STRUCTURAL AND FUNCTIONAL STUDIES OF NEJ1

STRUCTURAL AND FUNCTIONAL STUDIES

OF

NON-HOMOLOGOUS END-JOINING REGULATOR 1 (NEJ1)

By

MARGARET SULEK, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

For the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2007) McMaster University

(Biochemistry and Biomedical Sciences) Hamilton, Ontario

TITLE:	Structural and Functional Studies of Non-Homologous End-Joining Regulator 1 (Nej1)
AUTHOR:	Margaret Sulek, B.Sc. (McMaster University)
SUPERVISOR:	Dr. Murray Junop
NUMBER OF PAGES:	xvi, 135

ABSTRACT

Repair of double-strand breaks is critical for the preservation of genomic integrity and cellular viability. A predominant pathway implicated in the repair of such lesions is the evolutionarily conserved non-homologous end-joining (NHEJ) pathway. Among the major constituents of this pathway in *Saccharomyces cerevisiae* is Nej1, for which a clear biochemical function has not been determined. The results presented in this work demonstrate that Nej1 exhibits a DNA binding activity comparable to Lif1, with an apparent dissociation constant of 1.8 μM. Characterization of the DNA binding activity revealed that although short ~20 bp substrates can suffice, binding is enhanced with longer substrates (>300). This DNA binding activity supports the hypothesis that Nej1 plays a direct role in the repair of DNA double-strand breaks. Structure-function studies indicated that the C-terminus of Nej1 is not only required, but is sufficient, for mediating DNA interactions. Structural characterization revealed that Nej1 exists as a dimer, and that residues 1-244 are sufficient for dimer formation. Examining the ability of this truncated Nej1 (aa 1-244) to perform NHEJ, revealed a defect in end-joining as measured by an *in vivo* plasmid repair assay. Preliminary functional and structural studies on the Nej1-Lif1 complex suggest that the proteins stably co-purify and the complex binds DNA with a higher affinity than each independent component. The significance of these results is discussed with reference to current literature on Nej1 and other end-joining factors (mammalian and yeast), specifically the recently identified putative mammalian homologue of Nej1, XLF. Collectively, these results demonstrate that although there are

several functional similarities, there also appear to be important differences in the structure-function relationships of Nej1 and XLF, and Nej1/XLF and Lif1/Xrcc4.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my gratitude to my supervisor Dr. Murray Junop, for granting me the opportunity to carry out my graduate studies in his laboratory and allowing me to work on this exciting project. His positive attitude, endless patience and continuous encouragement made for an ideal, relaxed atmosphere in which learning was enjoyable. I am especially grateful for the all his helpful suggestions, and the countless discussions on many topics, without which this project, as well as my education, would not have been as successful. He is a true mentor, both inside and outside of the laboratory, a great role model and I wish him great success in the future.

I would also like to thank my supervisory committee, Dr. David Andrews and Dr. Dr. Xu-Dong Zhu, who provided me with a lot of guidance throughout my studies. Their expertise and numerous helpful suggestions have been invaluable to me and to the success of this project.

I would like to thank the very person who discovered Nej1, Dr. Jef Boeke, for agreeing to collaborate with us on a particular aspect of this project and his student, Robert Yarrington, for carrying out the *in vivo* experiments on the various Nej1 mutants. Their suggestions about the project were extremely valuable, and their patience and continued kindness is much appreciated.

I would like to thank Dr. Grahm McGibbon for taking the time to teach me how to carry out the in-gel trypsin digestion and for showing me how tandem mass

spectrometry is performed. I am grateful for his expertise in mass spectrometry and for his patience in explaining the theory to me.

I would like to thank all the past and present members of Dr. Junop's lab for being a truly great team of people to work with. Kun, Rachael, Sri, Tracy, Andrew W., Sean, Sara, Ashley, Raj, Andrew K., Akil, Uyen, and Luke, all of who have made the lab environment very pleasant and fun. I appreciate all the helpful discussions regarding my project, and I am grateful for their support and their endless words of encouragement. I wish them the best of luck in their future endeavors.

There have been a number of people from the Department who have contributed to the success of this project. Specifically, I would like to thank Dr. Li and his students for being so kind and generous, in allowing me to work in their lab on the experiments that involved radioactivity. I would like to thank Naveen Kumar for teaching me how to radiolabel and purify DNA. I would like to thank Mark Pereira for initially training me on how to carry out analytical ultracentrifugation, and Victor Pau for his suggestions on how to analyze analytical ultracentrifugation data. I would also like to extend my appreciation to Akeel Baig for teaching me about immunobloting and for spending the time to discuss various matters surrounding my project.

I would like to express my final gratitude toward my family, my loving parents Elzbieta and Zygmunt Sulek, and my caring brother Mike, who were truly a tremendous support network for me during my graduate studies. I would also like to thank Bojan for being understanding and always encouraging me, and my friends for being patient and kindhearted.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1 – INTRODUCTION

CHAPTER 3 – RESULTS

CHAPTER 4 – DISCUSSION

APPENDIX I

LIST OF TABLES

CHAPTER 3 - RESULTS

APPENDIX I

LIST OF ABBREVIATIONS

Amp	ampicillin
ATM	ataxia-telangiectasia mutated
ATP	adenosine triphosphate
ATR	ATM-related
BLAST	basic local alignment search tool
βME	beta-mercapto-ethanol
B-NHEJ	back-up non-homologous end joining
BRCT	breast cancer carboxy-terminal
BSA	bovine serum albumin
CD	circular dichroism
Cm	chloramphenicol
DNA	deoxyribonucleic acid
DNA-PK	DNA dependent protein kinase
DNA-PKcs	DNA dependent protein kinase catalytic subunit
D-NHEJ	DNA-PK dependent non-homologous end joining
Dnl4	Yeast DNA ligase 4
dATP	2'-deoxyadenosine 5'-triphosphate
dNTP	2-deoxynucleotide-5'-triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
DSB	double-strand break
DTT	dithiothreitol

EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
GFP	green fluorescent protein
HCA	hydrophobic cluster analysis
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
H2A	histone 2A
H2AX	histone 2A variant
HR	homologous recombination
IPTG	isopropyl-thio-β-D-galactopyranoside
IR	ionizing radiation
K_D	dissociation constant
kDa	kilodalton
LB	luria-bertani
LDAO	N-lauryl-N, N-dimethylamine N-oxide
LIC	ligation independent cloning
LIF1	ligase interacting factor 1
MAD	multiple-anomalous dispersion
MAT	mating type
MR	molecular replacement
MRX	Mre11/Rad50/Xrs2 complex
MS	mass spectrometry
NEJ1	non-homologous end-joining regulator 1

NHEJ	non-homologous end joining
NLS	nuclear localization signal
OD	optical density
PCR	polymerase chain reaction
PIKK	phosphatidylinositol kinase-related kinase
PMG	ploymicrogyria
PMSF	phenylmethanesulphonylfluoride
PNK	polynucleotide kinase
POL4	Polymerase 4
POLX	Polymerase X
RNA	ribonucleic acid
RS-SCID	radiosensitivity severe combined immune deficiency
SCID	severe combined immunodeficiency
SDM	site-directed mutagenesis
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TdT	terminal deoxy-nucleotidyl transferase
TBE	Tris Borate EDTA
TEV	tobacco etch virus
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
V(D)J	variable diverse and joining
XLF	Xrcc4 like factor
Xrcc4	X-ray cross complementation group 4

CHAPTER 1 – INTRODUCTION

1.1 DNA Damage and Double-Strand Breaks

Among the multitude of life threatening events that cells encounter on a daily basis, DNA damage is a frequent occurrence. It is estimated that mammalian cells are insulted with at least 10 000 different DNA damaging events every day (Lindhal *et al.,* 1993). DNA can be damaged in many ways, ranging from base alteration or deletion, formation of crosslinks with the complementary strand, with chemicals or proteins, and strand breakage giving rise to single-strand breaks or double-strand breaks (DSBs). Although each of these DNA damaging events can lead to disastrous outcomes for the cell, the most lethal type of DNA damage is the DSB. These lesions are particularly dangerous because both strands of the duplex are damaged and neither strand can provide physical integrity or information content to restore the break.

Pathological as well as physiological agents can cause DSBs. Pathological agents include ionizing radiation, such as X-rays and gamma rays, which produce free radicals that attack the phosphodiester backbone of the duplex (Morgan *et al.,* 1996). Moreover, chemotherapeutic drugs such as bleomycin and mitomycin C, directly interact with DNA compromising its integrity resulting in lesions on both strands of the duplex (Kamiguchi *et al.,* 1995; Watanabe *et al.,* 1999).

DNA DSBs can also arise due to naturally occurring cellular processes. Oxidative metabolism, which uses electrons from glucose metabolism to drive the synthesis of ATP through the electron transport chain, can generate oxygen free radicals

that are subsequently available to attack DNA (Lodish *et al.,* 2004). Single-strand breaks or other types of lesions may cause stalled replication forks during DNA synthesis, which may in turn be converted to DSBs (Michel *et al.,* 1997). In addition to DNA lesions that arise unintentionally, DSBs are required in specialized pathways, essential for maintaining immunediversity. For example, a process called V(D)J (variable, diversity, joining) recombination is responsible for the combinatorial joining of the V, D an J DNA segments that encode the heavy or light chains of the Ig receptor (Lewis *et al.,* 1997). In order for this to occur, site-specific proteins initiate cleavage of DNA to produce DSBs that are required for recombination of the segments. Furthermore, DSBs are critical intermediates in immunoglobulin class-switch recombination, during meiotic recombination and yeast mating-type switching (Chaudhuri *et al.,* 2004; Keeney *et al.,* 1997; Schiestl *et al.,* 1992).

Regardless whether the DSBs are generated unintentionally or by design, it is absolutely critical that each DSB is repaired. If unrepaired, or if repaired incorrectly, these lesions may lead to undesirable outcomes for a cell including chromosomal rearrangements associated with further genomic instability and even cell death. Additionally, DSBs may lead to the activation of oncogenes or inactivation of tumour-supressor genes. To avoid such disastrous outcomes resulting from improper repair of DSBs, eukaryotes employ an army of DNA response and repair factors to ensure survival and propagation of healthy cells.

1.2 DNA Damage Response and Chromatin Remodeling

Upon the formation of a DSB, the DNA damage response pathways are activated, including sensor, transducer and effector proteins (Fig. 1.1) (Lobrich *et al.,* 2005). Central to the DSB response are ATM and ATR protein kinases, Tel1 and Mec1 respectively in *S. cerevisiae*, responsible for the transduction of the damage signal to numerous downstream effectors, particularly the cell cycle checkpoint proteins. Thus when a DSB forms, the cell halts progression through the cell cycle to allow time for repair. This regulation mechanism provides the cell with time to asses the extent of damage, and determines whether repair of damage is feasible or if apoptosis is the ultimate fate of the cell. Indeed, if a cell sustains a vast amount of damage it may be more beneficial to remove this cell from the population, through programmed cell death, mediated by pathways that rely on p53 signaling. On the other hand, if repair is a reasonable option, chromatin remodeling enzymes are required to alter DNA structure before the cell can proceed with restoration efforts. Figure 1.2 illustrates the events that occur following the formation of a double-strand break, and the possible outcomes for a cell.

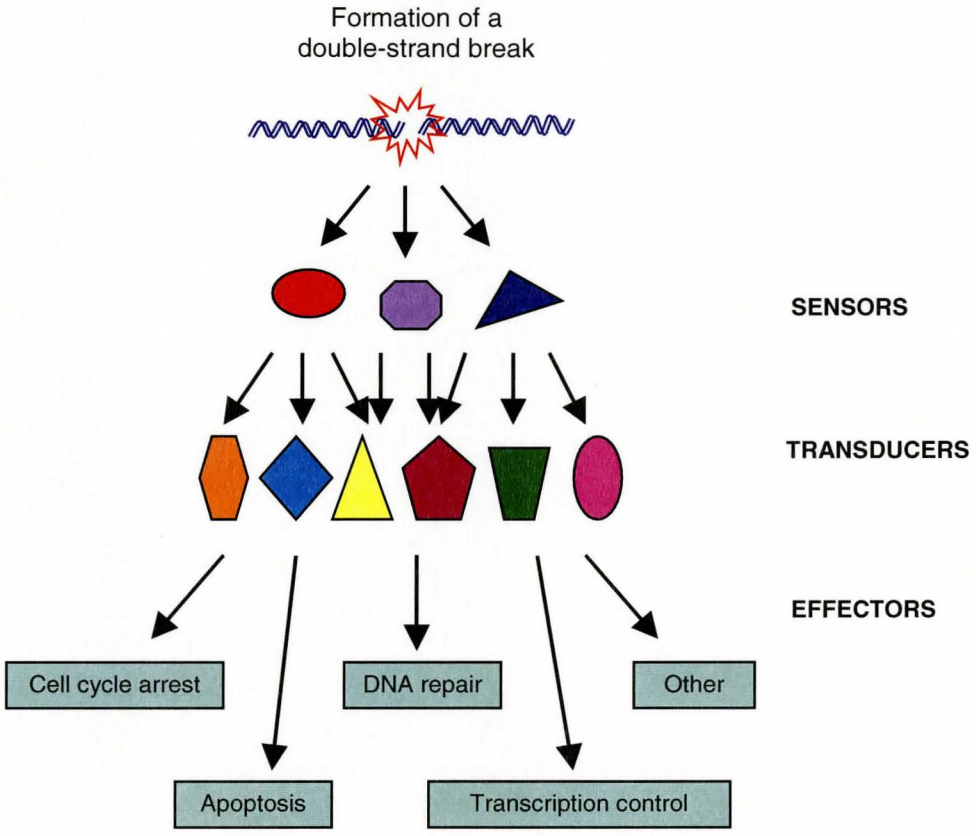

Figure 1.1 DNA damage response cascade. Upon the formation of a DSB, the DNA damage response pathways are activated, including sensor, transducer and effector proteins. A DSB can have downstream effects such as cell cycle arrest, DNA repair, transcriptional control and apoptosis.

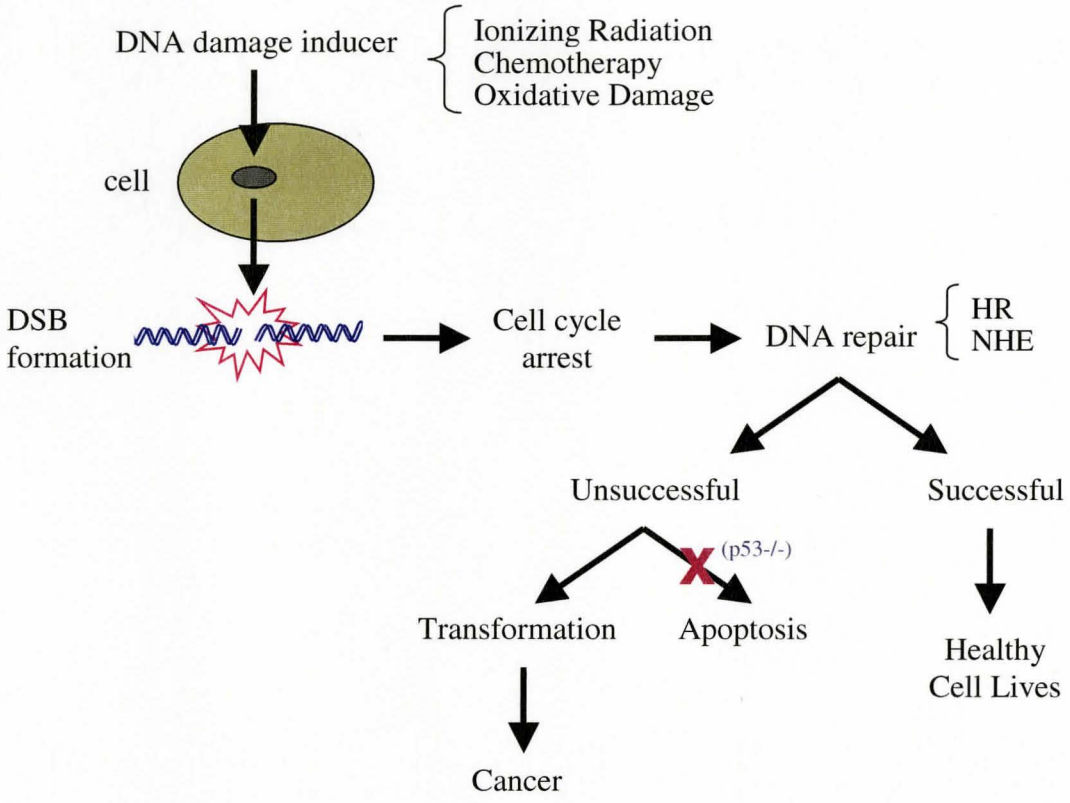

Figure 1.2 Events that occur after a DSB is formed and the possible outcomes for a cell. Upon the formation of a double-strand break, the cell cycle is arrested at which point the cell has the opportunity to repair the lesion via HR or NHEJ. If the repair is successful, the cell cycle progresses and the cell lives. If repair is unsuccessful, the cell may initiate apoptosis to remove this cell from the population. If however, the cell possesses defects in apoptotic or mitotic checkpoint pathways, neoplastic transformation may occur, leading to an uncontrolled growing population of cells (cancer).

Indeed, several studies demonstrated that one of the initial events following a DSB is the rapid phosphorylation of mammalian histone H2AX, homologous to yeast histone H2A (Rogakou *et al.,* 1999; Downs *et al.,* 2000). The phosphorylation of chromatin extends up to 50 kb from the site of the DSB in yeast (Downs *et al.,* 2004) and

is responsible for recruitment of various chromatin-modifying complexes which alter chromatin structure and subsequently provide access of DNA repair machinery (Ataian *et al.,* 2006). Interestingly, the ability of bacterial NHEJ proteins to function in yeast, despite the absence of histones in bacteria, suggests that their functions are more important than direct interactions between NHEJ and histones.

