 Apr 6];86(3):379–89. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0030576534&partnerID=tZOtx3y1>
135. Lim D-S, Vogel H, Willerford DM, Sands AT, Platt KA, Hasty P. Analysis of ku80-Mutant Mice and Cells with Deficient Levels of p53. *Mol Cell Biol* [Internet]. 2000 Jun 1 [cited 2016 Apr 6];20(11):3772–80. Available from: <http://mcb.asm.org/cgi/doi/10.1128/MCB.20.11.3772-3780.2000>
  136. Nussenzweig A, Chen C, da Costa Soares V, Sanchez M, Sokol K, Nussenzweig MC, et al. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. [Internet]. Vol. 382, *Nature*. 1996. p. 551–5. Available from: <http://pubget.com/paper/8700231?institution=cam.ac.uk%5Cnpapers3://publication/doi/10.1038/382551a0>
  137. Holcomb VB, Vogel H, Hasty P. Deletion of Ku80 causes early aging independent of chronic inflammation and Rag-1-induced DSBs. *Mech Ageing Dev* [Internet]. 2007 Jan [cited 2016 Apr 6];128(11–12):601–8. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-37349074760&partnerID=tZOtx3y1>
  138. Vogel H, Lim DS, Karsenty G, Finegold M, Hasty P. Deletion of Ku86 causes early onset of senescence in mice. *Proc Natl Acad Sci U S A* [Internet]. 1999 Sep 14 [cited 2016 Apr 6];96(19):10770–5. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0032827888&partnerID=tZOtx3y1>
  139. Li H, Vogel H, Holcomb VB, Gu Y, Hasty P. Deletion of Ku70, Ku80, or both causes early aging without substantially increased cancer. *Mol Cell Biol* [Internet]. 2007 Dec [cited 2016 Apr 6];27(23):8205–14. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-36849088515&partnerID=tZOtx3y1>
  140. Li H, Choi YJ, Hanes MA, Marple T, Vogel H, Hasty P. Deleting Ku70 is milder than deleting Ku80 in p53-mutant mice and cells. *Oncogene* [Internet]. 2009 Apr 23 [cited 2016 Apr 6];28(16):1875–8. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-67349207122&partnerID=tZOtx3y1>
  141. Difilippantonio MJ, Zhu J, Chen HT, Meffre E, Nussenzweig MC, Max EE, et al. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* [Internet]. 2000;404(6777):510–4. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4721590&tool=pmcentrez&rendertype=abstract>
  142. Koike M, Shiomi T, Koike A. Ku70 Can Translocate to the Nucleus Independent of Ku80 Translocation and DNA-PK Autophosphorylation. *Biochem Biophys Res Commun* [Internet]. 2000;276(3):1105–11. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0006291X00935678>

143. Koike M, Ikuta T, Miyasaka T, Shiomi T. Ku80 can translocate to the nucleus independent of the translocation of Ku70 using its own nuclear localization signal. *Oncogene* [Internet]. 1999;18(52):7495–505. Available from: <http://www.nature.com/doifinder/10.1038/sj.onc.1203247>
144. Lim JW, Kim KH, Kim H. NF- $\kappa$ B p65 regulates nuclear translocation of Ku70 via degradation of heat shock cognate protein 70 in pancreatic acinar AR42J cells. *Int J Biochem Cell Biol* [Internet]. 2008 Jan [cited 2016 Apr 8];40(10):2065–77. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-48949117384&partnerID=tZOtx3y1>
145. Subramanian C, Jarzembowski J a, Opiari AW, Castle VP, Kwok RP. HDAC6 Deacetylates Ku70 and Regulates Ku70-Bax Binding in Neuroblastoma. *Neoplasia* [Internet]. 2011 Aug [cited 2016 Apr 8];13(8):726–34. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-80051496447&partnerID=tZOtx3y1>
146. Subramanian C, Opiari AW, Bian X, Castle VP, Kwok RPS. Ku70 acetylation mediates neuroblastoma cell death induced by histone deacetylase inhibitors. *Proc Natl Acad Sci U S A* [Internet]. 2005 Mar 29;102(13):4842–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15778293>
147. Cohen HY, Lavu S, Bitterman KJ, Hekking B, Imahiyerobo TA, Miller C, et al. Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol Cell*. 2004;13(5):627–38.
148. Gomez J a, Gama V, Yoshida T, Sun W, Hayes P, Leskov K, et al. Bax-inhibiting peptides derived from Ku70 and cell-penetrating pentapeptides. *Biochem Soc Trans* [Internet]. 2007 Aug;35(Pt 4):797–801. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17635151>
149. Walden EA. Assessing the role of Ku70 vWA domain phosphorylation in the inhibition of Aurora B and activation of the DNA damage response. University of Western Ontario; 2017.
150. Xing M, Yang M, Huo W, Feng F, Wei L, Jiang W, et al. Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway. *Nat Commun* [Internet]. 2015;6:1–12. Available from: <http://dx.doi.org/10.1038/ncomms7233%5Cn10.1038/ncomms7233>
151. Ochi T, Blackford AN, Coates J, Jhujh S, Mehmood S, Tamura N, et al. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science* (80- ) [Internet]. 2015;347(6218):185–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25574025>
152. Craxton A, Somers J, Munnur D, Jukes-Jones R, Cain K, Malewicz M. XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair. *Cell Death Differ* [Internet]. 2015 Jun [cited 2016 Apr 8];22(6):890–7. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4423191&tool=pmcentrez&rendertype=abstract>