1.3 Repair of DNA DSBs

DNA DSBs can be repaired by one of two pathways, homologous recombination (HR) or non-homologous end-joining (NHEJ), two different repair mechanisms requiring a distinct set of proteins (Fig. 1.3). Both HR and NHEJ pathways exist in mammals as well as in simpler eukaryotes such as yeast *Saccharomyces cerevisiae*. In mammals the search for a homologous chromosome for repair via HR is not only inefficient but also more risky due to the extensive amount of DNA and repetitive sequences, which can be mistaken as the true homologous sequence. Therefore, mammalian cells predominantly utilize NHEJ to repair DSBs, while HR is the main repair pathway in lower organisms.

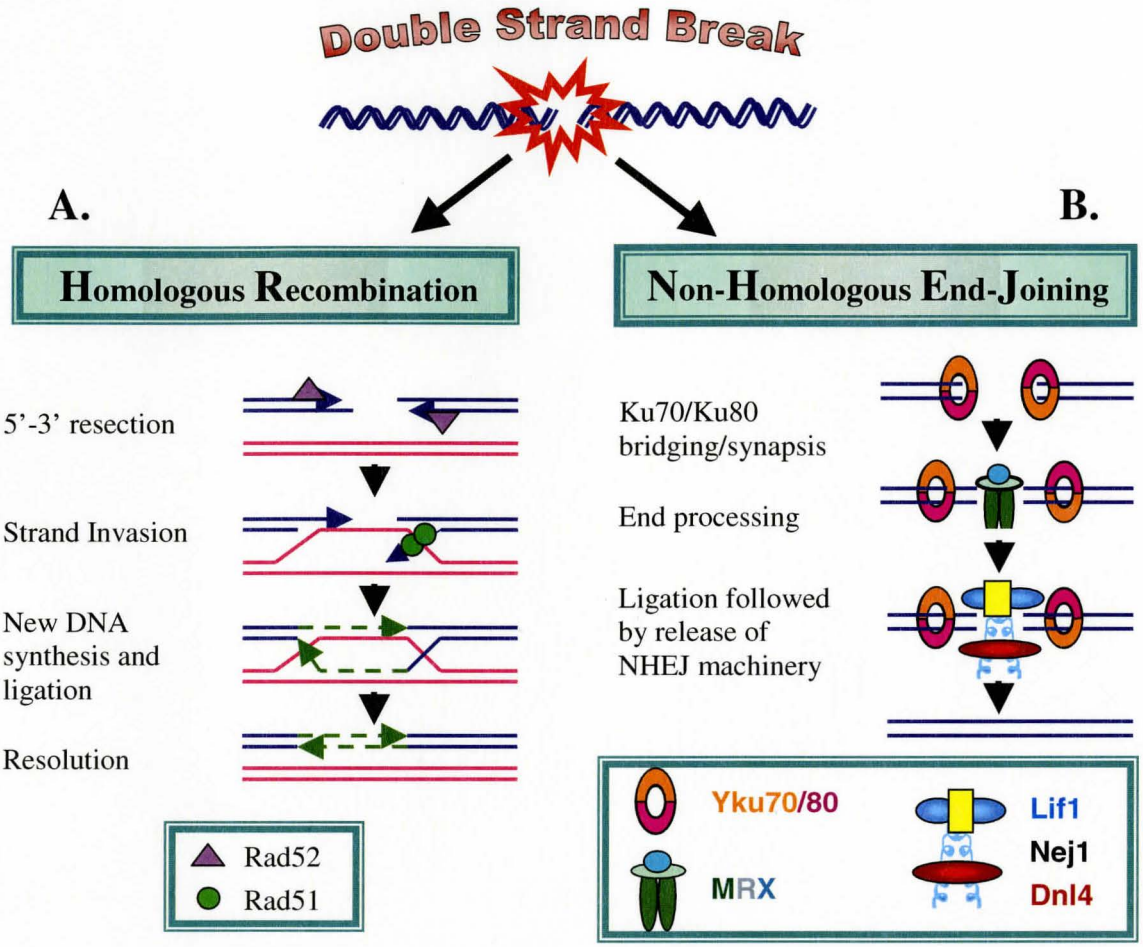

Figure 1.3 Two pathways for repairing double-strand breaks. (A) Homologous recombination relies on a homologous chromosome to serve as a template for repair. Because of this requirement HR is restricted to S, G_2 and M phases of the cell cycle. (B) Non-homologous end-joining repairs a DSB by ligating the two broken ends without the requirement for homology. Although NHEJ may operate in all stages of the cell cycle, it is predominant in the G_0 and G_1 phases.

In HR, repair of the DSB is dependent on the presence of a homologous sequence that serves as a template for DNA synthesis (Dudas *et al.*, 2004). The ideal donor is the sister chromatid to which a chromosome is paired following replication, or the homologous chromosome in diploid cells (Fig. 1.3). To initiate repair via HR, the ends are first resected in the 5' to 3' direction by nucleases, after which Rad51 monomers

polymerize at the 3' end and perform strand invasion. DNA polymerase extends the DNA, which is subsequently religated by a DNA Ligase to faithfully restore the sequence at the break. HR is the more accurate repair pathway, and is employed in the late synthesis, G_2 and M phases of the cell cycle, when a sister chromatid is available (Aylon *et al.,* 2004).

NHEJ, by contrast, does not require an undamaged partner to repair the DSB, and relies on little or no sequence homology (Moore *et al.,* 1996). Instead, a team of proteins is recruited to the DSB, where they stabilize, process and ligate the DNA ends (Daley *et al.,* 2005a). While NHEJ can ligate essentially any two DNA ends, restoration of the DNA is usually accompanied by insertions or deletions (Varga *et al.,* 2005). This is largely attributed to the action of processing enzymes that trim the DNA to generate compatible ends required for ligation. In the case of generating antibodies and antigen receptors, end processing plays a crucial role in creating unique V(D)J segments responsible for creating immunediversity. However, in the repair of DNA DSBs, errors associated with their repair are dangerous and believed to act as precursors in the development of cancer (Elliott *et al.,* 2002). Although NHEJ may operate in all stages of the cell cycle, it is predominant in the G_0 and G_1 phases.

1.4 Non Homologous End Joining (*Homo sapiens*)

Efficient NHEJ in *H. sapiens* relies on a minimum of 6 factors: Ku, DNA-PKcs, Artemis, Xrcc4, DNA Ligase IV and XLF which perform distinct roles during repair of DSBs. The initial step in NHEJ involves the binding the heterodimeric Ku70/Ku80 (Ku) complex to the DSB, to juxtaposition the DNA ends. Ku binds to the ends of the DNA with high affinity (Blier *et al.,* 1993), and this critical event serves to limit nucleolytic degradation that would otherwise lead to the loss of genetic information. Although not sequence specific, the DNA binding affinity of Ku (2.4 nM) (Blier *et al.,* 1993) can be attributed to the quasisymmetrical ring-like structure of the heterodimer, which allows for the DNA to weave through an opening formed by both subunits (Walker *et al.,* 2001). Moreover, the interior of the cavity is abundant in positively charged residues further stabilizing DNA-Ku interactions. Although Ku initially binds to DNA ends, it is able to translocate inward, allowing for subsequent repair proteins to gain access to the lesion.

Ku plays a central role in NHEJ, acting as a scaffold for downstream repair factors that are recruited to the site of the DSB. The DNA end bound Ku (regulatory subunit) recruits and activates DNA-dependent protein kinase catalytic subunit (DNA-PKcs), a serine-threonine kinase of the PIKK family responsible for phosphorylating several of the downstream NHEJ factors (Collis *et al.,* 2005). Ku together with DNA-PKcs constitute the DNA-PK complex (Dvir *et al.,* 1992), responsible for aligning and bridging the DNA ends (Spagnolo *et al.,* 2006).

Since the formation of DSBs is often accompanied with the formation of incompatible termini (3' phosphate or 5' hydroxyl), the DNA ends often require

processing prior to ligation (Ma *et al.,* 2005a). Processing is carried out by several factors, among them Artemis, which contains both endonuclease and exonuclease activities (Ma *et al.,* 2002). Artemis interacts with, and is phosphorylated by DNA-PKcs (Ma *et al.,* 2002), and has been implicated in processing a variety of DNA structures including gaps, flaps and hairpins (Ma et al., 2005b), often generated during V(D)J recombination.

The final ligation step requires the ATP-dependent DNA Ligase IV in a 2:1 complex with Xrcc4 (Junop *et al.,* 2000; Modesti *et al.,* 2003), in association with the recently discovered Cernunnos/XLF (described in section 1.8). Although Xrcc4-DNA Ligase IV is critical for end-joining, it is completely dependent on DNA-PKcs/Ku for its targeting to breaks. Additional studies indicated that Xrcc4 is required to stabilize DNA Ligase IV and stimulate its activity (Grawunder *et al.,* 1997; Modesti *et al.,* 1999). Characterization of the DNA binding activities of each protein indicated that DNA Ligase IV has weak DNA binding activity, while Xrcc4 binds nicked and linear DNA in a co-operative manner (Modesti *et al.,* 1999). The X-ray crystal structure of Xrcc4 revealed that this protein exists as an unusually elongated molecule (Junop *et al.,* 2000; Sibanda *et al.,* 2001; Meesala 2006). Xrcc4 exists as a dimer, with each monomer composed of an N-terminal 'head' domain containing a β-sandwich, followed by an extended C-terminal α-helical tail (Fig. 1.4). Additional studies illustrated that residues 1-200 of Xrcc4 are sufficient for DNA binding, and that phosphorylation abolishes its DNA binding activity (Modesti *et al.,* 1999), raising the possibility that this could serve as an end-joining regulation mechanism. Also worth noting is that this fragment (aa 1-200) can fully

complement the V(D)J defect in XR-1 cells *in vivo* (Mizuta *et al.,* 1997; Leber *et al.,* 1998).

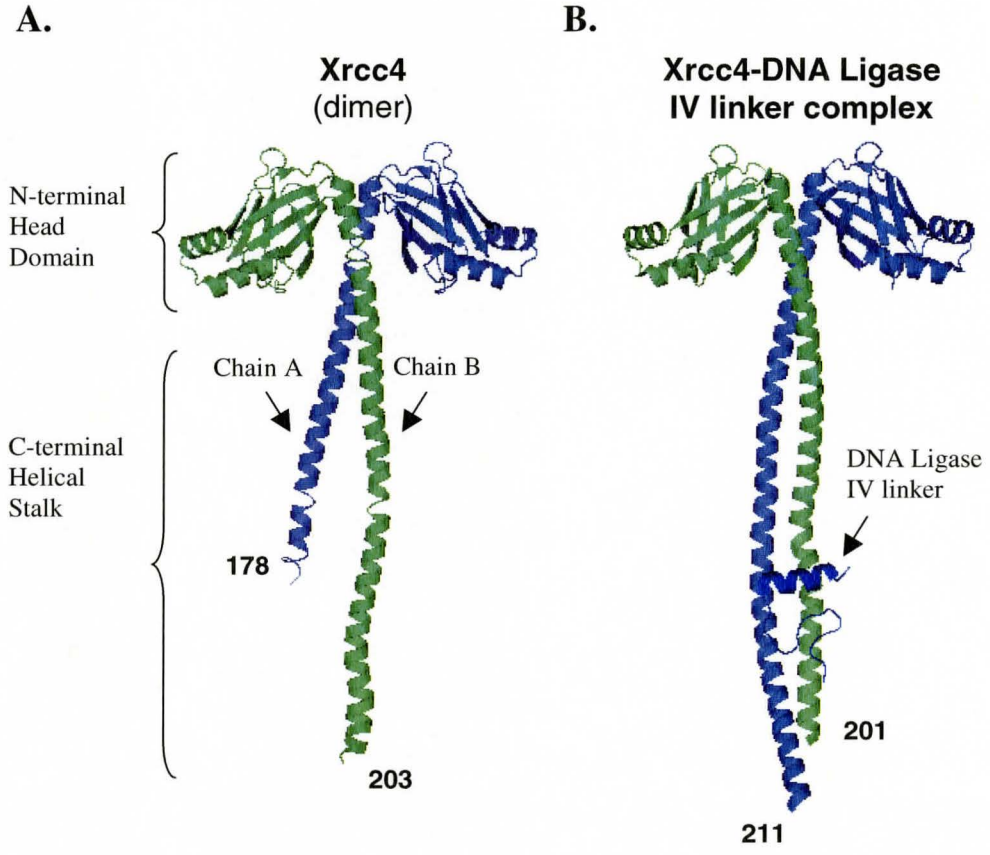

Figure 1.4 X-ray crystal structures determined of Xrcc4. (A) X-ray crystal structure of Xrcc4 (1-203) revealed that the protein exists as an unusually elongated dimer (Junop *et al.,* 2000). Xrcc4 is composed of an N-terminal globular 'head' domain, followed by an extended C-terminal α-helical stalk. (B) In 2001, the structure of Xrcc4 in complex with the Xrcc4 interaction domain of DNA Ligase IV was determined, illustrating a conformational change in the C-terminus upon binding of DNA Ligase IV (Sibanda *et al.,* 2001). PyMOL software was used to generate the pictures (De Lano *et al.,* 2002).

In addition to the abovementioned core NHEJ proteins, other factors are often required for accurate repair of DSBs. Polynucleotide kinase (PNK) which possesses both

5' DNA kinase and 3' DNA phosphatase activities (Karimi-Busheri *et al.,* 1999) is thought to be responsible for processing DSBs with 5'-OH termini (Chappell *et al.,* 2002). End processing enzymes such as DNA polymerases μ and λ of the Pol X family of polymerases, are implicated in gap filling during the rejoining of the two non-complementary ends (Daley *et al.,* 2005b). Notably, both of these polymerases associate with DNA-bound Ku (Mahajan *et al.,* 2002; Ma *et al.,* 2004). Another Pol X family member, terminal deoxynucleotidyl transferase, or TdT, has been implicated in the addition of non-templated nucleotides to DNA ends, and is thought to be responsible for increasing the diversity of junctions generated by V(D)J recombination. Additional factors such as the WRN and BLM helicases are suspected to unwind and process DNA ends (Chen *et al.,* 2003; Lengland *et al.,* 2002). FEN-1, an endonuclease capable of cleaving 5' flaps has also been implicated in NHEJ (Wu *et al.,* 1999).

A continued interest towards a comprehensive understanding of the NHEJ pathway has resulted in the design of clever biochemical experiments, which have revealed astonishing complexity regarding the mechanistic details of this DNA repair pathway. The current iterative model of NHEJ indicates that repair of the damaged duplex occurs through ligation of one DNA strand at a time (Ma *et al.,* 2005). This allows the cell to overcome a critical stage when the two strands are at risk of diffusing. Furthermore, DSBs with fully compatible overhangs do not require processing by a nuclease or polymerase, and thus, these lesions can be directly ligated using DNA Ligase IV (Ma *et al.,* 2005). Finally, current studies reveal the existence of a back-up NHEJ pathway (B-NHEJ), which operates with slower kinetics and is independent of DNA-PK

(Wang *et al.,* 2003). The DNA-PK dependent pathway (D-NHEJ), which also requires Xrcc4/DNA Ligase IV, is the most prevalent end-joining pathway and will simply be referred to as NHEJ in further discussion.

1.5 Non-Homologous End-Joining (*S. cerevisiae*)

Understanding of complex systems in higher eukaryotes (mammals) has been aided by the study of simple organisms such as the budding yeast *S. cerevisiae*. *S. cerevisiae* has been extensively characterized due to the high evolutionary conservation of many processes and pathways, and its ease of genetic manipulation. Since most of the mammalian NHEJ factors have functional counterparts in *S. cerevisiae*, it is not surprising that this organism serves as a valuable model for understanding mammalian end-joining.

As in mammalian NHEJ, Yku70/Yku80 is thought to initially recognize the DSB and perform DNA alignment and end-bridging functions (Fig. 1.3). Although *S. cerevisiae* lacks clear DNA-PKcs and Artemis homologues, the current model suggests that the end-bridging and processing functions are carried out by the MRX complex (Chen *et al.,* 2001), consisting of Mre11, Rad51 and Xrs2. This complex exists in a 2:2:1 stoichiometry respectively and is the only NHEJ component also involved in HR. Several studies indicate that the MRX complex is rapidly recruited to DSBs in a DNA damage dependent manner, and this is thought to occur simultaneously with the recruitment of Ku. As with mammalian Ku, Yku70 has been reported to be required for the recruitment of the ligation complex, Lif1-Dnl4. Once this complex is engaged,

consisting of a Lif1 dimer in association with the monomeric Dnl4, it is responsible for the ATP-dependent ligation of the damaged duplex. The X-ray crystal structure of Lif1 in complex with the tandem BRCT domains of Dnl4 has recently been elucidated (Dore *et al.,* 2006), illustrating remarkable structural similarity to the homologous human complex (Meesala 2006) (Fig 1.5).