153. Li G, Alt FW, Cheng HL, Brush JW, Goff PH, Murphy MM, et al. Lymphocyte-Specific Compensation for XLF/Cernunnos End-Joining Functions in V(D)J Recombination. *Mol Cell*. 2008;31(5):631–40.
154. Balmus G, Barros AC, Wijnhoven PWG, Lescale CC, Hasse HLHL, Boroviak K, et al. Synthetic lethality between PAXX and XLF in mammalian development. *Genes Dev*. 2016;30(19):2152–7.
155. Tomkinson AE, Totty NF, Ginsburg M, Lindahl T. Location of the active site for enzyme-adenylate formation in DNA ligases. *Proc Natl Acad Sci U S A* [Internet]. 1991;88(2):400–4. Available from: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list\\_uids=1988940](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=1988940)
156. Chaikuad A, Knapp S, von Delft F. Defined PEG smears as an alternative approach to enhance the search for crystallization conditions and crystal-quality improvement in reduced screens. *Acta Crystallogr D Biol Crystallogr* [Internet]. 2015;71(Pt 8):1627–39. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4528798&tool=pmcencetrez&rendertype=abstract>
157. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr Sect D Biol Crystallogr*. 2011;67(4):271–81.
158. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, et al. PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr Sect D Biol Crystallogr*. 2002;58(11):1948–54.
159. Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr Sect D Biol Crystallogr*. 2012;68(4):352–67.
160. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr*. 2010;66(4):486–501.
161. Andres SN. XRCC4 AND XLF IN MAMMALIAN DNA DOUBLE-STRAND BREAK REPAIR. McMaster University; 2011.
162. Andres SN, Vergnes A, Ristic D, Wyman C, Modesti M, Junop M. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Res* [Internet]. 2012;40(4):1868–78. Available from: <http://dx.doi.org/10.1093/nar/gks022>
163. Lee WKY. CHARACTERIZATION OF THE OLIGOMERIZATION STATE OF XRCC4. McMaster University; 2013.
164. Genersch E, Eckerskorn C, Lottspeich F, Herzog C, Kühn K, Pöschl E. Purification of the sequence-specific transcription factor CTCBF, involved in the control of human collagen IV genes: subunits with homology to Ku antigen. *EMBO J* [Internet]. 1995;14(4):791–800. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC398145/>

165. Hanakahi LA. 2-Step purification of the Ku DNA repair protein expressed in *Escherichia coli*. *Protein Expr Purif* [Internet]. 2007 Mar [cited 2016 Apr 8];52(1):139–45. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33845937721&partnerID=tZOtx3y1>
166. Junop MS, Modesti M, Guarné A, Ghirlando R, Gellert M, Yang W. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. *EMBO J*. 2000;19(22):5962–70.
167. Sawada M, Sun W, Hayes P, Leskov K, Boothman D a, Matsuyama S. Ku70 suppresses the apoptotic translocation of Bax to mitochondria. *Nat Cell Biol*. 2003;5(4):320–9.
168. Choi YJ, Li H, Son MY, Wang XH, Fornsaglio JL, Sobol RW, et al. Deletion of individual Ku subunits in mice causes an NHEJ-independent phenotype potentially by altering apurinic/apyrimidinic site repair. *PLoS One* [Internet]. 2014 Jan [cited 2016 Apr 6];9(1):e86358. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84899688897&partnerID=tZOtx3y1>
169. Kleckner IR., Foster MP. An Introduction to NMR based approaches for measuring protein dynamics. *Biochim Biophys Acta* . 2011;1814(8):942–68.
170. Wales TE, Engen JR. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev*. 2006;25(1):158–70.
171. Rossmann MG, Morais MC, Leiman PG, Zhang W. Combining X-ray crystallography and electron microscopy. *Structure*. 2005;13(3):355–62.
172. Bai X, Yan C, Yang G, Lu P, Ma D, Sun L, et al. An atomic structure of human  $\gamma$ -secretase. *Nature* [Internet]. 2015;525(7568):212–7. Available from: <http://www.nature.com/doifinder/10.1038/nature14892>
173. Mazumder S, Plesca D, Kinter M, Almasan A. Interaction of a cyclin E fragment with Ku70 regulates Bax-mediated apoptosis. *Mol Cell Biol* [Internet]. 2007 May [cited 2016 Apr 8];27(9):3511–20. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-34247564493&partnerID=tZOtx3y1>
174. Mazumder S, Plesca D, Almasan A. A jekyll and hyde role of cyclin E in the genotoxic stress response: Switching from cell cycle control to apoptosis regulation. *Cell Cycle* [Internet]. 2007;6(12):1437–42. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-35948999706&partnerID=tZOtx3y1>
175. Mukherjee S, Chakraborty P, Saha P. Phosphorylation of Ku70 subunit by cell cycle kinases modulates the replication related function of Ku heterodimer. *Nucleic Acids Res* [Internet]. 2016;44(16):7755–65. Available from: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkw622>
176. Ngo J, Matsuyama M, Kim C, Poventud-Fuentes I, Bates A, Siedlak SL, et al. Bax deficiency extends the survival of Ku70 knockout mice that develop lung and heart

- diseases. *Cell Death Dis* [Internet]. 2015;6(3):e1706. Available from: <http://dx.doi.org/10.1038/cddis.2015.11> %5Cnpapers3://publication/doi/10.1038/cddis.2015.11 %5Cn<http://www.ncbi.nlm.nih.gov/pubmed/25811803> %5Cn<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4385910>
177. Matsuyama S, Palmer J, Bates A, Poventud-Fuentes I, Wong K, Ngo J, et al. Bax-induced apoptosis shortens the life span of DNA repair defect Ku70-knockout mice by inducing emphysema. *Exp Biol Med* [Internet]. 2016;241(12):1265–71. Available from: <http://ebm.sagepub.com/lookup/doi/10.1177/1535370216654587>
  178. Vishnudas VK, Miller JB. Ku70 regulates Bax-mediated pathogenesis in laminin-alpha2-deficient human muscle cells and mouse models of congenital muscular dystrophy. *Hum Mol Genet* [Internet]. 2009 Dec 1 [cited 2016 Apr 8];18(23):4467–77. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-70449419707&partnerID=tZOtx3y1>
  179. Hung PJ, Chen B-R, George R, Liberman C, Morales AJ, Colon-Ortiz P, et al. Deficiency of XLF and PAXX Prevents DNA Double-Strand Break Repair by Non-homologous End Joining in Lymphocytes. *Cell Cycle* [Internet]. 2016;16(3):0. Available from: <http://dx.doi.org/10.1080/15384101.2016.1253640>
  180. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* [Internet]. 2005;434(7035):913–7. Available from: <http://www.nature.com/doifinder/10.1038/nature03443>
  181. Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* [Internet]. 2005;434(7035):917–21. Available from: <http://www.nature.com/doifinder/10.1038/nature03445>
  182. Konecny GE, Kristeleit RS. PARP inhibitors for BRCA1/2-mutated and sporadic ovarian cancer: current practice and future directions. *Br J Cancer* [Internet]. 2016;115(10):1157–73. Available from: <http://www.nature.com/doifinder/10.1038/bjc.2016.311>
  183. Dedes KJ, Wilkerson PM, Wetterskog D, Weigelt B, Ashworth A, Reis-Filho JS. Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations. *Cell Cycle*. 2011;10(8):1192–9.

## Appendices

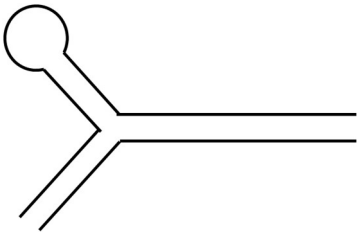
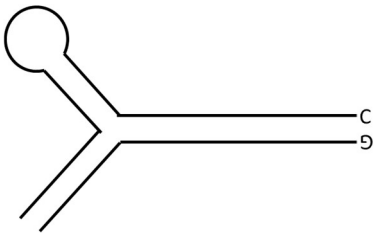
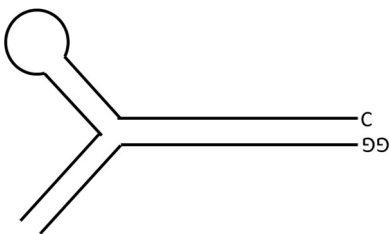
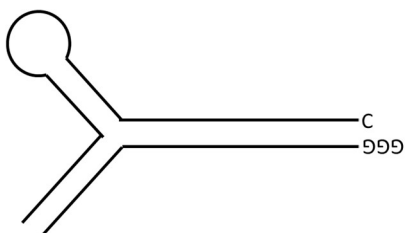
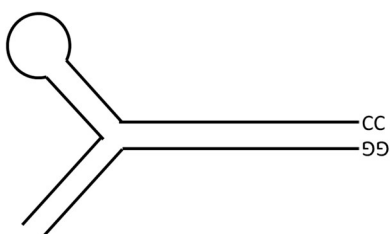
**Appendix 1. List of oligonucleotides.** Oligonucleotides synthesized and utilized for crystallography are listed here.