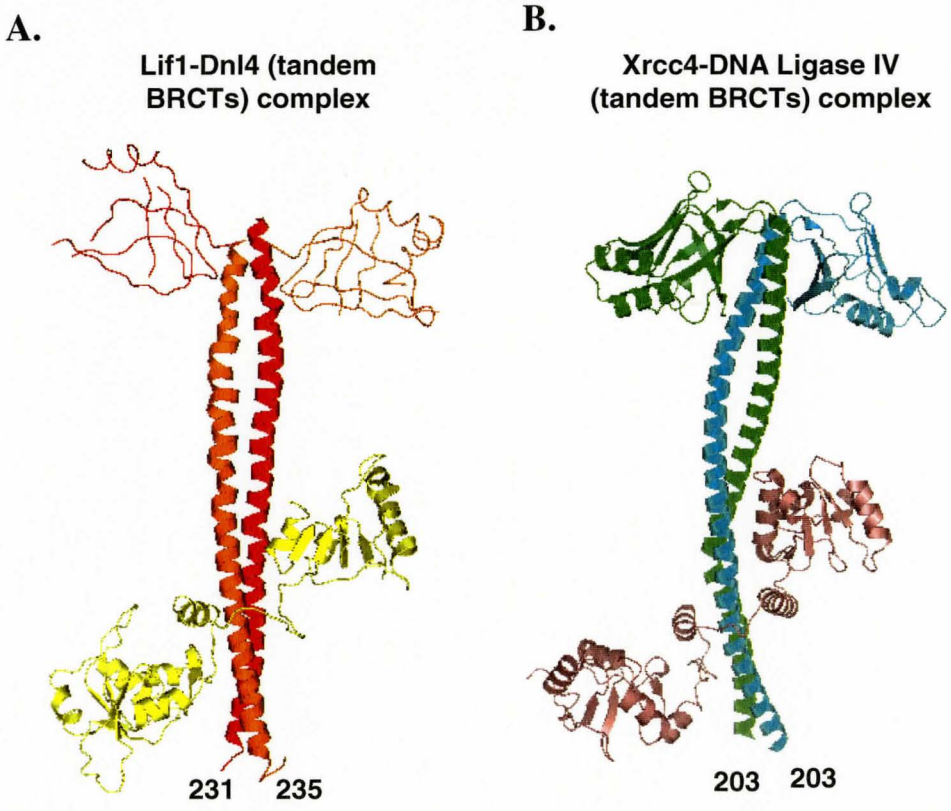

Figure 1.5 X-ray crystal structures of yeast and human ligation complexes. (A) The structure of Lif1 dimer (aa 1-246) (red/orange) in complex with Dnl4 (aa 680-944) (yellow) was determined to 3.9 Å resolution (Dore *et al.,* 2006). The structure shows remarkable structural similarity to the Xrcc4-DNA Ligase IV tandem-BRCT complex, shown in (B). A dimer of Xrcc4 (aa 1-203) (blue/green) in complex with DNA Ligase IV (aa 654-911) (pink) (Meesala 2006). One variation worth pointing out is the difference in orientation of the head domain with respect to the C-terminal helical stalk. PyMOL software was used to generate the pictures (De Lano *et al.,* 2002).

Several studies have implicated additional factors such as Rad27, a DNA structure specific endonuclease, and Pol4, a Pol X family member, in end-processing (Tseng *et al.,* 2004) and gap filling (Wilson *et al.,* 1999) respectively. As such, the NHEJ machinery represents a versatile task force for the proficient repair of not only complementary but also non-complementary DNA molecules.

It has become increasingly apparent that protein-protein interactions are important not only for stabilizing the end-joining machinery, which in turn is required for efficient repair of the damaged duplex. A recent study reported interactions between Yku80-Dnl4, Xrs2-Lif1 and Yku80-Mre11 (Palmobos *et al.,* 2005), highlighting the importance of protein-protein contacts within the NHEJ repair machinery (Fig. 1.6). Furthermore, recent biochemical data indicates that the proteasome is recruited to DSBs in a Pol4-dependent manner, suggesting that the proteasome may facilitate entry of NHEJ proteins to the lesion or be responsible for post-repair cleanup of end-joining factors (Krogan *et al.*, 2004).

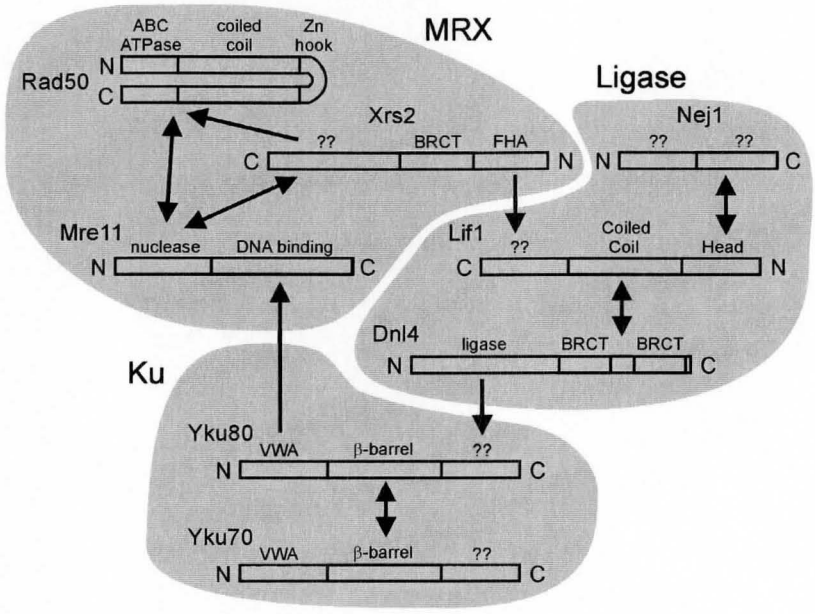

Figure 1.6 Illustrative summary of the interaction domains between the key NHEJ factors in *S. cerevisiae*. Interactions were determined using yeast-two hybrid analysis. Double-headed arrows represent interactions that were detected by reciprocal bait-prey pairings. Single-headed arrows represent interactions that were detected only in one orientation, with the arrow pointing to the prey. Arrows point to the domain responsible for the interactions, and shading indicates stronger interactions. Illustration is adapted from Palmbos *et al.,* 2005.

In addition to their function in end-joining, several NHEJ proteins have been proposed to play a role in preventing the fusion of telomeres, cells' natural DBSs. Specifically, Yku70/Yku80, the MRX complex and Nej1, are all speculated to carry out additional roles in maintaining chromosome termini (Fisher *et al.,* 2005; Foster *et al.,* 2006; Liti *et al.,* 2003). The function of these proteins in end-joining drastically contrasts their proposed role in preventing the fusion of telomeres, posing an interesting biological paradox. Nej1, the most recent addition to the *S. cerevisiae* NHEJ repair machinery, is the focus of this research and thus will be described in detail below (section 1.6).

1.6 Non-Homologous End-Joining Regulator 1, Nej1

Yku70/Yku80, MRX and Lif1/Dnl4 were once thought to constitute the essential core components of the *S. cerevisiae* NHEJ pathway. It was not until 2001, that Ooi *et al.* reported an elegant study describing the identification of a novel NHEJ factor, called non-homologous end-joining regulator 1, otherwise known as Nej1 (Ooi *et al.,* 2001). The study presented compelling evidence indicating that Nej1 is required for NHEJ to the same extent that Dnl4 and Yku80, demonstrating that it is a critical component of this repair pathway.

Initial characterization of Nej1 revealed that this factor strongly interacts with Lif1 independently of Dnl4, and furthermore illustrated that formation of the Lif1-Dnl4 complex does not require Nej1. While amino acids 170-342 of Nej1 were shown be necessary and sufficient to interact with the globular head domain of Lif1 (aa 2-191), interactions with other NHEJ factors have not been demonstrated (Valencia *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001). The boundaries of these interaction domains were recently refined to include aa 173-432 of Nej1 and aa 1-157 of Lif1 (Fig. 1.7) (Deshpande *et al.,* 2007). Although an early report indicated that Nej1 might be responsible for nuclear localization of Lif1 (Valencia *et al.,* 2001), subsequent experimental evidence disproved this as a potential role for Nej1 (Kegel *et al.*, 2001; Ahnesorg *et al.,* 2006a).

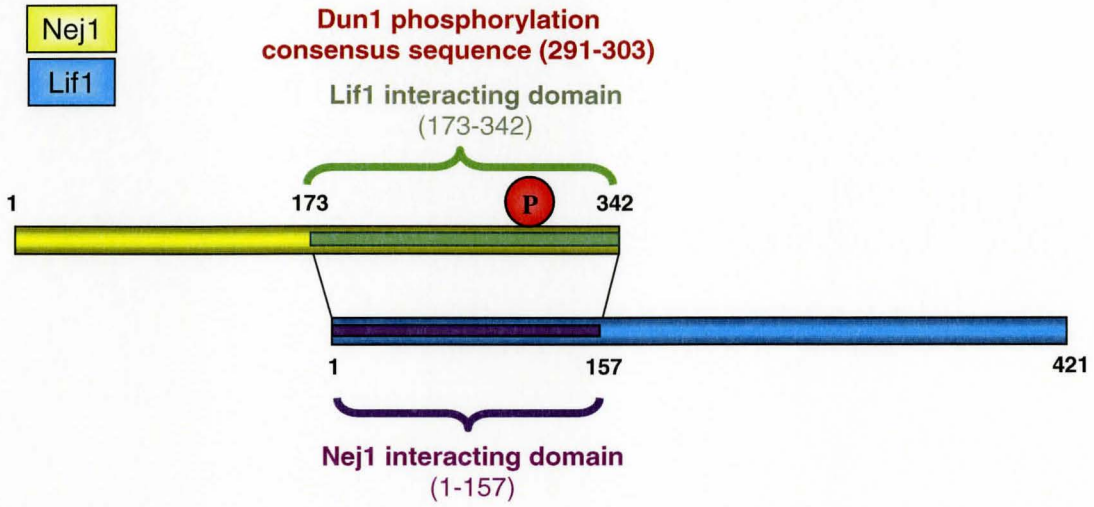

Figure 1.7 Functional domains of Nej1. The Lif1 interaction domain has been mapped out to the C-terminus of Nej1 (aa 173-342). Conversely, the N-terminus of Lif1 (aa 1-157) interacts with Nej1 (Ooi *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007).

Additional studies presented data to suggest that this nuclear, 39 kDa protein may function as a regulatory component of NHEJ. Indeed, several research groups demonstrated that expression of Nej1 is repressed in diploid *S. cerevisiae* via the Mata1-Matα2 transcriptional repressor which binds to a consensus sequence located upstream of the Nej1 reading frame (Valencia *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001). Thus, diploid yeast are deficient in Nej1, resulting in a clever control mechanism that down-regulates NHEJ, allowing for the damage to be repaired efficiently and accurately by homology driven repair (Fig. 1.8).

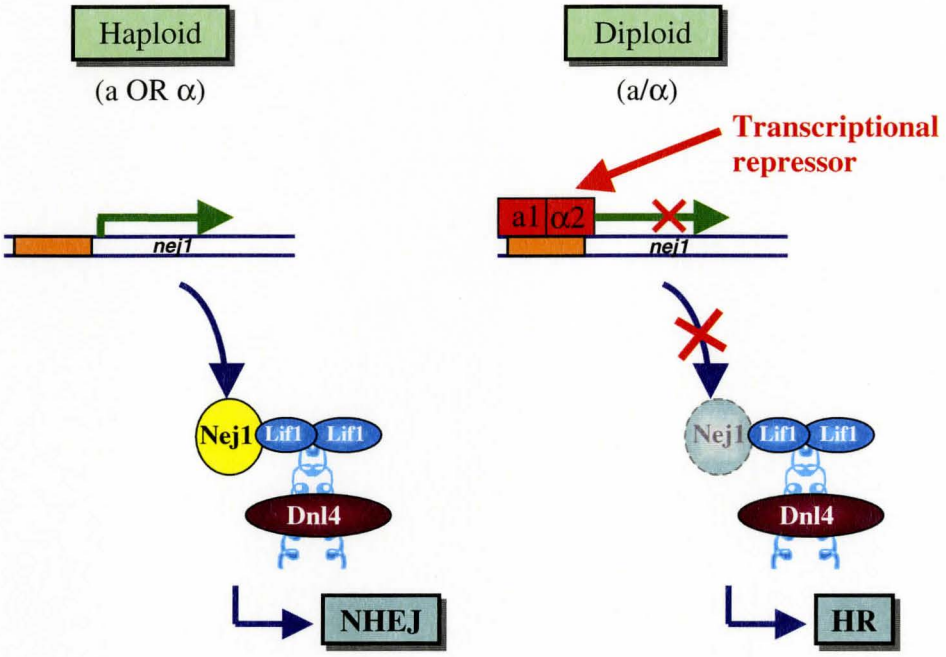

Figure 1.8 Expression of Nej1 is down-regulated in diploid *S. cerevisiae*. Haploid yeast (a or α) are unable to form the Mata1/Matα2 transcriptional repressor, allowing for Nej1 to be expressed and for the cells to utilize NHEJ to repair a DSB. In contrast, diploid yeast (a/α) contain this transcriptional repressor, leading to a down-regulation of Nej1 expression, allowing the cells to make use of a homologous chromosome to repair a DSB using homologous recombination.

To date a clear biochemical function has not been assigned to Nej1, despite the fact that numerous groups reported that yeast lacking Nej1 are defective in end-joining (Ooi *et al.,* 2001; Valencia *et al.,* 2001; Kegel *et al.,* 2001). In addition, Nej1 was shown to be required for efficient NHEJ even if Lif1 is over-expressed (Kegel *et al.,* 2001), indicating that each protein is distinct and both are required for efficient NHEJ. Further studies by Liti and colleagues demonstrated that Nej1 might be implicated in preventing telomere fusions in the absence of telomerase (Liti *et al.,* 2003). This suggests the

possibility that Nej1 may have more than one function, much like the dual functions of other NHEJ factors such as Ku and MRX, in end-joining and in prevention of chromosome fusions.

A recent study by Ahnesorg and Jackson demonstrated that Nej1 is rapidly phosphorylated in response to DNA damage, in a manner that relies on a cascade of DNA damage checkpoint kinases: Mec1, Rad53 and Dun1 (Ahnesorg *et al.,* 2006a). Using a mutational approach, the group identified a consensus Dun1 phosphorylation sequence in the C-terminus of Nej1 (aa 291-303) (Fig. 1.7). Interestingly, alanine substitutions at the putative phosphorylation sites, specifically S297/8, resulted in a decrease in NHEJ activity. Since phosphorylation occurs in the C-terminus of Nej1, which is responsible for interacting with Lif1, this suggests that phosphorylation may play a role in regulating interaction between these two proteins, and thereby NHEJ.

A study by Cliften and colleagues revealed that Nej1 is highly conserved within the *Saccharomyces* genus, with Nej1 orthologues being found in at least four other *Saccharomyces* species (Cliften *et al.,* 2003). Interestingly, a mammalian homologue of Nej1 has recently been identified and is described in detail in the following section (section 1.7).

1.7 Cernunnos/XLF

Recently, two groups independently discovered a novel NHEJ factor, Cernunnos/XLF. The hunt towards the identification of this protein came about after Dai and colleagues presented convincing evidence that NHEJ and V(D)J recombination

require an additional and yet uncharacterized factor (Dai *et al.,* 2003). Following this initial report, using yeast-two hybrid screening, one group identified an Xrcc4 interacting factor as a yet uncharacterized 33 kDa protein, well conserved throughout the vertebrate species (Ahnesorg *et al.,* 2006b). Sequence structure comparisons indicated that this protein is predicted to display a similar structural organization to Xrcc4, containing an N-terminal 'head' domain followed by a C-terminal coiled-coil domain (Fig. 1.9). Although primary sequence comparisons between this novel protein and Xrcc4 revealed a low level of sequence similarity, based on the above findings, the group named the protein Xrcc4-like factor or XLF.

The second research group employed an alternate approach. Buck and colleagues examined patients with general DNA repair defects, characterized by increased sensitivity to ionizing radiation and defective V(D)J recombination (Buck *et al.,* 2006). Subsequent analysis revealed that the patients carried various point or frameshift mutations in a protein that they named Cernunnos. Cernunnos and XLF both code for the same 33 kDa protein, and thus from this point onwards XLF/Cernunnos will simply be referred to as XLF.

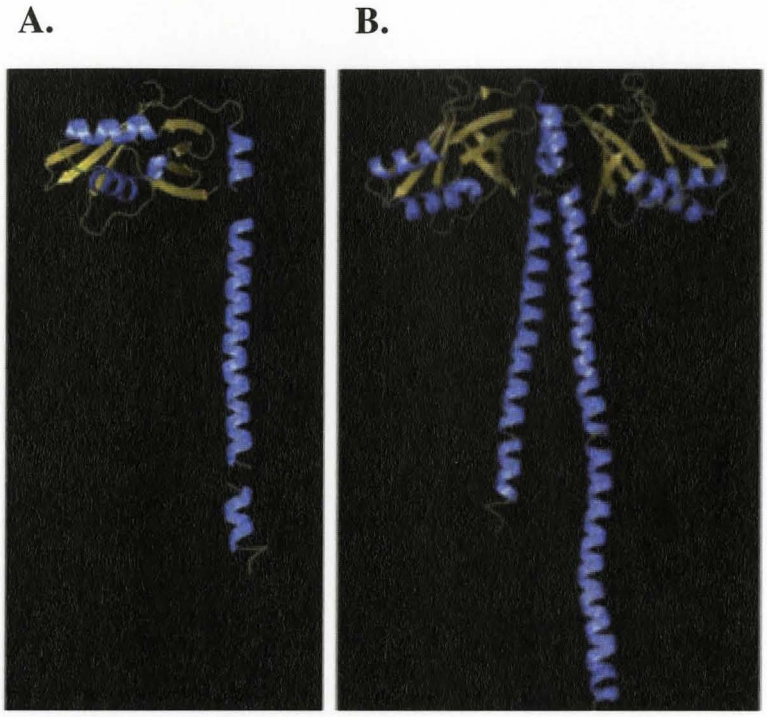

Figure 1.9 Model for the structure of XLF based on X-ray crystal structure of Xrcc4. The model of XLF in (A) was predicted based on secondary structure similarities to Xrcc4 (B). Figure is adapted from Ahnesorg *et al.,* 2006b.

Despite any obvious sequence similarity, using iterative BLAST and HCA manual alignment of multiple distantly related sequences, Callebaut *et al.* were recently able to provide compelling evidence to suggest that XLF may represent the *H. sapiens* homologue of *S. cerevisiae* Nej1 (Callebaut *et al.,* 2006). Consistent with this idea, both proteins appear to interact with the same functional components, XLF with Xrcc4-Ligase IV in humans (Ahnesorg *et al.,* 2006b; Callebaut *et al.,* 2006), and Nej1 with Lif1 in *S. cerevisiae* (Ooi *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001).