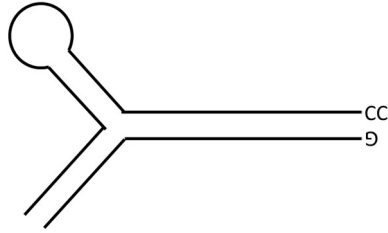
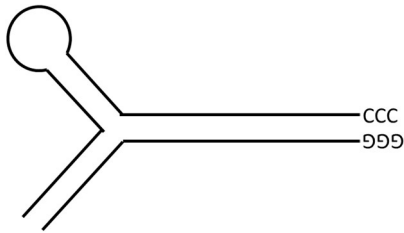
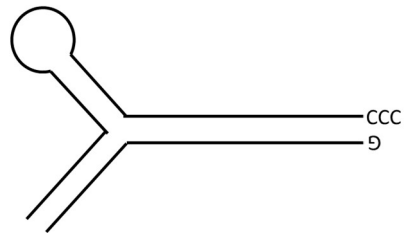
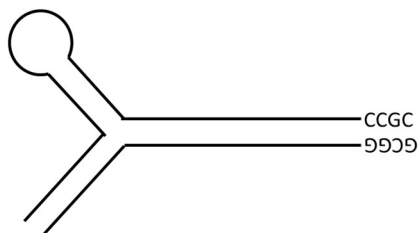
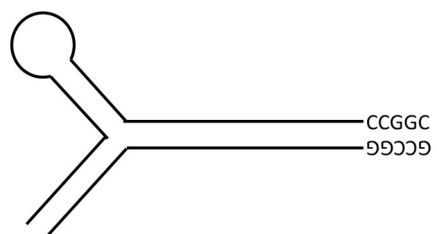
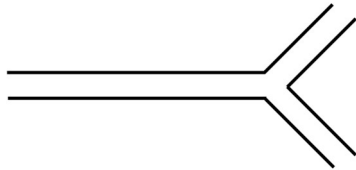
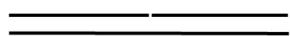
Oligonucleotide	Sequence (5'-3')
Loop14-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAAC
Loop15-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACC
Loop16-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACCC
Loop17-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACCCC
Loop18-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACCCGC
Loop19-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACCCGGC
Loop34-1	CGCGCCCAGCTTTCCCAGCTAATAAACTAAAAACCCAATAAACTAAAA CCCCC
Loop14-2	GTTTTTAGTTTATTGGGCGCG
Loop15-2	GGTTTTTAGTTTATTGGGCGCG
Loop16-2	GGGTTTTTAGTTTATTGGGCGCG
Loop17-2	GGGGTTTTTAGTTTATTGGGCGCG
Loop18-2	GCGGGTTTTTAGTTTATTGGGCGCG
Loop19-2	GCCGGTTTTTAGTTTATTGGGCGCG
Loop34-2	GGGGGTTTTTAGTTTATTGGGTTTTTAGTTTATTGGGCGCG
Y7-1	TTTATTACCTCCCCTACCCAC
Y7-2	GTGGGTAGGGGAGGATTGTTT
Y7-3	AAACAATTAATAAA
TS1	GATCCCTCTAGATAT
TS2	CGGGCCCTCGATCCG
TS3	CGGATCGAGGGCCCGATATCTAGAGGGATC
14b1T	TGGGCTGGTCGGGTT
14b2T	TGGGCTGGTCGGGTTT
14b3T	TGGGCTGGTCGGGTTTT
14b4T	TGGGCTGGTCGGGTTTTT
14b5T	TGGGCTGGTCGGGTTTTTT
14b10T	TGGGCTGGTCGGGTTTTTTTTTT
14bottom	ACCCGACCAGCCCA
KP1530b	CGCGAGCTTTCCCAGCTGATCCCTCTAGATAT

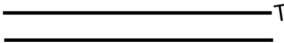
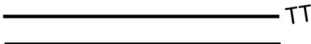
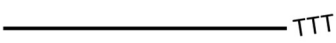
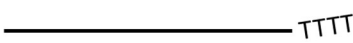
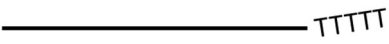
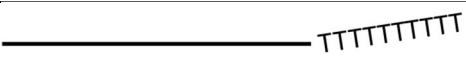
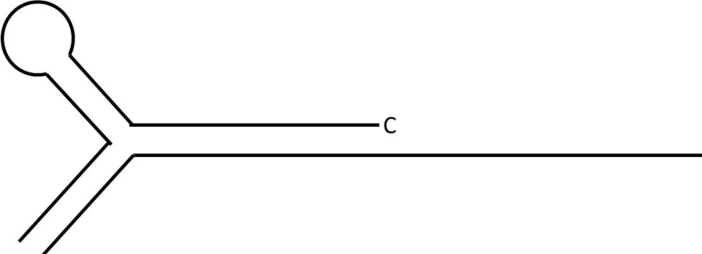
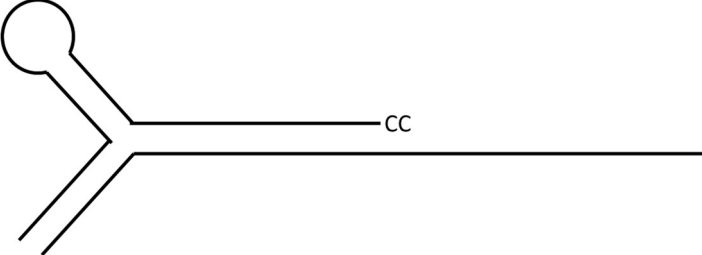
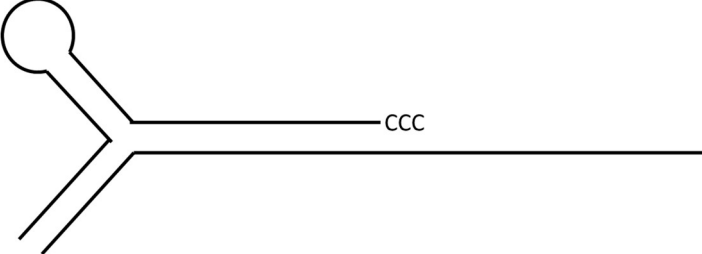