Initial functional characterization demonstrated that XLF interacts with DNA in a non-specific way, binding both linearized and circular DNA with similar efficiencies (Hentges *et al.,* 2006). Additional experiments revealed that XLF possesses a length-dependent DNA binding property, similar to what has been reported for Xrcc4 (Modesti *et al.,* 1999). Importantly, XLF was found to specifically stimulate DNA end-joining by Xrcc4-DNA Ligase IV complex *in vitro* (Ahnesorg *et al.,* 2006b; Lu *et al.,* 2007), suggesting that XLF may play a role in the final ligation step of NHEJ. Remarkably, recent experimental data provided significant insight into a potential role of XLF in end-joining. *In vitro* biochemical assays illustrated that XLF promotes ligation of mismatched and non-complementary DNA ends, ~8 to 150 fold depending on the type of mismatch (Tsai *et al.,* 2007). This activity was dependent on Ku and did not require additional processing factors such as nucleases or polymerases. In contrast to Xrcc4, which requires the N-terminal 200 residues for wild type activity, an autopsy on a patient expressing a truncated XLF (aa 1-196) revealed a common malformation of the human cerebral cortex, called polymicrogyria (PMG) (Cantagrel *et al.,* 2007), indicating that the C-terminus of XLF is critical for its function. Furthermore, targeted deletion of XLF exons 4 and 5 in murine cells resulted in sensitivity to ionizing radiation, DNA fragmentation and chromosomal abnormalities, as well as defective V(D)J recombination (Zha *et al.,* 2007). Collectively, the experimental evidence indicates that XLF, the newest addition to the components of the mammalian NHEJ pathway, plays a critical role in the repair of DNA DSBs.

1.8 Double-strand breaks and Disease

It has been widely accepted that unrepaired double-strand breaks pose major threats not only to genomic integrity but also cellular viability. For that reason, it is imperative that cells restore each DNA DSB efficiently and accurately, relying on the efforts of the NHEJ pathway. Thus, it is not surprising that this class of DNA repair proteins is extremely important for maintaining a healthy cell and organism. This is apparent in the number of cancers and immunological disorders that have been documented due to defects in various NHEJ factors.

For example, murine cells deficient in Ku70, Ku80, Xrcc4, and DNA Ligase IV, the four evolutionarily conserved NHEJ factors, present with premature senescence, sensitivity to DSB inducing agents, and severely impaired V(D)J recombination (Bassing *et al.,* 2002). Inactivation of any of these four genes in mice leads to multiple defects including growth retardation, sensitivity to ionizing radiation and severe combined immunodeficiency (SCID). SCID is characterized by the absence of mature B and T lymphocytes due to impaired V(D)J recombination (Nonoyama *et al.,* 1996). DNA-PKcs deficiency also results in a SCID phenotype and variable IR sensitivity, although, no growth or developmental abnormalities have been documented (Bosma *et al.,* 1983; Taccioli *et al.,* 1998). Furthermore, both Ku70/Ku80, and DNA-PKcs deficiencies result in increased levels of telomeric fusions (Bailey *et al.,* 1999). Patients defective in Artemis have been characterized with the severe radiation sensitivity in addition to SCID, categorized as RS-SCID (Moshous *et al.,* 2000). It is important to point out that absence of either Xrcc4 or DNA Ligase IV results in embryonic lethality caused by severe

apoptosis of post mitotic neurons (Barnes *et al.,* 1998; Frank *et al.,* 2000; Gao *et al.,* 2000).

In addition to many disorders being caused by deficiencies of NHEJ factors, mutations in many of these proteins also have severe consequences. For example, LigIV syndrome, a rare autosomal recessive disorder caused to hypomorphic mutations in DNA Ligase IV, is characterized by developmental delay, radiosensitivity, immunodeficiency and impaired end-joining (Enders *et al.,* 2006). The recently identified XLF was discovered mutated in patients exhibiting microcephaly and immunodeficiency (Buck *et al.,* 2006). Mutations in this critical component of NHEJ results in increased sensitivity to ionizing radiation, along with growth retardation, defective V(D)J recombination and defective end-joining, a similar phenotype to LigIV syndrome. In contrast to both DNA Ligase IV and XLF, no patients have been reported with mutations in either Ku70/Ku80 or Xrcc4, revealing a fundamental role of these proteins in NHEJ and in human health.

It is a well-established fact that defects in NHEJ proteins increase genomic instability. Although excessive DNA damage typically leads to apoptosis, a cell can acquire a growth advantage over the population, leading to selective proliferation of its descendents. The mechanisms that result in cancerous cells are not only countless but also variable. Frequently however, a general inability to restore the integrity of the duplex results in genomic translocations that lead to amplification of proto-oncogenes and subsequently transformation of healthy cells into cancerous cells. Moreover, it has been reported that Xrcc4/p53-, DNA Ligase IV/p53- and Artemis/p53- deficient mice develop medulloblastomas (Yan *et al.,* 2006; Lee e*t al.,* 2002; Rooney *et al.,* 2004).

1.9 Thesis Objectives

The main purpose of this study is to advance the current understanding of the role of Nej1 in the NHEJ DNA repair pathway. Although Nej1 is the most recent addition to the NHEJ machinery, and has been shown to be required for end-joining, a clear biological function has not been assigned for Nej1. Initial studies on Nej1 revealed that it interacts with another NHEJ protein, Lif1, and suggested that its primary role may be to regulate NHEJ in *S. cerevisiae*. Despite these proposed roles, this study addresses the hypothesis that Nej1 plays a direct role in the repair of DSBs.

Thus, the goal of this research is to characterize the function and structure of Nej1, and subsequently compare and contrast it with the current knowledge of its mammalian homologue XLF, its interacting partner Lif1, and the corresponding human homologue of Lif1, Xrcc4. Biochemical characterization of Nej1, both functionally and structurally should provide information regarding the physical properties of Nej1 and in turn reveal invaluable insights into the mechanism of this protein in end-joining.

CHAPTER 2 – MATERIALS AND METHODS

2.1 General Materials

Nej1 and Lif1 protein expression vectors (MJ4108 and MJ4102 respectively) were previously constructed by Beena Mistry and kindly provided by Dr. Junop. The pSLO103 vector was generously donated by Dr. Jef Boeke (John Hopkins University). The pLIC vectors were a kind gift from Dr. Stephen Bottomley (Monash University).

Pfu and Taq Polymerases were purified in our lab (see Appendix II). All other restriction enzymes purchased form New England Biolabs (Mississauga, ON). Primers were obtained from Mobix (McMaster University, Hamilton, ON) or Invitrogen (Burlington, ON). Sequencing was preformed at Mobix, McMaster University (Hamilton, ON). Tobbaco Etch Virus (TEV) protease was purified in our lab (see Appendix II).

A plasmid miniprep kit was purchased from Sigma (Oakville, ON) or a manual alkaline lysis method was used to extract plasmid DNA from *E. coli*. MinElute PCR purification spin columns and the DNA agarose gel extraction kit were obtained from Qiagen (Mississauga, ON).

All *E. coli* plasmid propagating cell lines, including Top10 and DH5α were originally purchased from Invitrogen (Burilngton, ON). Rosetta (DE3) and BL21 Star cells containing pLysRARE, protein expression cell lines were obtained from Novagen (Mississauga, ON). All original cell lines were subsequently cultured in our lab to create

permanent large scale lab stocks. All *S. cerevisiae* strains, including BY8920 and BY4741 were a kind gift from Dr. Boeke (John Hopkins University).

Chemicals for the preparation of LB media and buffer solutions were purchased from Bioshop Canada, Ltd. (Burlington, ON). Ampicillin and kanamycin were obtained from Fisher Scientific (Ottawa, ON), chloramphenicol was purchased from Calbiochem (San Diego, CA) and IPTG was obtained from Bioshop Canada, Ltd. (Burlington, ON). Trypsin and chymotrypsin, used for partial proteolysis experiments, were purchased from Winley-Morris Co. (Montreal). Protease inhibitors, including PMSF and pepstatin A, benzamidine and leupeptin were purchased from Bioshop Canada, Ltd. (Burlington, ON). Protein concentrators, 30K, 10K and 3K, were obtained from Pall Corporation (Mississauga, ON).

For all chromatographic procedures, elution profiles were monitored with an in-line UV-900 detector at 280 nm on an ÄKTA explorer workstation, both purchased from Amersham Biosciences/GE Healthcare (Baie d'Urfe, QC). Protein concentrations were determined by the Bradford assay using the Bradford dye purchased from BioRad (Mississauga, ON).

The 35 bp DNA substrate used for EMSA studies was a generous gift from Dr. Junop. PNK enzyme was purchased from Fermentas (Burlington, ON) and the [γ-^{32}P] ATP was obtained from Perkin Elmer (Woodbridge, ON), and both kindly donated by Dr. Li from McMaster University (Hamilton, ON). For EMSA studies, the gels were visualized and quantitated using ImageQuant, which was a generous gift from Dr. Li. To

analyze co-operativity a program called Sigma Plot was used, supplied by Dr. Wright at McMaster University (Hamilton, ON).

For circular dichroism experiments, the spectra were recorded and visualized using the program Jasco J 600, a kind gift from Dr. Ananthanarayanan at McMaster University (Hamilton, ON). For analytical ultracentrifugation experiments the ultracentrifuge was obtained from Beckman Coulter (Mississauga, ON), and data analysis was performed using the programs SEDNTERP and Optima XL-A/XL-1 analysis software within Origin version 6.0 (Microcal), courtesy of Dr. Yang.

N-terminal amino acid sequence of peptide fragments derived from Nej1 proteolysis was generated using a 492 gas-phase/pulsed-liquid Procise automated sequencer (Applied Biosystems) at McGill University (Montreal, Canada) and analyzed using the Model 610A data analysis program. MS/MS identification of proteins was performed at McMaster University (Hamilton, ON).

Antibodies towards Nej1 were generated at the Central Animal Facility at McMaster University (Hamilton, ON). The donkey anti-rabbit IgG antibody was a kind gift from Dr. Haslam, and was originally purchased from Amersham Biosciences/GE Healthcare (Baie d'Urfe, QC). Western Lightning detection system was purchased from Perkin Elmer (Woodbridge, ON), and the XRA film used to expose the immunoblots a generous gift from Dr. Haslam, originally obtained from Kodak (Mississauga, ON).

For crystallography experiments, The Classics screen and the MbClassI screen were purchased from Nextal/Qiagen (Mississauga, ON). The trays and screw on lids were also obtained from Nextal/Qiagen (Mississauga, ON). The Membrane kit was

purchased from Sigma Aldrich (Mississauga, ON). The Additive Screens, I, II and III, were obtained from Hampton Research (Aliso Viejo, CA). The 15 bp DNA used for crystallization studies was purchased from Invitrogen (Burlington, ON). The X-ray diffraction images were visualized using Crystal Clear software package (Rigaku/MSC Ltd., La Jolla, CA).

2.2 Construct Design and Site Directed Mutagenesis

The Nej1 expression vector (pMJ4108) was generated by cloning full-length *S. cerevisiae NEJ1* into NcoI/XhoI sites of pPROEX-HTb (Gibco BRL), creating a TEV cleavable His_6 fusion. The Lif1 expression vector (pMJ4102) was generated by cloning full-length *S. cerevisiae LIF1* into NcoI/XhoI sites in pPROEX-HTb, creating a TEV cleavable His_6 fusion. These two constructs were constructed by a previous student, Beena Mistry, and generously donated for studies described here.

Nej1 C-terminal truncations, specifically Nej1 1-269, 1-252 and 1-244 (pMJ4356, pMJ4367 and pMJ4358 respectively) were created by performing site directed mutagenesis on pMJ4108 to incorporate appropriate stop codons. Likewise, alanine substitutions were introduced in Nej1 by site directed mutagenesis on pMJ4108 to create Nej1 (K283A, K285A, K287A) (K283A, K285A, K287A), Nej1 (K291A, K293A, R295A), and Nej1 (K332A, R333A, K334A) (pMJ4318, pMJ4323 and pMJ4319 respectively). These same Nej1 truncations and point mutations were introduced via site directed mutagenesis into a yeast compatible vector, pSLO103, for *in vivo* analysis. A

typical site directed mutagenesis reaction and the program are outlined in Table A.II in Appendix I.

To verify whether amplification was successful, 10 µl of the site directed mutagenesis reaction was subjected to electrophoresis on a 1 % agarose gel. The DNA was visualized by ethidium bromide staining and detected using a UV light detection system. Successfully amplified reactions were treated with 1 µl of DpnI enzyme for 1 hour at 37°C to digest the original template DNA. Subsequently, 2-5 µl of DNA was used to transform 50 µl of Top10 or DH5α CaCl2 competent cells using the heat shock method (Frobisher 1968). Following transformation, the reaction was briefly spun down to obtain a cell pellet which was resuspended in 100 µl of LB and plated on a LB-Amp plate. The resulting single colonies were cultured overnight in a shaker in 5 mL of LB-Amp at 37°C. The DNA was extracted using the alkaline lysis miniprep method and the quality of the DNA was confirmed by subjecting the plasmid DNA to electrophoresis on a 1 % agarose gel. Finally, the plasmids were verified by sequencing to ensure that the desired mutations were successfully introduced, and that no other mutations were introduced unintentionally.

Nej1 N-terminal truncations, specifically Nej1 170-342 and Nej1 268-342, were created by PCR amplifying the desired *NEJ1* fragments, and inserting them into a pLIC-Nus vector using a method previously described (Cabrita *et al.,* 2006), to create pMJ4521 and pMJ4524 respectively. In brief, ligation independent cloning vectors (pLIC) were digested with SacII restriction endonuclease, after which the linearized plasmid was subjected to electrophoresis on a 1 % agarose gel and subsequently purified. The purified

linearized vector was treated with T4 DNA Polymerase (30 minutes, room temperature) in the presence of dTTP to create 5' overhangs, complementary to the incoming insert. The enzyme was inactivated by incubation at 75°C for 20 minutes.

Meanwhile, the insert was prepared by performing PCR on pMJ4108 using primers designed to contain a start codon at an appropriate position in addition to a short stretch of DNA complementary to the destination vector. The PCR products obtained were purified using a spin column (exclusion limit 5 kb) and subsequently treated with T4 DNA Polymerase (30 minutes, room temperature) in the presence of dATP to create 3' overhangs, complementary to the vector containing 5' overhangs. Following treatment, T4 DNA Polymerase was inactivated by incubation at 75°C for 20 minutes. The vector and insert were mixed in a 1:2 molar ratio and the DNA was allowed to anneal overnight at room temperature. Subsequently, 2-3 µl of the reaction was used to transform 50 µl of Top10 cells via the heat shock method of DNA transformation, and all cells were plated on LB-Amp plates. The resulting colonies were cultured overnight in 5 mL of LB-Amp and the DNA was extracted using the alkaline lysis miniprep method. The DNA obtained was subjected to electrophoresis on a 1 % agarose gel where the size of the resulting plasmid determined whether the annealing reaction was successful and the plasmid of interest was indeed obtained. Plasmids of the expected size were verified by sequencing.

The Nej1-Lif1 co-expression vector (pMJ4216) was generated by PCR amplifying the promoter-*NEJ1*-terminator region from pMJ4099, containing SwaI restriction sites, and ligating this fragment into the PshAI site of pMJ4102. Ligation

reactions were transformed into Top10 or DH5α CaCl2 competent cells using the heat shock method. Resulting colonies were cultured and the DNA was extracted using alkaline lysis miniprep method. A diagnostic restriction digest was performed on the resulting DNA samples to verify that the PCR fragment was successfully ligated in.

All constructs used in this study are summarized in Table A.I in Appendix I. All primer sequences appear in Table A.III in Appendix I. All constructs were fully verified by DNA sequencing. In the case of pMJ4216, E130G and K221Δ mutations resulted during cloning.

2.3 Protein Expression

Recombinant His_6-tagged Nej1 proteins (full-length Nej1, Nej1 1-269, 1-252 and 1-244, 170-342, 268-342 Nej1 (K283A, K285A, K287A), Nej1 (K291A, K293A, R295A) and Nej1 (K332A, R333A, K334A)) were expressed in Rosetta (DE3) *E. coli* (from pMJ4108, pMJ4356, pMJ4367, pMJ4358, pMJ4318, pMJ4323 pMJ4319, pMJ4521 and pMJ4524 respectively), under the control of the T7 RNA polymerase promoter. A saturated 10 mL overnight culture was used to inoculate 1 L of LB containing 100 µg/ml ampicillin and 25 µg/ml chloramphenicol. Cultures were grown at 37°C with vigorous shaking to an optical density at 600 nm (OD_{600}) of 0.5, and induced for 4 hours at 25°C in the presence of 1 mM IPTG. Cells were harvested by centrifugation at 2000 g, frozen in liquid nitrogen and stored at -80°C.

Recombinant full-length His_6-tagged Lif1 was expressed from pMJ4102, grown and induced as described above for full-length Nej1, with the exception that induction was carried out at 37°C for 3 hours.

Recombinant full-length Lif1-Nej1 protein complex was expressed in BL21 Star cells containing pLysSRARE from pMJ4216, under the control of the T7 RNA polymerase promoter. A saturated 10 mL overnight culture was used to inoculate 1 L of LB containing 100 µg/ml kanamycin and 25 µg/ml chloramphenicol. Cultures were grown at 37°C with vigorous shaking to an OD_{600} of 0.5 and induced for 3 hours at 37°C in the presence of 1 mM IPTG. Cells were harvested by centrifugation at 2000 g, frozen in liquid nitrogen and stored at -80°C.

2.4 Protein Purification

2.4.1 General Procedures for Purification of all Proteins

Prior to lysis, 1 mg of DNaseI was added to resuspended cells along with protease inhibitors (PMSF, benzamidine, pepstatin A and leupeptin) to a final concentration of 1 mM. The sample was passed through the French Press at 20 000 psi, 4 times to efficiently lyse the cells, and subsequently all four protease inhibitors were added to a final concentration of 1 mM. The lysate was clarified by centrifugation (40 min, 48344 g, 4°C), treated with leupeptin and benzamidine, at 1 mM final concentration, and filtered prior to carrying out column chromatography. During chromatographic procedures, peak fractions were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis

(SDS-PAGE) and protein concentrations were determined using the Bradford assay (Bradford 1976).