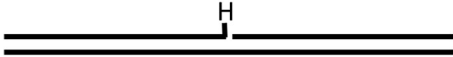
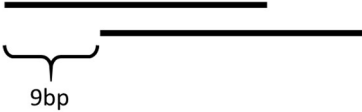
<b>KP1630b</b>	<b>CGCGAGCTTTCCCAGCTGATCCCTCTAGATATC</b>
<b>KP1730b</b>	<b>CGCGAGCTTTCCCAGCTGATCCCTCTAGATATCA</b>
<b>18stall-1</b>	<b>TGATGCGTc (3' dideoxycytosine)</b>
<b>18stall-2</b>	<b>(5' phosphate) GTCAGGCTG</b>
<b>18stall-3</b>	<b>CAGCCTGACGACGCATCA</b>
<b>stack9-1</b>	<b>ATGATTAGAACGGACACTGGATTGTGACCT</b>
<b>stack9-2</b>	<b>TCTAATCATAGGTCACAATCCAGTGTCCGT</b>

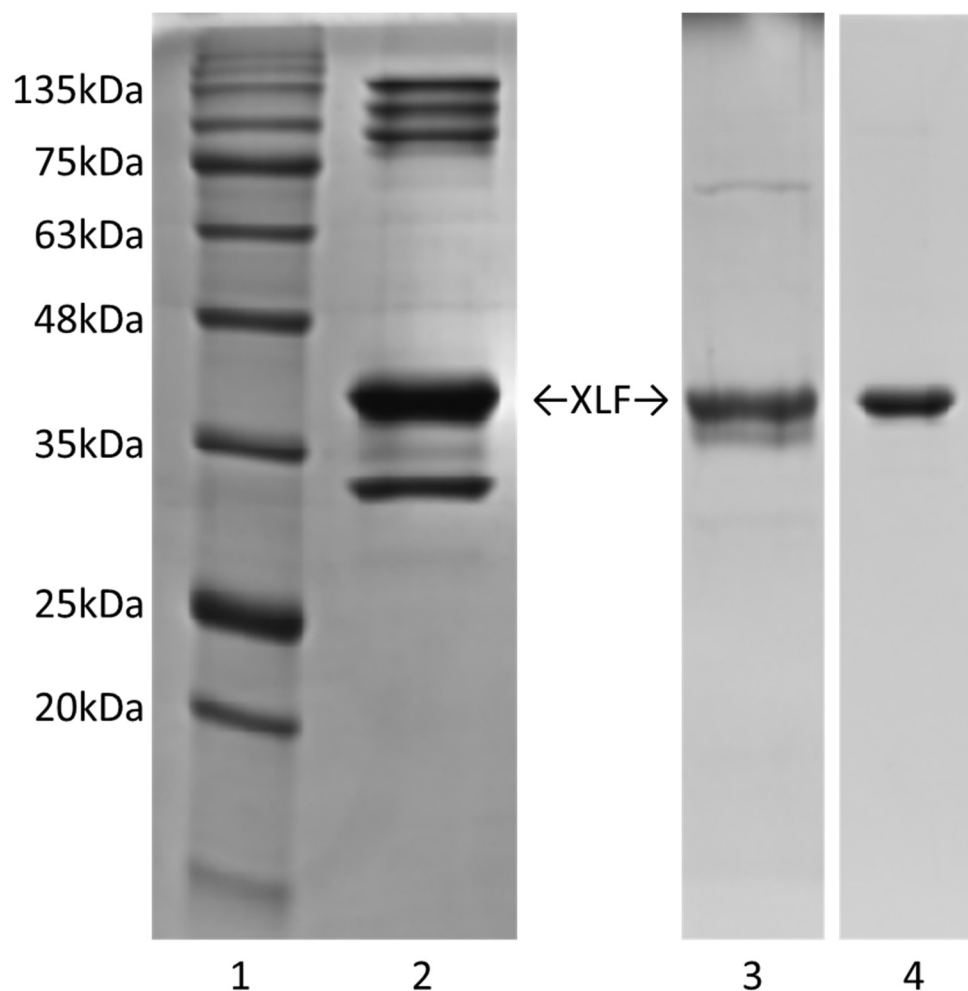
### Appendix 2. Structure of annealed DNA substrates.

Name	Component	Illustrated Structure
Loop14	Loop14-1 & loop14-2	
Loop15	Loop15-1 & loop15-2	
Loop15/16	Loop15-1 & loop16-2	
Loop15/17	Loop15-1 & loop17-2	
Loop16	Loop16-1 & loop16-2	

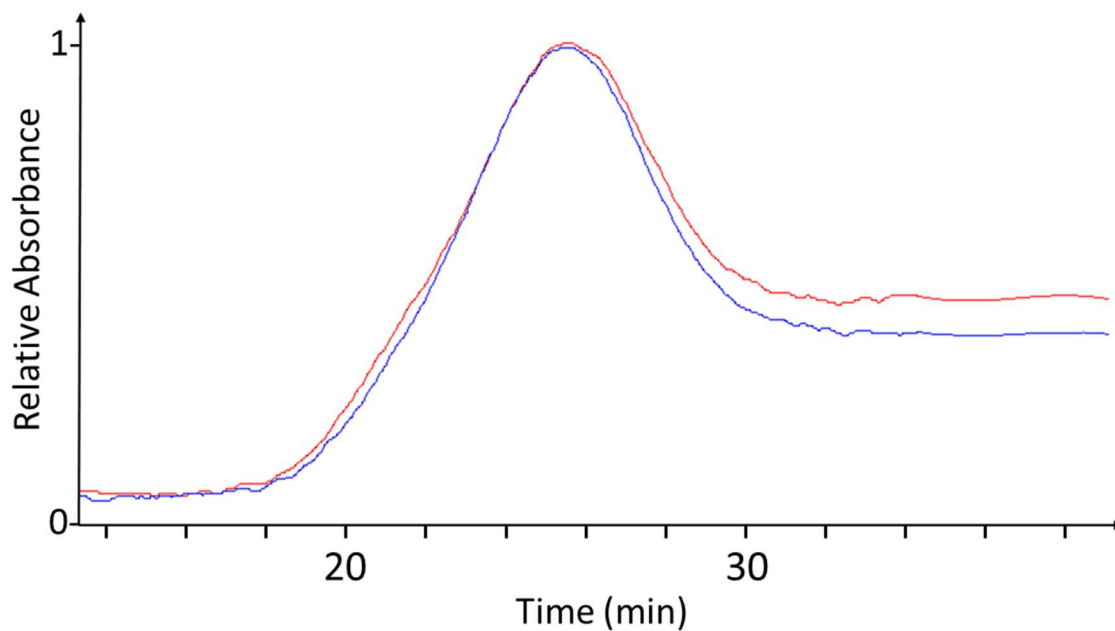
Loop16/15	Loop16-1 & loop15-2	
Loop17	Loop17-1 & loop17-2	
Loop17/15	Loop17-1 & loop15-2	
Loop18	Loop18-1 & loop18-2	
Loop19	Loop19-1 & loop19-2	
Y7	Y7-1 & Y7-2 & Y7-3	
TS	TS1 & TS2 & TS3	

14b1Tail	14b1T & 14bottom	
14b2Tail	14b1T & 14bottom	
14b3Tail	14b1T & 14bottom	
14b4Tail	14b1T & 14bottom	
14b5Tail	14b1T & 14bottom	
14b10Tail	14b1T & 14bottom	
KP1530	Loop15 & KP1530b	
KP1630	Loop16 & KP1630b	
KP1730	Loop17 & KP1730b	

<b>18stall</b>	<b>18stall-1, 18stall-2, &amp; 18stall-3</b>	
<b>Stack9</b>	<b>Stack9-1 &amp; stack9-2</b>	



**Appendix 3.** Comparison of XLF expression and purification using different methods. Lane 1 shows the Trident Prestained Protein Ladder (GeneTex); lane 2 shows XLF expressed in *E. coli* BL21 (DE3) and purified using the method indicated in Chapter 2.2.2, but using MonoQ in place of Heparin column; lane 3 shows XLF expressed in *E. coli* Rosetta™(DE3)pLysS and purified using the same method as lane 1; lane 4 shows XLF expressed in *E. coli* Rosetta™(DE3)pLysS, and purified using Heparin column as the second chromatography method.



**Appendix 4. Ku70-DNA bound in solution.** The complex formation of Ku70 and loop16 was evaluated with size-exclusion chromatography using similar procedure to Chapter 2.6 with a dual-detection fast protein liquid chromatography system. The red line indicates absorbance at 230nm, the blue line indicates absorbance at 280nm. DNA and Ku70 were both detected by gel electrophoresis of fractionated samples, indicating interaction.