2.4.2 Purification of full-length His_6-tagged Nej1

Cell pellets were resuspended in Nickel-Column Buffer A containing 20 mM 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) pH 7.0, 2 mM Beta-mercapto-ethanol (βME), 5% glycerol, 500 mM KCl, 10 mM imidazole and 0.06% N-lauryl-N, N-dimethylamine N-oxide (LDAO). Following lysis using French Press, the filtered clarified lysate was applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating Nickel Affinity column, pre-equilibrated with Nickel-Column Buffer A. Successive washes were performed at 15 mM, 30 mM and 45 mM imidazole, and Nej1 was eluted in a 210 mM imidazole fraction. The sample was diluted to 150 mM KCl in the presence of 20 mM Tris pH 7.0, 5 mM dithiothreitol (DTT), 1 mM ethylenediamine tetraacetic acid (EDTA) and 5% glycerol, filtered and injected onto a SP Sepharose column pre-equilibrated at 150 mM KCl. The resin was washed at 200 mM, 250 mM KCl, 400 mM KCl, 600 mM KCl and 1 M KCl. The 250 mM fraction containing Nej1 was subjected to tobacco etch virus (TEV) digestion in the presence of TEV buffer (50 mM Tris pH 8, 0.5 mM EDTA) and a final concentration of ~83 mM KCl. TEV was used at a final concentration of 0.2 mg/mL and the reaction proceeded for 1 hour and 40 minutes, and subsequently stopped by the addition of 1 mM benzamidine. Following TEV cleavage, the reaction was re-applied through a SP Sepharose column, where Nej1 was recovered during a gradient from 200 mM KCl to 500 mM KCl (1.3 mL/min, 90 min). Fractions containing Nej1

were pooled and buffer exchanged using Hiprep 26/10 Desalting column 20 mM Tris pH 7.0, 10 mM DTT, 1 mM EDTA, 175 mM KCl and 10% glycerol. Typical protein yields were 0.5 mg of Nej1 per litre of cell culture.

2.4.3 Purification of His_6-tagged Nej1 1-244

The cell pellet was resuspended in Nickel-Column Buffer A containing 20 mM Tris pH 8.5, 2 mM βME, 10 mM imidazole, 500 mM KCl and 5 % glycerol. Following lysis using French Press, the clarified lysate was applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating column, and purified as described above. The 220 mM imidazole fraction was diluted to 150 mM KCl in the presence of 20 mM Tris pH 7.5, 5 mM DTT, 1 mM EDTA and 5% glycerol, filtered and injected onto Q Sepharose pre-equilibrated at 150 mM KCl. The protein recovered in the 200 mM KCl wash was subjected to TEV digestion at a final TEV concentration of 0.2 mg/ml in the presence of TEV buffer (50 mM Tris pH 8, 0.5 mM EDTA). The reaction was subsequently re-applied to a Q Sepharose column and the cleaved Nej1 1-244 was recovered in the unbound fraction, buffer exchanged into 20 mM Tris pH 7.0, 5 mM DTT, 1 mM EDTA, 175 mM KCl using a Hiprep 26/10 Desalting column. Typical protein recovery was in the range of 1.5 mg of Nej1 (1-244) per litre of cell culture.

2.4.4 Purification of His_6-tagged Nej1 1-252 and Nej1 1-269

The same approach was used to purify both Nej1 1-269 and Nej1 1-252. The cell lysate was prepared as described above for Nej1 1-244. Following lysis using French Press, the clarified lysates were applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating column, pre-equilibrated with Nickel-Column Buffer A. Successive washes were performed at 15 mM, 30 mM and 45 mM imidazole, and Nej1 was eluted in a 210 mM imidazole fraction. Purity of the fraction was confirmed by SDS-PAGE and the protein was frozen in liquid nitrogen in 30% glycerol and stored at –20°C until needed.

2.4.5 Purification of His_6-NusA-Nej1 170-342

The cell lysates were prepared as outlined for Nej1 1-244 (section 2.4.3). Clarified lysates were applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating column, pre-equilibrated with Nickel-Column Buffer A and recovered in the 210 mM imidazole fraction. The protein was subjected to TEV cleavage as outlined above. The sample was loaded onto a SP Sepharose column, in the same buffers as described for Q Sepharose (section 2.4.3), and recovered in the 300 mM KCl wash. The salt concentration was adjusted to 150 mM KCl and the protein was concentrated using a 10K macrosep. A typical protein yield was in the range of 1.3 mg of Nej1 170-342 per litre of cell culture.

2.4.6 Purification of His_6-NusA-Nej1 245-342

Cell lysates were prepared as outlined for Nej1 1-244 (section 2.4.3). Following lysis by French Press and centrifugation, the sample salt concentration was brought to 2M KCl by gradually adding the salt and stirring while on ice. The sample was applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating column, and purified as described above. The 210 mM imidazole fraction containing Nej1 245-342 was subjected to a TEV digest as described above. The sample was purified using a 5 mL Ni^{2+}charged Hi-Trap Chelating Nickel Affinity column and the protein of interest was recovered in the unbound fraction. Nej1 245-342 was then applied to a SP Sepharose column in the same Q Sepharose buffers described (section 2.4.3) and recovered in the unbound fraction. The protein was subsequently concentrated, frozen in liquid nitrogen and stored at –20°C. A typical yield was 0.05 mg per litre of cell culture.

2.4.7 Purification of His_6-NusA-Nej1 268-342

His_6-NusA-Nej1 268-342 was purified as described above for His_6-NusA-Nej1 245-342, except high salt conditions were not used during the initial Nickel Affinity column. Also, no further purification was done after the second Nickel Affinity column, as the protein obtained was a homogeneous mixture. The protein was buffer exchanged into 20 mM Tris pH 7.0, 5 mM DTT, 1 mM EDTA, 175 mM KCl and 5% glycerol using a Hiprep 26/10 Desalting column. Typical protein yield was in the range of 0.1 mg of pure Nej1 268-342 per litre of cell culture.

2.4.8 Purification of Nej1 (K283A, K285A, K287A) and Nej1 (K291A, K293A, R295A)

For each mutant, cell pellets were resuspended in Nickel-Column Buffer A containing 20 mM Tris pH 7.0, 2 mM βME, 5% glycerol, 500 mM KCl and 10 mM imidazole. For each of the point mutants, the clarified lysate was applied to a 5 mL Ni^{2+}charged Hi-Trap Chelating column, and purified as described above for the full-length protein. The recovered Nej1 was diluted to 130 mM KCl in the presence of 20 mM Tris pH 7.0, 5 mM DTT, 1 mM EDTA and 5% glycerol, and applied to an SP Sepharose column. For Nej1 (K283A, K285A, K287A), the protein was recovered in the 400 mM wash, further subjected to a TEV digest, and re-applied over a SP Sepharose column. The protein was eluted during a gradient from 200 mM to 500 mM KCl (1.3 mL/min, 50 min), and buffer exchanged as for the wild type protein. For Nej1 (K291A, K293A, R295A), the protein was eluted from an SP Sepharose column at 250 mM KCl, subjected to TEV cleavage and re-applied to an SP Sepharose column. Nej1 (K291A, K293A, R295A) eluted during a gradient from 200 mM to 500 mM KCl (1.3 mL/min, 60 min), was buffer exchanged in the same way as full-length Nej1.

2.4.9 Purification of Nej1 (K332A, R333A, K334A), Nej1 170-342 and Nej1 268-342

For each protein, cell pellets were resuspended in Nickel-Column Buffer A containing 20 mM Tris pH 7.0, 2 mM βME, 5% glycerol, 500 mM KCl and 10 mM imidazole. The clarified lysate was applied to a Ni^{2+}charged Hi-Trap Chelating column,

pre-equilibrated with 20 mM Tris pH 8.5, 2 mM βME, 5% glycerol, 500 mM KCl, and 10 mM imidazole, and subsequently recovered in the 210 mM imidazole fraction. The protein was subjected to TEV digestion, re-applied to a Nickel column and recovered in the unbound fraction. The protein was then buffer exchanged into 20 mM Tris pH 7.0, 1 mM EDTA, 150 mM KCl and 5 % glycerol, and concentrated.

2.4.10 Purification of His_6-tagged Lif1

Cell pellets were resuspended in Nickel-Column Buffer A containing 20 mM Tris pH 7.0, 2 mM βME, 5% glycerol, 500 mM KCl, 10 mM imidazole and 0.06% LDAO. Following lysis via French Press, the clarified lysate was subjected to a 5 mL Ni^{2+} charged Hi-Trap Chelating column, in the presence of 20 mM Tris pH 8.5, 2 mM βME, 5% glycerol, 500 mM KCl, 10 mM imidazole and 0.06% LDAO. The sample was washed as described above and Lif1 eluted in the 210 mM imidazole fraction. The Lif1 fraction was diluted to 150 mM KCl and subjected to further purification by passing it over a Q Sepharose column pre-equilibrated with 20 mM Tris pH 8.5, 150 mM KCl, 5 mM DTT, 1mM EDTA and 5% glycerol. After a 250 mM KCl wash, Lif1 eluted during a gradient from 250 mM to 600 mM KCl (1.7 ml/min, 92 minutes). The recovered Lif1 was brought to 1.5 M NH_4S0_4, applied to a Phenyl Sepharose column pre-equilibrated with 1.5 M NH_4S0_4, and recovered during a gradient from 1.5 M NH_4S0_4 to 0 M NH_4S0_4 (1 ml/min, 100 minutes). Lif1 was then buffer exchanged using Hiprep 26/10 Desalting

column into 20 mM Tris pH 7.0, 150 mM KCl, 5 mM DTT and 1 mM EDTA and concentrated using a 30K microsep.

2.4.11 Purification of His_6-tagged Nej1-Lif1 full-length complex

Cells were resuspended in 20 mM Tris pH 7.0, 2 mM βME, 5% glycerol, 500 mM KCl, 10 mM imidazole. Following lysis via French Press, the clarified lysate was applied to a 5 mL Ni^{2+} charged Hi-Trap chelating column, washed at 60 mM imidazole and subsequently eluted with 210 mM imidazole. The recovered protein was applied to a Q Sepaharose column and eluted during a gradient from 200 to 600 mM KCl (80 min, 1 mL/min). The Nej1-Lif1 containing fractions were pooled and further purified using a Superdex200 HiPrep 16/60 gel filtration column, equilibrated with 20 mM Tris pH 7.0, 1 mM EDTA, 200 mM KCl.

2.5 Electrophoretic Mobility Shift Assay (EMSA)

To generate a radiolabeled 35 bp DNA substrate, a 35 nt oligonucleotide (5'CCCTGTGCGACGCTAGCGTGCGGCCTGGTCTGTCC3') was labeled with [γ-^{32}P] ATP using Polynucleotide Kinase (PNK) and purified using denaturing PAGE. The radiolabeled oligonucleotide was annealed to its complementary non-radiolabeled oligonucleotide in an equimolar ratio, by heating to 95°C for 5 minutes and slowly cooling to room temperature. To generate a 340 bp substrate, PCR was used to obtain a PCR fragment which was subsequently radiolabeled as described above. DNA binding

reactions were carried out in a 20 µl reaction containing 10 mM Tris pH 7.0, 10 mM DTT, 1 mM EDTA, 30 mM KCl, 1 mg/ml bovine serum albumin (BSA) and 15 % glycerol. DNA binding reactions were incubated for 20 min at room temperature prior to being resolved on an 8% polyacrylamide gel. Electrophoresis was performed at 4°C for 1.5 hours at 150 V in 1X Tris Borate EDTA (TBE) buffer. Gels were exposed to a phosphorimaging screen and developed on a Typhoon Scanner. The amount of bound oligonucleotide was measured by quantitating intensity of the bands and the apparent dissociation constant (K_D) was determined from the protein concentration at which half of the duplex DNA was protein-bound (Carey 1991). Hill plots to assess co-operativity were generated using the program SigmaPlot.

For competition assays, Nej1 (8 µM) was initially incubated with a 35 bp ^{32}P-labeled DNA probe (2 µM) and allowed to come to equilibrium, after which the complex was challenged with increasing concentrations of cold competitor DNA in the form of nicked, linear or supercoiled plasmid (pUC19). The plasmid DNA was added at 1X microgram excess (0.0261 pmol in lane 4) up to 64X microgram excess over the 35 bp DNA (1.67 pmol in lane 11) and the reactions were resolved by EMSA. All assays were performed in triplicate, unless stated otherwise.

2.6 Circular Dichroism (CD) Spectroscopy

Proteins were buffer exchanged into 20 mM Phosphate pH 7.0, 150 mM KCl and 5 % glycerol and analyzed by plane polarized light from 199 nm to 250 nm, using Jasco J

600. Eight spectra were recorded and averaged, using a 0.1 cm path length cell at room temperature. Baseline spectra were recorded using 20 mM Phosphate pH 7.0, 150 mM KCl and 5 % glycerol, and subtracted from the protein spectra. The data was analyzed using CONTINLL, CDSTAR and SELCON3 (Johnson 1988).

2.7 Gel Filtration Chromatography

Gel filtration experiments were performed using fast protein liquid chromatography on a Superdex 200 equilibrated with buffer containing 20 mM Tris-HCl pH 7.0, 5 mM DTT, 1 mM EDTA and 200 mM KCl. The column was run at a flow rate of 0.5 ml/min. Elution profiles were monitored with an in-line UV-900 detector at 280 nm on a ÄKTAexplorer workstation, and peak fractions were analyzed by SDS-PAGE. Eight proteins of known molecular weight (Thyroglobulin 6690000, Ferritin 440000, Catalase 232000, Aldolase 158000, Albumin 67000, Ovalbumin 43000, Chymotrypsinogen 25000, RibonucleaseA 13700) were used as standards to calibrate the column. The molecular weight of sample proteins was determined using a logarithmic curve of the molecular weight plotted against V_e/V_0.

2.8 Analytical Ultracentrifugation

Sedimentation equilibrium experiments were performed at 4°C using a Beckman-Coulter XL-A analytical ultracentrifuge with an An-60 Ti rotor. Protein samples were buffer exchanged into a buffer containing 20 mM Tris pH 7.0, 150 mM KCl. Three

samples of each protein were prepared at OD_{280} of 0.2, 0.4 and 0.6. For full-length Nej1 the absorbance of the cell was measured at 280 nm at speeds of 6,000 rpm, 8,000 rpm, 10,000 rpm and 12,000 rpm. For Nej1 1-244 the rotor speeds used were 12,000 rpm, 14,000 rpm, 16,000 rpm and 18,000 rpm. The equilibrium data was analyzed using the Optima XL-A/XL-I analysis software within Origin version 6.0. The partial specific volumes were calculated from the amino acid composition of the proteins using SEDNTERP developed by Hayes, Laue, and Philo.

2.9 Partial Proteolysis

Full-length Nej1 (5 μg), was subjected to partial proteolysis by either chymotrypsin or trypsin (0.4 μg, 0.6 μg or 1.0 μg of protease) in the presence of 50 mM HEPES pH 7.5, 40 mM KCl, 10 mM $MgCl_2$ at room temperature. The reaction was allowed to proceed for a total of 30 minutes, with samples removed at five-minute intervals. Reactions were stopped by the addition of SDS loading dye containing 5 mM PMSF, and subsequently resolved via SDS-PAGE (150 V, 1 hour) and visualized by coomassie brilliant blue staining.

2.10 *In Vivo* Plasmid Repair Assay

This assay was performed as described previously, using BY8920 (containing wild type Nej1) and BY4741 (*nej1Δ*) *S. cerevisiae* strains (Ooi *et al.,* 2001). In brief, 400 ng of desired Nej1 WT or mutant expression vector (pSLO103 derivative) was

transformed one at a time, in parallel with 400 ng of substrate vector (either EcoRI "cut" pRS416 plasmid mimicking a double-strand break or "uncut" pRS416 vector) and plated on appropriate selection media. The number of colonies resulting from the "cut" transformation was scored as a percentage of the number of transformants arising from the "uncut" transformation to yield efficiency of repair of "cut" substrate or NHEJ. All levels of NHEJ were standardized to those observed in BY4741.

2.11 N-Terminal Amino Acid Sequencing of Nej1

N-terminal amino acid sequence of peptide fragments derived from Nej1 proteolysis was generated using a 492 gas-phase/pulsed-liquid Procise automated sequencer and analyzed using the Model 610A data analysis program.

2.12 In-Gel Trypsin Digest and MS/MS Identification of Proteins

Following SDS-PAGE and detection of proteins via coomassie staining, gel bands containing were excised from the gel and de-stained using 50 µl of water for 30 minutes by agitation at 1250 rpm. The gel slice was then shrunk by treatment with 50 µl of 50% CH_3CN 50 mM NH_4CHCO_3 twice for 20 minutes with agitation. The gel slice was dehydrated with CH_3CN for 20 minutes without agitation. Next, the gel slice was rehydrated with 50 µl of digest buffer containing 10% CH_3CN 50 mM NH_4CHCO_3, for 15 minutes at 100°C. To this solution, 1 µg of trypsin was added and incubated overnight at 37°C. The next day, the liquid was removed and the gel treated and

agitation with 50 µl of aqueous 5% CH_3CO_2H, followed by 10 µl of (1:1) CH_3CN/H_2O containing 0.2% HCO_2H. The extracts were combined and, by ZipTip-C18 (Millipore), tryptic peptides were concentrated by elution with 8 µl of (1:1) CH_3CN/H_2O containing 0.2% HCO_2H. Half of the solution was loaded into a long microspray needle (Proxeon) and analyzed using a Waters Q-Tof Global Mass Spectrometer (calibrated for MS and MS/MS with Glufibrinopeptide) equipped with a nanoelectrospray source. The capillary voltage was typically 1.0-1.5 kV, the cone voltage was 35-100 V and the collision energy was optimized for fragmentation. Mass spectra in TOF-MS and MS/MS modes were obtained in a mass range of 50-1800 m/z with a resolution of 8000 FWHM. Argon was used as collision gas. The data were obtained and processed using Mass Lynx 4.0. To identify interacting proteins, database searches were performed on peak lists obtained from the MS/MS mass spectra, which represent peptides produced by in-gel trypsin digestion of stained gel bands. Features corresponding to 2+ or 3+ charged ions were expected to correspond preferentially to ionized tryptic peptides and were selected manually for fragmentation-based sequencing by tandem mass spectrometry. The MS/MS data so obtained were processed by standard methods to generate MS/MS peak lists that are compared via software algorithms to sequence ions predicted for all the tryptic peptides derived from proteins known in the database. Individual peak lists were submitted in open searches against the nrNCBI database using parameters pertaining to the above methods in the web interface of Mascot (Matrix Science). Quadrupole/time-of-flight mass spectrometry measurements offered highly accurate data to match peptide hits

with significance scores providing confidently identified protein assignments. In each case, the predicted match was confirmed by inspection of the MS/MS mass spectra.

2.13 Generation of Nej1 Specific Antibodies

An antibody towards Nej1 was raised in a rabbit, using purified recombinant Nej1 1-244 as the antigen, according to standard immunological procedures (see materials and methods) (Drenckhahn *et al.,* 1993). In brief, 1 mL of the antigen was prepared by mixing 0.5 mL of purified recombinant Nej1 1-244 and 0.5 mL of Freund's Complete Adjuvant. Once emulsified, the antigen (100 μg) was injected into the rabbit, equally over 4 sites (0.25 mL in each) using the sub-cutaneous method of injection. Prior to injection, a pre-immunization sample was collected, by withdrawing 1 mL of blood and mixing it with an anticoagulating agent, 3.8 % NaCitrate, in a 1:10 ratio with the blood. The unclotted blood was then spun down at 2000 g for 5 minutes, after which the serum was collected and stored at -80°C. Exactly 10 days following the initial immunization, a 1 mL test bleed sample was collected in the same manner as the pre-immunization sample, and tested for the presence of a Nej1 specific antibody. Subsequent boosts of the antigen were administered every 21 days post initial immunization, in the same manner, except that Freund's Incomplete Adjuvant was used instead. Test bleed samples (10 mL) were collected exactly 10 days following each injection and tested for response via immunoblotting. For each a sample of the serum was retained and stored at -80°C. A total of 3 boosts were administered prior to a final antibody harvest. All immunological

procedures were performed according to the guidelines on antibody production, outlined by the Canadian Council on Animal Care.

2.14 Immunoblotting

To prepare yeast lysates, yeast strains were cultured overnight in synthetic complete media lacking histidine (SC-His), while shaking at 30°C. The next day, a subculture was set up using a 1:50 dilution of the overnight, saturated culture and the cells were allowed to grow at 30°C with vigorous shaking. Cell growth was monitored by measuring OD_{600}, and once this value reached 0.6-0.8, the required OD units of cells were collected, spun down and frozen at –20°C.

Yeast lysates were prepared by either the alkaline lysis method (Kushnirov 2000) or using the glass bead lysis method (Harlow et al., 2006). For alkaline lysis, 2.5 OD units of cells were resuspended in 100 µl of sterile ddH_20, to which 100 µl of 0.2 M NaOH was added, and the mixture was allowed to incubate for 5 minutes at room temperature. The reaction was spun down briefly in a microcentrifuge and the supernatant was discarded. The cell pellet was resuspended in 40 µl of a modified loading buffer (0.06 M Tris pH 6.8, 2% SDS, 4% βME, 5% glycerol and 0.0025% Bromophenol blue) and boiled for 3 minutes. The debris was pelleted and the supernatant was transferred into a fresh eppendorf tube and used for loading on a gel.

The glass bead lysis method was performed by resuspending 1 OD units of cells in 40 µl of TLB buffer (80 mM Tris·HCl pH 6.8, 20% SDS, 100 mM DTT, 10% glycerol and 0.1% Bromophenol blue), and immediately boiled for 2 minutes. To this, 0.12 mg

(approximately 100 µl volume in an eppendorf) of glass beads was added and the reaction was vortexed for 1 minute on highest setting. An additional 160 µl of TLB buffer was added and the mixture was boiled again for 2 minutes, and vortexed again for 1 minute on highest setting. The beads were pelleted using a microcentrifuge and 150 µl of the yeast lysate was recovered for loading on a gel.

All samples were first subjected to SDS-PAGE on a 12% gel (150V, 1 hour). Before assembling the transfer apparatus, sponges and whattman paper sheets were soaked in transfer buffer (25 mM Tris pH 6.8, 190 mM glycine, 10% methanol and 0.05% SDS). In addition, the PVDF membrane was activated in 100% methanol for 15 minutes and next equilibrated in transfer buffer for at least 15 minutes. Subsequently, the proteins were transferred to a PVDF membrane at 25V for 1.5 hours in the transfer buffer. Once the transfer was complete, the membrane was blocked using 5% skim milk in 1X TBST buffer (10 mM Tris pH 7.4, 150 mM NaCl and 0.1% Tween 20) for 15 minutes, rocking on a slow setting. Next, the membrane was incubated with the 1° antibody (rabbit anti-Nej1 1-244) for 1.5 hours (1/1000 dilution) while slowly rocking. Subsequently, three 5-minute washes were carried out using 1X TBST buffer while rocking at a fast speed. Next, the 2° antibody (donkey anti-rabbit IgG) was incubated with the membrane for 45 minutes (1/10000 dilution) while slowly rocking. Once again, residual antibody was washed away during three 5-minute washes using 1X TBST buffer, while rocking at a fast speed. Finally, the membrane was immersed in Western Lightning solution for 1 minute, prior to exposure of the membrane onto an XRA film. All immunoblotting procedures were carried out at room temperature.

2.15 Crystallization and X-Ray Data Collection

Protein Crystallization

All crystallization trials were performed at either 4°C or 20°C by the hanging-drop vapour-diffusion method. Drops containing 1 µl of the protein solution and 1 µl of the condition were suspended over 1 ml NH_4S0_4 (1 M, 1.5 M, 2 M or 2.5 M as indicated). Where indicated, additives were added to the drop at a final concentration of 20 mM. Alterations to this general set up are stated within text.

Crystallization of Nej1-DNA complex—Purified recombinant full-length Nej1 was incubated with a 15 bp DNA substrate in a 1:1.3 molar ratio. Drops containing 1 µl of Nej1-DNA and 1 µl of the condition were suspended over 1 ml of 1.5 M NH_4S0_4. For crystal growth conditions see Table A.IV in Appendix I.

Crystallization of Nej1 1-244—For crystal trials, 1µl of protein and 1 µl of the condition, or 2 µl of protein and 1 µl of the condition, were suspended over 500 µl of 1.5 M NH_4S0_4. For crystal growth conditions see Table A.V and Table A.VI in Appendix I.

X-ray Data Collection

Nej1-DNA complex—Crystals of full-length Nej1 bound to15 bp DNA, grown in condition 4, (see Appendix I, Table A.IV) were mounted and exposed for two minutes of X-rays at the National Synchrotron Light Source, Brookhaven National Laboratory.

Several images 0.5° in width were collected, with diffraction reaching 30Å. Images were viewed using the program Crystal Clear.

Nej1 1-244- Crystals of Nej1 1-244 were transferred to a cryo buffer containing prior to placing the crystal in the X-ray beam. Two 15-minute exposures were performed, recording two images 0.5° in width and 90° apart, with diffraction reaching a maximum resolution of 10Å. The data was collected using a Cu-anode rotating X-Ray machine, and the images were viewed using the program Crystal Clear.

CHAPTER 3 – RESULTS

3.1 Functional Characterization of Nej1

Full-length Nej1 (39 kDa) was over expressed in *E. coli* as a hexa-histidine fusion. To obtain a highly pure sample of Nej1, the protein of interest was initially purified using metal ion affinity chromatography using a Ni^{2+} charged resin to trap the His_6-tagged Nej1. Subsequent purification using cation exchange chromatography exploited the positively charged Nej1 protein and resulted in the elimination of further contaminants. TEV cleavage was employed to remove the hexa-histidine fusion and the sample was re-purified using cation exchange chromatography to yield a near homogeneous Nej1 sample. Fig. 3.1 depicts an SDS-PAGE gel representative of the final purified full-length Nej1 protein.

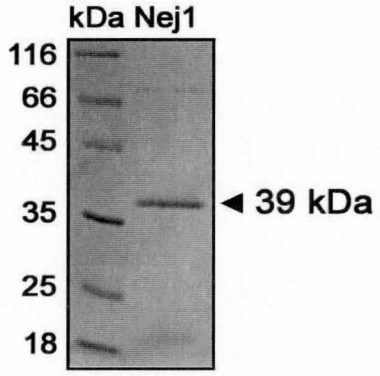

Figure 3.1 Purified recombinant full-length Nej1. Nej1 was initially purified using Nickel affinity chromatography, after which the His_6 fusion was removed and subsequent purification steps resulted in near homogeneous Nej1 protein. Nej1 (39 kDa) obtained after a final buffer exchange, was analyzed by SDS-PAGE. kDa - kilodalton.

While most proteins involved in the NHEJ pathway possess DNA binding activity, to date it has not been shown whether Nej1 also possesses the ability to interact with DNA. Based on the hypothesis that Nej1 plays a direct role in the repair of DSBs, a possible way that Nej1 may participate in end-joining is by interacting with DNA. To investigate this possibility, an EMSA was carried out to determine if Nej1 possesses DNA binding activity. DNA binding reactions were carried out with ^{32}P-labeled 35 bp DNA in the presence of increasing amounts of purified recombinant Nej1. The reactions were resolved by native PAGE and the resulting species quantified by phosphorimaging analysis (Fig. 3.2A). The K_D was determined by plotting the fraction of bound DNA against protein concentration (Fig. 3.2B). This data illustrates that Nej1 possesses DNA binding activity and binds the 35 bp DNA with an apparent K_D of 9.5 μM *in vitro*. This protein-DNA interaction was fully stable up to 120 mM KCl and essentially diminished above 300 mM KCl (data not shown).

Given that DSBs can occur anywhere, it is reasonable and even expected that proteins involved in end-joining bind DNA independently of sequence. The lack of sequence specificity was confirmed for Nej1 by testing DNA binding activity with 3 additional small DNA substrates sharing unrelated sequence (data not shown). Consistent with other proteins involved in NHEJ, the DNA binding activity of Nej1 is not dependent on DNA sequence. This suggests that DNA binding is primarily mediated by electrostatic interactions between positively charged residues in Nej1 and the negatively charged DNA phosphates. As such, the affinity of such an interaction is expected to be

weaker then those observed for sequence specific protein interactions (Kneale 1994), thus the apparent dissociation constant determined for Nej1 is in the expected range.

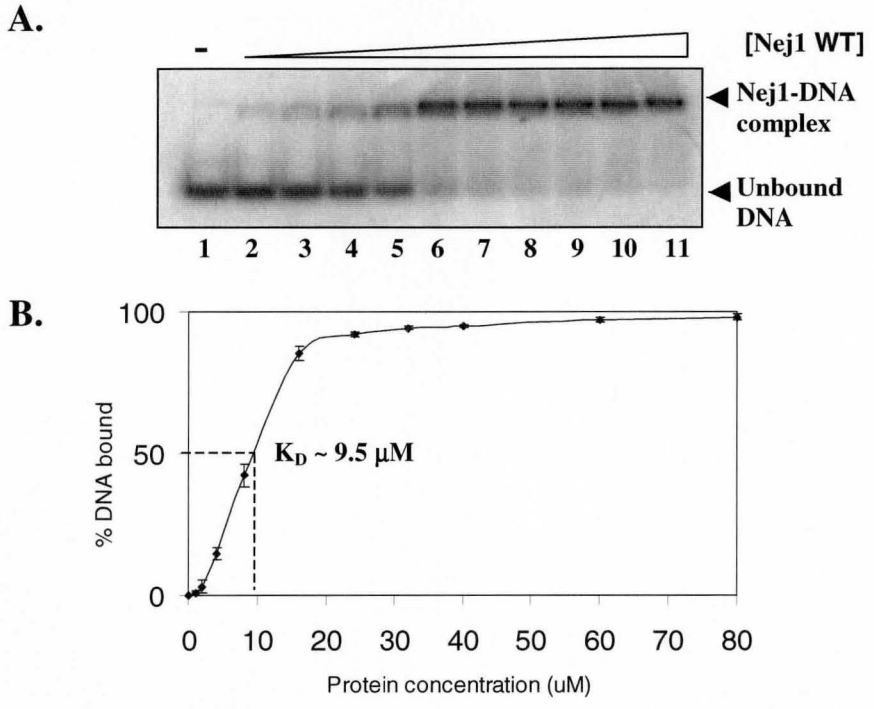

Figure 3.2 Nej1 DNA binding on a 35 bp DNA substrate. (A) Binding of Nej1 to a 35 bp ^{32}P-labeled DNA probe. In each reaction 2 µM of DNA was incubated with increasing amounts of Nej1 protein (1µM lane 2 to 80 µM lane 11), analyzed by EMSA and visualized via autoradiography. (B) The fraction of DNA bound was plotted versus the protein concentration, and the K_D determined as the concentration of protein at which 50% of the DNA is bound. Error bars represent the standard deviation of 3 independent experiments.

While proteins involved in DNA repair and recombination frequently posses no sequence-specific DNA binding preference, interestingly these same proteins often exhibit altered affinities towards different DNA structures. Human Xrcc4, for instance,

has been shown to bind DNA with no sequence specificity, but exhibit a marked preference for double-strand DNA containing either a nick or an end (Modesti *et al.,* 1999). This feature may serve as an advantage allowing DNA repair proteins to quickly recognize a specific type of damage, for efficient repair. In fact, preferential DNA binding of other NHEJ proteins, such as Ku, to certain types of DNA has been demonstrated (Blier *et al.,* 1993). In contrast, a recent study on human and *S. pombe* XLF reported equal DNA binding towards linear and closed circular plasmid DNA (Hentges *et al.,* 2006).

To determine whether Nej1 exhibits DNA structure-specific binding activity, a competition EMSA was carried out. In these reactions, Nej1 was first incubated with a 35 bp DNA probe under conditions where all DNA was bound with protein. This Nej1-35 bp DNA complex was subsequently challenged with different amounts of cold competitor DNA in the form of linear, nicked or supercoiled plasmid. Once again reactions were resolved by native PAGE and subjected to autoradiography. All forms of competitor DNA tested competed essentially equally well for Nej1 binding (Fig. 3.3), indicating that *in vitro* Nej1 exhibits no DNA-structure specificity. This lack of structure-dependent DNA binding behaviour has also been reported for Lif1 (Teo *et al.,* 2000), the *S. cerevisiae* homologue of Xrcc4. Thus it is possible that, while the yeast NHEJ proteins (Nej1, Lif1) have no preference towards a particular form of DNA, the mammalian end-joining factors (Xrcc4, Ku) have evolved to distinguish different DNA structures, allowing for more efficient recognition and repair of the damaged site.

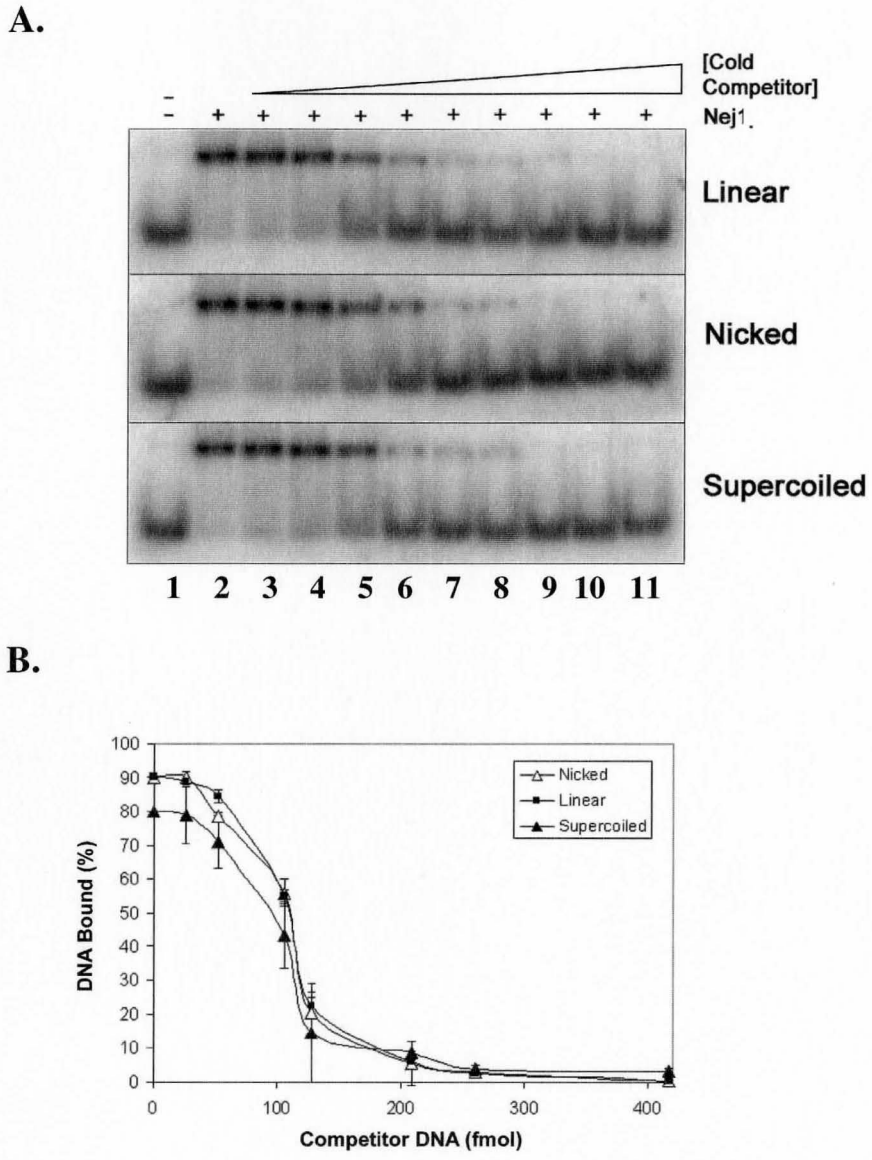

Figure 3.3 Nej1 exhibits no DNA-structure specificity. (A) Nej1 was initially incubated with a 35 bp ^{32}P-labeled DNA probe, after which a cold competitor DNA, in the form of linear, nicked, and supercoiled plasmid (2876 bp) was added at 1x microgram excess (0.0261 pmol in lane 3) up to 64 microgram excess over the 35 bp DNA (1.67 pmol in lane 11) and the reactions were resolved by EMSA. (B) Graphical representation of results in A. Reactions were analyzed by EMSA and visualized via autoradiography. Error bars represent the standard deviation of 2-3 separate experiments.