**Appendix 5. List of crystallization conditions for Ku70 and X-ray screening outcomes.**

Crystallization condition	X-ray screening result
0.2M potassium sulphate, 20% w/v PEG 3350	Salt
0.2M potassium formate pH7.3, 20% w/v PEG3350	Salt
0.2M potassium acetate, 20% w/v PEG3350	Salt
1M NaH <sub>2</sub> PO <sub>4</sub> /K <sub>2</sub> HPO <sub>4</sub> pH6.9	Salt
24% w/v PEG 1500, 20% v/v glycerol	Salt
0.2M Lithium sulphate, 0.1M Sodium phosphate dibasic pH4.2, 10% v/v 2-propanol	Salt
1M Potassium sodium tartrate, 0.1M Imidazole pH8, 0.2M Sodium chloride	Salt
0.1M Bis-Tris pH5.5, 25% w/v PEG3350	Non-diffracting crystal
0.2M Calcium acetate, 0.1M Imidazole pH8, 20% w/v PEG1000	Non-diffracting crystal
0.2M Calcium acetate, 0.1M Imidazole pH8, 10% w/v PEG8000	Non-diffracting crystal
0.2M Lithium sulphate, 0.1M Tris pH7, 1M Potassium sodium tartrate	Non-diffracting crystal
0.8M potassium sodium tartrate, 0.1M Tris pH8.5, 0.5% w/v PEG5000 MME	Non-diffracting crystal
Condition above with 0.1M Strontium chloride hexahydrate	Salt



**Appendix 6. List of crystallization conditions for Ku70-DNA and X-ray screening outcomes.** The three bolded conditions underwent extensive optimization varying multiple different components for crystallization.

Crystallizing condition	DNA substrate	X-ray screening result (Resolution)
0.15M DL-Malic Acid pH 7, 20% w/v PEG3350	Loop14	Non-diffracting crystal
0.2M magnesium chloride, 0.1M sodium citrate pH5.5, 40% v/v PEG400	Loop14	Non-diffracting crystal
0.2M sodium malonate pH7, 20% w/v PEG3350	Loop14	Non-diffracting crystal
0.8M lithium chloride, 0.1M Tris pH8.5, 32% w/v PEG4000	Loop14	Non-diffracting crystal
0.2M sodium chloride, 0.1M HEPES pH7.5, 25% w/v PEG3350	Loop14	Protein (15Å)
1M lithium chloride, 30% w/v PEG6000, 0.1M sodium acetate	Loop14	Salt
0.1M Lithium sulphate monohydrate, 0.1M Sodium citrate tribasic dihydrate pH5.5, 20% w/v PEG1000	Loop16	Non-diffracting crystal
0.2M ammonium formate pH6.6, 20% w/v PEG3350	Loop16	Non-diffracting crystal
0.2M Ammonium sulphate, 0.1M BIS-TRIS pH6.5, 18% v/v PEG400	Loop16	Non-diffracting crystal
0.6M sodium chloride, 0.1M MES pH6.5, 20% w/v PEG4000	Loop16	Non-diffracting crystal
<b>0.2M sodium chloride, 0.1M HEPES pH7.5, 25% w/v PEG3350</b>	<b>Loop16</b>	<b>Protein (15Å)</b>
<b>0.1M Tris pH7, 20% w/v PEG1000</b>	<b>Loop16</b>	<b>Protein (20Å)</b>
<b>0.1M Tris pH7, 20% w/v PEG2000 MME</b>	<b>Loop16</b>	<b>Protein (20Å)</b>
0.2M calcium chloride, 0.1M Tris pH8.5, 25% w/v PEG4000	Loop16	Salt
0.2M magnesium chloride, 0.1M sodium citrate pH5.5, 40% v/v PEG400	Loop16	Salt
0.2M magnesium chloride, 0.1M Bis-Tris pH5.5, 25% w/v PEG3350	Loop16	Salt
0.2M magnesium chloride, 20% w/v PEG3350	Loop16	Salt
0.8M lithium chloride, 0.1M Tris pH8.5, 32% w/v PEG4000	Loop16	Salt
0.8M potassium sodium tartrate, 0.1M Tris pH8.5, 0.5% w/v PEG5000 MME	Loop16	Salt
1M lithium chloride, 30% w/v PEG6000, 0.1M sodium acetate	Loop16	Salt
2.5M Sodium chloride, 0.1M Imidazole pH8, 0.2M Zinc acetate	Loop16	Salt