Subsequent characterization revealed that Nej1 displayed equal binding affinity for short DNA substrates containing a hairpin, nick, blunt end, or 5'or 3' 5-nucleotide overhangs (data not shown). Furthermore, the binding affinity towards single stranded DNA substrates was comparable to that of double strand DNA (data not shown). Therefore, when assembled into a repair complex *in vivo*, Nej1 might be expected to contribute directly to the stability of the protein-DNA interactions within the complex without necessarily directly interacting with DNA ends. DNA binding to substrates shorted than 15 bp could not be detected, suggesting that when Nej1 is assembled into a higher-order NHEJ repair complex, it interacts with at least 15 bp of DNA.

In addition to DNA structure specificity, some end-joining proteins, specifically human Xrcc4, demonstrated an unusual requirement for large DNA substrates (~ 300 bp) in order to form stable protein-DNA complexes (Modesti *et al.,* 1999). Although Nej1 is capable of binding short DNA substrates as shown above, the prospect for enhanced binding to larger DNA substrates required investigation. DNA was therefore performed on a ^{32}P-labeled 340 bp substrate, and assayed by native PAGE as outlined above. Interestingly, Nej1 bound this longer DNA substrate with higher affinity than the 35 bp substrate, with an apparent K_D of 1.8 μM (Fig. 3.4).

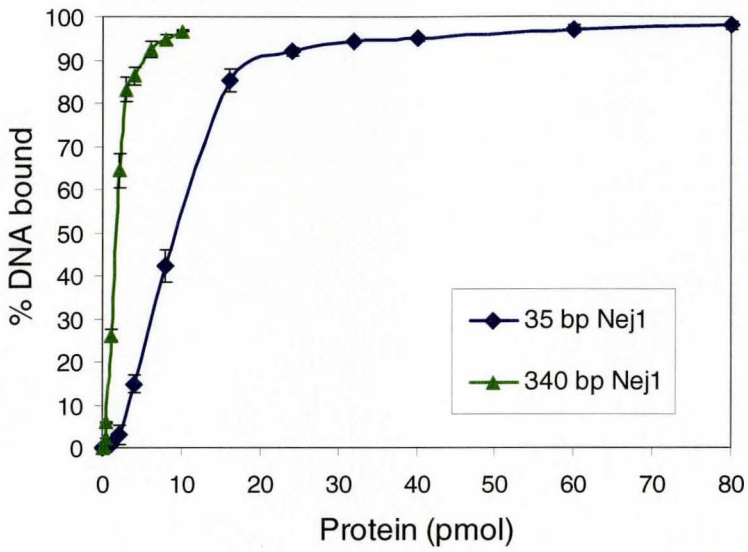

Figure 3.4 Comparison of Nej1 DNA binding affinities on 35 bp and 340 bp substrates. Nej1 was incubated in the presence of 35 bp ^{32}P-labeled DNA (2 μM) or 340 bp ^{32}P-labeled DNA (0.2 μM) and analyzed by EMSA. The fraction of DNA bound was plotted versus the protein concentration, and the K_D determined as the concentration of protein at which 50% of the DNA is bound. Nej1 exhibited an increased DNA binding affinity towards the longer DNA substrate, 9.5 μM vs. 1.8 μM for the 35 bp and 340 bp substrates respectively. Error bars represent the standard deviation of 2-3 independent experiments.

Although initial studies of XLF DNA binding were conducted with 1000 bp DNA substrates (Hentges *et al.,* 2006), a recent report demonstrated that while 15 bp could theoretically suffice as a DNA binding substrate, a much larger DNA fragment (~80 bp) is actually required to form a stable complex (Lu *et al.,* 2007). The mechanistic basis for this unusual preference remains unclear; however, this behaviour is consistent given the fact that XLF family members, including Nej1 and XLF, are predicted to be structurally and possibly functionally similar to Xrcc4, which also shares this unusual property. This concept is further expanded on in the discussion (Chapter 4).

3.2 Functional Comparison of Nej1 and Lif1

Several reports demonstrated that Nej1 strongly interacts with another NHEJ factor, called Lif1 (Ooi *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007). Although it was previously demonstrated that Lif1 possesses an ability to interact with DNA (Teo *et al.,* 2000), a thorough characterization of this property was not carried out. Thus, the DNA binding properties of Lif1 were explored in an effort to better understand the purpose of Nej1 and Lif1 in end-joining.

In order to perform a functional comparison of Nej1 and Lif1 DNA binding properties, pure recombinant Lif1 protein was required. Lif1 was over-expressed and purified to near homogeneity as seen in Fig. 3.5.

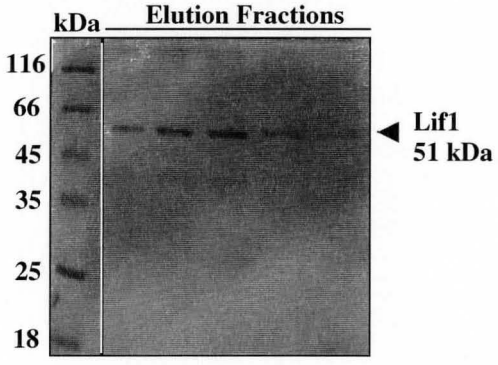

Figure 3.5 Purified recombinant full-length Lif1. SDS-PAGE of Lif1 fractions recovered during a gradient on a Phenyl Sepharose column. Lif1 migrated close to its predicted molecular weight of 51 kDa (including His_6 fusion).

Since the DNA binding of Lif1 has only been reported for larger DNA substrates (~300 bp) (Teo et al, 2000), the affinity of Lif1 towards both 35 bp and 340 bp DNA

substrates was investigated, to allow for direct comparison in the activities of both Nej1 and Lif1. Surprisingly, unlike its human homologue Xrcc4, Lif1 displayed substantial DNA binding towards a short 35 bp substrate, with a K_D of 13.5 μM. This data revealed that like Nej1, Lif1 also showed increased affinity towards longer DNA. Furthermore, both Lif1 and Nej1 exhibited similar affinities towards the 340 bp substrate, with apparent K_D values of 3.0 and 1.8 μM, respectively (Fig. 3.6). While these results suggest a strong similarity in the DNA binding behaviour of Lif1 and Nej1, they also highlight a potentially significant difference between these proteins and their human homologues, in that Xrcc4 and XLF seem to require larger DNA substrates for stable interactions.

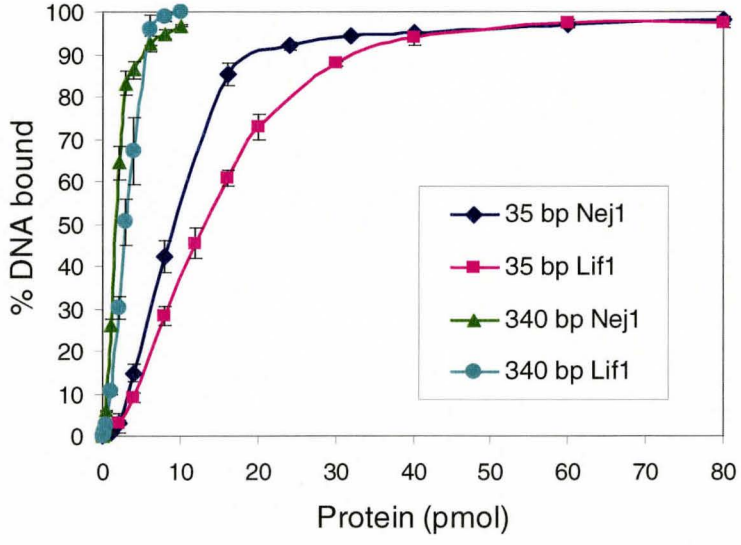

Figure 3.6 Comparison of Nej1 and Lif1 DNA binding affinities. Nej1 and Lif1 were incubated individually in the presence of 35 bp ^{32}P-labeled DNA (2 μM) or 340 bp ^{32}P-labeled DNA (0.2 μM) and analyzed by EMSA. Both proteins exhibited increased DNA binding affinities towards the longer DNA substrate, 9.5 μM vs. 1.8 μM for Nej1, and 13.5 μM vs. 3.0 μM for Lif1, for 35 bp and 340 bp substrates respectively. Error bars represent the standard deviation of 3 independent experiments.

The mammalian homologue of Lif1, Xrcc4, has been shown to bind DNA with a high degree of co-operativity (Modesti *et al.,* 1999). Although the functional significance of this DNA binding property remains unclear, this co-operative DNA binding behaviour suggests that the formation of higher-order complexes of Xrcc4 are favored in their ability to interact with DNA. To determine whether Lif1 and Nej1 share similar co-operative DNA binding properties as Xrcc4, Hill Plots were generated from which co-operativity coefficients (n) were obtained. A Hill coefficient greater than one indicates positive co-operative binding. Figure 3.7 shows the Hill Plots generated for the interactions of each protein on both 35 bp and 340 bp DNA substrates. This analysis revealed clear co-operative DNA binding behaviour for both Nej1 and Lif1. Thus, while DNA binding affinity varied with the length of substrate, binding co-operativity did not; suggesting, as one would anticipate, that co-operative binding behaviour is independent of the substrate. Characterization of the structural nature of the protein-protein interactions that drive this observed co-operativity would greatly impact the understanding of how these proteins function mechanistically in NHEJ. A summary of the DNA binding affinities (K_D) and co-operativity coefficients (n) can be found in Table 3.1. A more in-depth discussion of Nej1 DNA binding properties is provided in section 4.1.

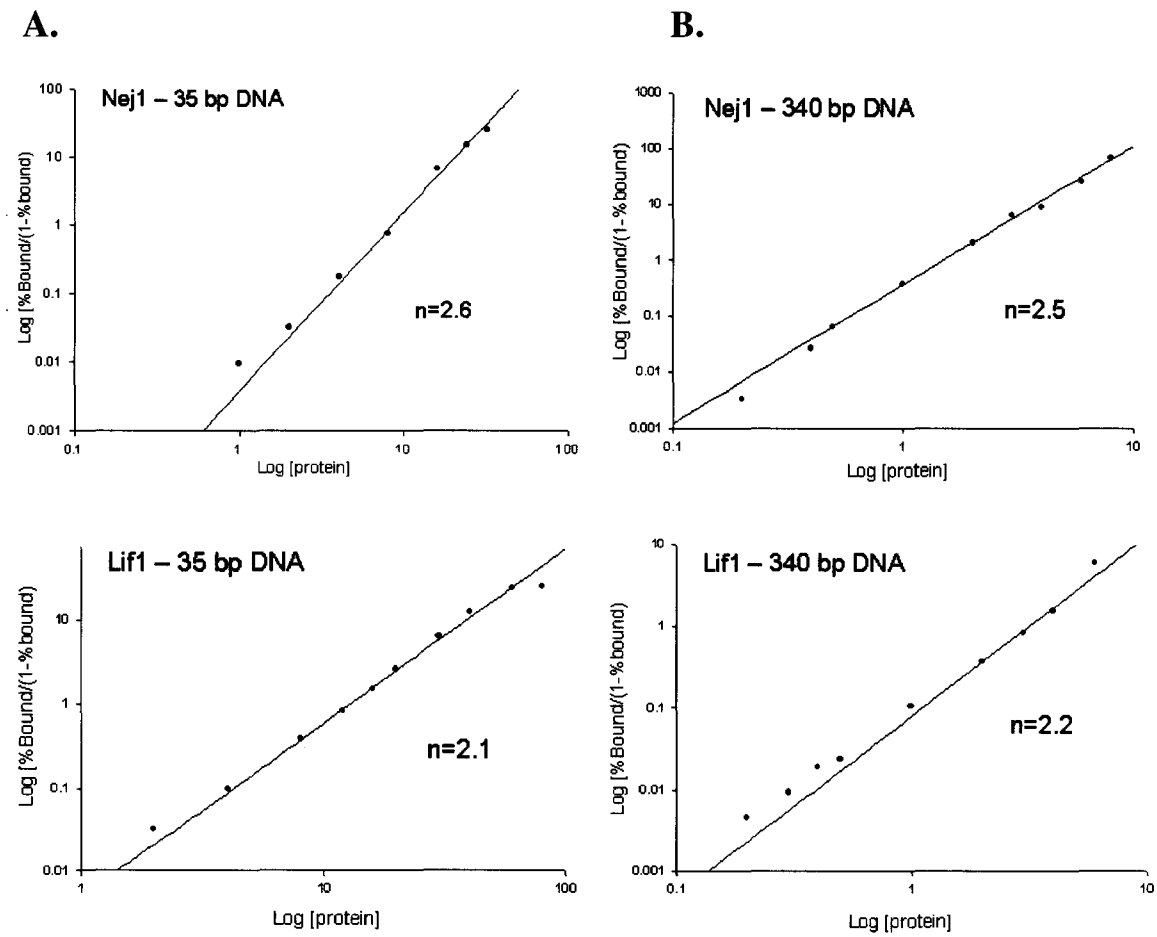

Figure 3.7 Nej1 and Lif1 Hill Plots. Hill plots were generated for DNA binding data of both (A) 35 bp and (B) 340 bp substrates for Nej1 and Lif1. Hill plots for Nej1 and Lif1 indicate that all protein-DNA interactions are co-operative. Hill Plots were generated using SigmaPlot.

Table 3.1 Summary of apparent dissociation constants and co-operativity coefficients obtained for Nej1 and Lif1.

	35 bp DNA		340 bp DNA	
	K_D (μM)	Hill Coefficient (n)	K_D (μM)	Hill Coefficient (n)
Nej1	9.5	2.6	1.8	2.5
Lif1	13.5	2.1	3.0	2.2

3.3 Functional Analysis of Nej1 Amino-Terminal Constructs

The finding that Nej1 has the ability to interact with DNA provided supporting evidence for the hypothesis that Nej1 may be directly involved in end-joining. Having characterized this DNA binding activity and compared it to the binding properties of Lif1, the next step was to investigate the structure-function relationship of Nej1. The initial approached was to identify domains or motifs in Nej1 based on homology to other proteins. However, similarity searches based on primary sequence and secondary structure failed to retrieve proteins or motifs with significant homology to Nej1 (data not shown).

Nevertheless, initial clues regarding the structural properties of Nej1 were obtained by carefully observing the behaviour of Nej1 during purification. While purifying full-length Nej1 it was evident that a significant amount of degradation was occurring (data not shown). In fact, if the protein was exposed to room temperature for an extended period of time, the majority of the purified protein would eventually degrade into a distinct species between 25 and 30 kDa (Figure 3.8, compare lane 2 or 11 with lane 9 or 18). In order to determine whether this cleavage product represents a stable domain within Nej1, partial proteolysis was used to assess if there are stable domains within Nej1.

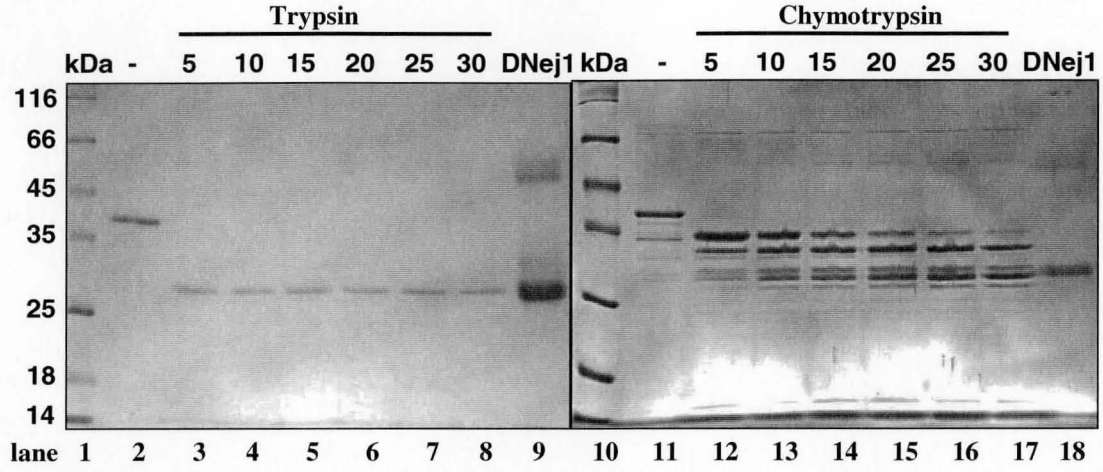

Figure 3.8 Partial proteolysis of Nej1 confirms presence of a stable domain. SDS-PAGE showing a time course proteolysis of Nej1 (5 μg or 90 μM) by chymotrypsin and trypsin (0.4 μg). Rapid cleavage of full-length Nej1 (lane 2 and 11) is observed, followed by the formation of a stable domain resistant to further proteolysis. The stable cleavage product is of similar molecular weight to the degraded Nej1 (DNej1) observed during purification (lane 9 and 18). Results are representative of proteolysis obtained with 0.4 μg, 0.6 μg or 1.0 μg of protease. The presence of DNA (at 1.5x or 135 μM) had no effect on the proteolysis pattern of Nej1 (data not shown).

A time course proteolysis, with either trypsin or chymotrypsin, shows the formation of a product resistant to further cleavage, similar in migration to the fragment observed during purification (Fig. 3.8). In a similar experiment, the proteolysis of Nej1 in the presence of DNA (1.5 molar excess) resulted in no change in the observed cleavage pattern. To eliminate the possibility that Nej1 may be inactive under proteolysis reaction conditions, an EMSA was carried out in both buffer conditions confirming that DNA binding is unaffected (data not shown). Since DNA binding is occurring in the reaction but not altering the proteolysis of Nej1, it suggests that the on and off rate of Nej1 for DNA is rapid, leaving Nej1 susceptible to cleavage.