3% w/v Dextran sulphate sodium salt, 0.1M BICINE pH8.5, 15% w/v PEG20000	Loop16	Salt
0.2M Magnesium acetate tetrahydrate, 20% w/v PEG3350	Loop18	Salt
0.2M Magnesium chloride, 20% w/v PEG3350	Loop18	Salt
0.2M Potassium citrate tribasic monohydrate, 20% w/v PEG3350	Loop18	Salt
0.2M Potassium sulphate, 20% w/v PEG3350	Loop18	Salt
0.2M Sodium acetate trihydrate, 20% w/v PEG3350	Loop18	Salt
0.2M Zinc acetate dihydrate, 20% w/v PEG3350	Loop18	Salt
0.2M Ammonium acetate, 0.1M Sodium citrate tribasic dihydrate pH5.5, 24% v/v PEG400	Loop19	Non-diffracting crystal
4% v/v (+/-)-2-Methyl-2,4-pentanediol, 0.1M Citric acid pH3.5, 20% w/v PEG1500	Loop19	Non-diffracting crystal

**Appendix 7. Optimization of Wizard I #10 and MCSG II #80 for Ku70-loop16 crystal growth and X-ray diffraction.**

Varying conditions	Resolution
0.1M Tris pH7, 20% w/v PEG2000 MME; additive 1M Sodium malonate pH7	20Å
0.1M Tris pH7, 20% w/v PEG2000 MME; additive 30% w/v 1,5-Diaminopentane dihydrochloride	20Å
<b>0.1M Tris pH7, 20% w/v PEG1000; 0.2M sodium iodide</b>	<b>8Å</b>
Further additive screen	
0.16% w/v 3-Aminobenzenesulfonic acid, 5-Sulfosalicylic acid dihydrate, p-Coumaric acid, PIPES, Terephthalic acid, Vanillic acid mixture	16Å
0.2% w/v (±)-2-Methyl-2,4-pentanediol, 1,2,3-Heptanetriol, Diethylenetriaminepentakis(methylphosphonic acid), D-Sorbitol, Glycerol mixture	20Å
0.2% w/v Barbituric acid, Betaine anhydrous, Phloroglucinol Resorcinol, Tetrahydroxy-1,4-benzoquinone hydrate mixture	15Å
0.2% w/v 2,5-Pyridinedicarboxylic acid, Pyromellitic acid Salicylic acid, trans-1,2-Cyclohexanedicarboxylic acid, trans-Cinnamic acid mixture	20Å

**Appendix 8. Ku70-DNA, Ku70-DNA-PAXX crystallization conditions.** Listed are the conditions that generated crystals. The first column indicates the name of the crystallization kit where the condition was obtained; second column, salt identity and concentration; third column, buffer identity, concentration and pH; fourth column, percentage concentration by weight of PEG. None of the Ku70 crystals under these conditions yielded diffraction data. Ku70-DNA and Ku70-DNA-PAXX crystals yielded diffraction to varying degrees of resolution.

Ku70-DNA	Initial hits		
MCSG I #95	0.2M sodium chloride	0.1M HEPES pH7.5	25% w/v PEG3350
MCSG II #80		0.1M Tris pH7	20% w/v PEG1000
MCSG I #95 optimize 1	0.2M sodium chloride	0.1M HEPES pH7.5	20% w/v PEG2000
MCSG II #80 optimize 1	0.2M sodium iodide	0.1M Tris pH7	20% w/v PEG1000
MCSG I #82	0.2M magnesium formate		15% w/v PEG3350
MCSG II #83		0.1M Na <sub>2</sub> HPO <sub>4</sub> pH4.2	40% w/v PEG300
PAXX	Initial hits		
MCSG III #31		0.1M Bis-Tris Propane pH7	1M ammonium citrate tribasic pH7
MCSG II #36		0.1M Tris pH8.5	1.4M ammonium tartrate dibasic
MCSG I #9	0.2M magnesium chloride	0.1M HEPES pH7.5	25% w/v PEG3350
MCSG I #12	0.2M calcium chloride	0.1M Tris pH8.5	20% w/v PEG4000
MCSG I #42		0.1M Bis-Tris Propane pH7	1.4M sodium malonate pH 7.0
MCSG I #44	0.1M sodium chloride	0.1M Bis-Tris pH6.5	1.5M Ammonium Sulphate
MCSG I #72	0.1M potassium thiocyanate		30% w/v PEG2000 MME
MCSG I #96	0.2M lithium sulphate	0.1M HEPES pH7.5	25% w/v PEG3350
MCSG II #33	0.2M sodium fluoride		20% w/v PEG3350

## Curriculum Vitae

**Name:** Huasheng Wang

**Post-secondary Education and Degrees:** McMaster University  
Hamilton, Ontario, Canada  
2011-2015 B.Sc.

The University of Western Ontario  
London, Ontario, Canada  
2015-2017 M.Sc.

**Honours and Awards:** Province of Ontario Graduate Scholarship  
2015-2017

**Related Work Experience** Teaching Assistant  
The University of Western Ontario  
2015-2017

### Publications:

Qiaojuan Shi, Marco R. Straus, Jeremy J. Caron, Huasheng Wang, Yu Seon Chung, Alba Guarné, & Joseph E. Peters. (2015). Conformational toggling controls target site choice for the heteromeric transposase element Tn7. *Nucleic Acids Research*, 43(22), 10734–10745. <https://doi.org/10.1093/nar/gkv913>