To determine the boundaries of the stable fragment observed during purification, and generated during proteolysis, so that it could be further characterized, the fragment was analyzed by N-terminal amino acid sequencing. The results indicated that this stable Nej1 product contains an intact N-terminus indicating that degradation was occurring only in the C-terminus (data not shown). Inspection of the secondary structure prediction of the C-terminus of Nej1 revealed that this region is the most unstructured part of the protein, and thus may explain the susceptibility of the C-terminus for proteolytic cleavage (Appendix I, Fig. A.I). To determine the exact cleavage site and thus determine the C-terminal boundary, the stable fragment was analyzed by mass spectrometry. Despite repeated attempts however, a reproducible mass could not be obtained and thus at best, only an estimate of ~29 kDa could be made (n = 4; data not shown).

Nevertheless, based on evidence from partial proteolysis, as well as careful examination of the amino acid sequence and secondary structure prediction of Nej1, three C-terminal deletion constructs were designed: Nej1 1-244, Nej1 1-252 and Nej1 1-269, with predicted molecular weights of 27.9 kDa, 28.9 kDa and 31.0 kDa respectively (Fig. 3.9). Each of these constructs was designed in hopes of generating a pure and stable N-terminal domain protein sample for further characterization.

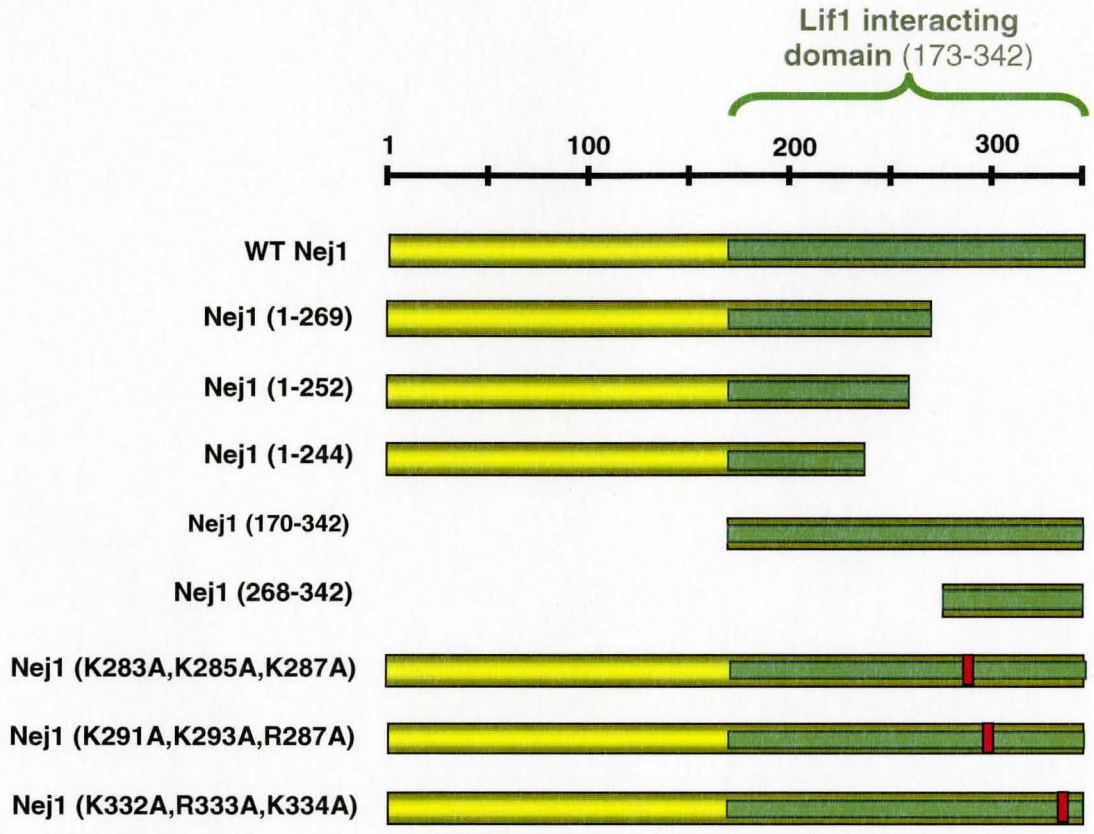

Figure 3.9 Nej1 deletion and point mutation constructs. The above constructs were designed based on the behaviour of Nej1 during purification, experimental evidence from partial proteolysis, secondary structure prediction and the distribution of amino acids. The boundaries of the Lif1 interaction domain, as determined previously (Ooi *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007) are indicated in green. Red boxes indicate location of point mutations.

The three Nej1 truncations were expressed, purified and assayed for DNA binding activity in the same manner as full-length Nej1. Unexpectedly, none of the truncated Nej1 proteins exhibited any DNA binding activity (Fig. 3.10A, data not shown), suggesting that residues 270 to 342 are required for DNA binding. This is surprising

because even shorter truncated forms of human Xrcc4 (aa 1-200) retain DNA binding activity (Modesti *et al.,* 1999) and can fully complement the V(D)J defect in XR-1 cells *in vivo* (Mizuta *et al.,* 1997; Leber *et al.,* 1998).

To eliminate the possibility that these truncations are inactive due to protein unfolding, the secondary structure content of the proteins was investigated using CD Spectroscopy. The spectrum obtained for Nej11-244 showed a negative minimum at 207 nm and a shoulder around 220 nm, indicative of the presence of both α-helical and β-sheet structures (Johnson 1988), suggesting that the protein was not unfolded (Fig. 3.10B). In addition, the spectrum revealed an increase in the α-helical content of the protein as compared to full-length Nej1, which is consistent with the removal of a largely unstructured C-terminus. Collectively, this analysis suggests that Nej1 requires the C-terminus to interact with DNA.

The absence of DNA binding activity upon removal of Nej1 C-terminus was interesting because to date there have been no reports localizing the DNA binding site for any XLF family members. In the case of human Xrcc4, Modesti and colleagues localized DNA binding to the first 200 amino acids of the protein (Modesti *et al.,* 1999). In this study it was further demonstrated that the entire C-terminal region from residues 201-336 was dispensable for both *in vitro* and *in vivo* NHEJ and V(D)J recombination. The functional significance of the C-terminal region of Xrcc4 remains unclear. This region of the protein is presumed to be largely unstructured, since it is highly sensitive to protease treatment and was irresolvable in the X-ray crystal structure (Junop *et al.,* 2000). Nej1 sensitivity to proteolysis is very similar to Xrcc4 and is consistent with the limited

amount of predicted secondary structure within the C-terminal region of both Xrcc4 and Nej1. Nevertheless, despite these similarities, it is evident that the organization of functional domains in these two proteins is very different, since removing the C-terminus of Nej1 abolished its ability to interact with DNA. The significance of this finding is discussed in section 4.5.

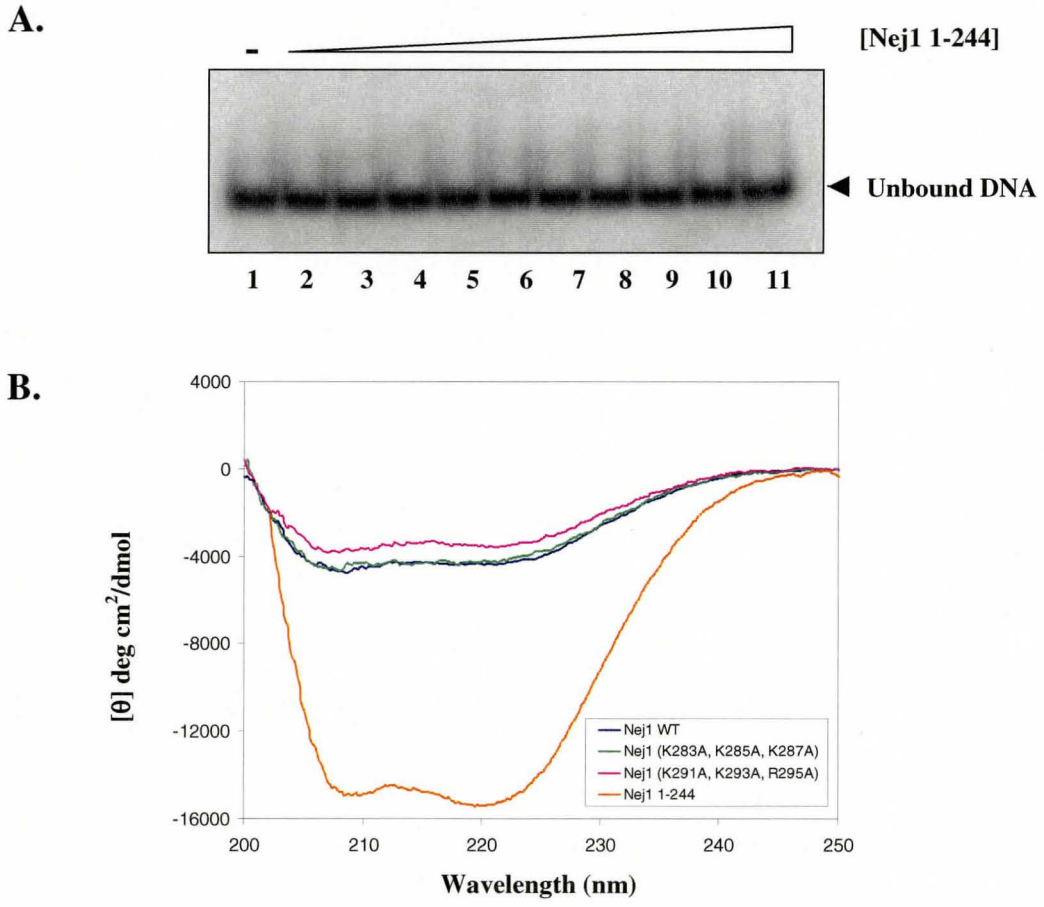

Figure 3.10 Nej1 possesses a proteolytically susceptible C-terminus that is required for DNA binding. (A) Nej1 1-244 does not exhibit DNA binding activity towards a ^{32}P-labeled 35 bp DNA probe. In each reaction 2 μM of DNA was incubated with increasing amounts of Nej1 protein (1μM lane 2 to 80 μM lane 11) and analyzed by EMSA. (B) CD spectroscopy of Nej1 WT, Nej1 1-244, Nej1 (K281A, K283A, K285A) and Nej1 (K293A, K295A, R297A) obtained at room temperature, showing an increase in α-helical content of the truncated protein.

3.4 Functional Analysis of Nej1 Carboxyl-Terminal Constructs

Although frequently a specific domain may be required for mediating a specific activity, it does not necessarily follow that that domain on its own will be sufficient to perform that function. Thus to further investigate the structure-function relationship of Nej1, it was necessary to perform additional characterization. Having established that the C-terminal ~70 residues of Nej1 are necessary for maintaining Nej1-DNA interactions (section 3.3), the next logical step was to investigate whether the C-terminus is sufficient for sustaining these contacts. To examine this possibility, two constructs expressing the C-terminus of Nej1 were designed. Since the C-terminus of Nej1 (~ aa 270-342) is predicted to be the least structured region of the protein, the first construct was engineered to contain residues 170-342. This boundary was based on previously published yeast two-hybrid data, which indicated that this fragment of the protein can be successfully expressed and is functional in an assay for Lif1 interaction (Frank-Vaillant *et al.,* 2001). The second construct was designed based on the fact that the region encompassing aa 270-342 was shown to be required for DNA binding (section 3.3). Thus, a construct expressing only these residues, specifically aa 268-342, was also constructed. Anticipating the fact that the C-terminus on its own may be susceptible to degradation and difficult to purify, both truncations were engineered in vectors containing either His_6 or NusA-His_6 TEV-cleavable, N-terminal fusions. An illustration of the C-terminal constructs is shown in Fig. 3.9 (section 3.3).

To precisely determine if the C-terminus of Nej1 constitutes the DNA binding domain, and is sufficient to promote DNA interactions, the two N-terminal truncations,

Nej1 170-342 and Nej1 268-342, were over-expressed, purified and tested for DNA binding activity. Both of these C-terminal fragments demonstrated similar DNA binding properties compared to full-length Nej1 on linearized plasmid DNA (Fig. 3.11 and data not shown). In contrast, no binding could be detected with either Nej1 1-244 or Nej1 1-269 under the same conditions. Together these results provide convincing evidence suggesting that the C-terminus of Nej1 contains the DNA binding domain. Although the C-terminus of Nej1 does not resemble a typical DNA binding domain, this highly basic domain is clearly both necessary and sufficient to promote DNA binding interactions. Indeed, domains with high concentrations of basic amino acids have previously been implicated in playing an important role in DNA binding (Busch *et al.,* 1990).

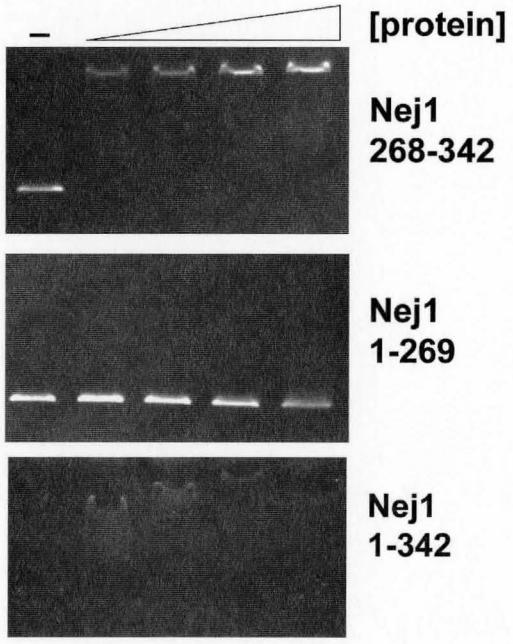

Figure 3.11 Nej1 C-terminal 72 amino acids are sufficient for DNA binding. Linearized pUC19 (100 ng) was incubated with increasing amounts of Nej1 268-342, Nej1 1-269 or Nej1 1-342 (0, 0.25, 0.5, 1 and 2 ng), resolved by EMSA (1% agarose, 100 V, 40 minutes) and visualized by ethidium bromide staining.

3.5 Mutational Analysis of Nej1

Once the DNA binding domain was localized to the C-terminal ~70 residues of Nej1, the next step was to investigate the residues within this region to further study the DNA binding properties. Inspection of the amino acid distribution within the C-terminus of Nej1 revealed an abundance of positively charged residues (Appendix I, Fig. A.II). Clusters of three neighboring lysine or arginine residues were simultaneously substituted for alanines to generate the mutants: Nej1 (K283A, K285A, K287A), Nej1 (K291A, K293A, R295A) and Nej1 (K332A, R333A, K334A) (Fig. 3.9). These mutants were purified and tested for DNA binding activity as previously described for full-length Nej1. As illustrated in Fig. 3.12, the DNA binding activity of these mutants was reduced relative to wild type protein, with apparent K_D values of 47 μM, 14 μM and 75 μM for Nej1 (K283A, K285A, K287A), Nej1 (K291A, K293A, R295A) and Nej1 (K332A, R333A, K334A), respectively. CD Spectroscopy analysis of Nej1 (K283A, K285A, K287A) and Nej1 (K291A, K293A, R295A) revealed that no significant structural changes had occurred as a result of introducing these amino acid substitutions (Fig. 3.10). These results not only further implicate the C-terminal region of Nej1 in DNA binding, but also suggest that efficient DNA binding is a result of numerous and perhaps co-operative interactions within the C-terminal ~70 residues of Nej1. A summary of the DNA binding results for all constructs is provided in Fig. 3.13.

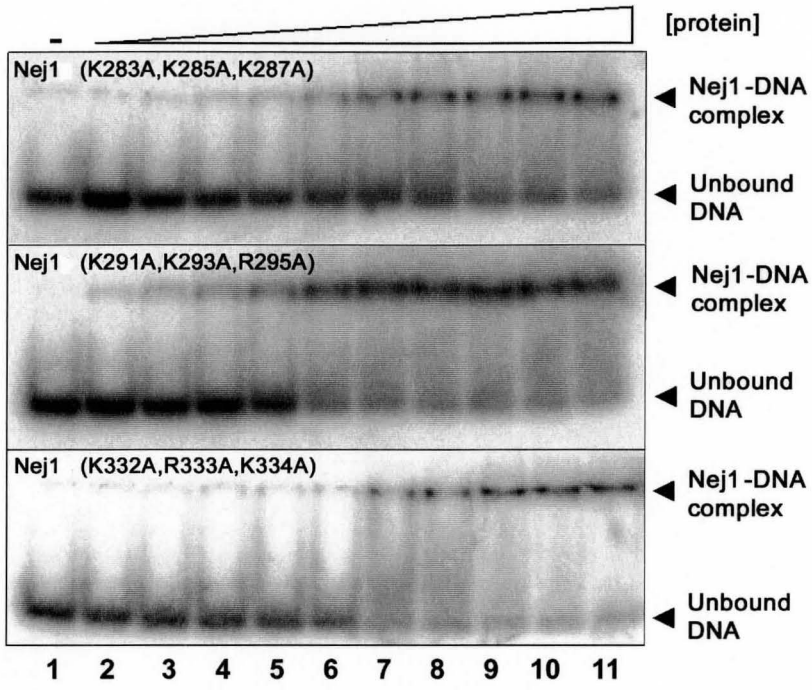

Figure 3.12 DNA binding activity of all Nej1 point mutants is reduced relative to WT Nej1. In each reaction, 2 μM of ^{32}P-labeled 35bp DNA was incubated with increasing amounts of protein (1 μM lane 2 to 80 μM lane 11 for Nej1 (K283A, K285A, K287A) and Nej1 (K291A, K293A, R295A), and 2 μM lane 2 to 300 μM lane 11 for Nej1 (K332A, R333A, K334A)), analyzed by EMSA and visualized via autoradiography. Nej1 (K283A, K285A, K287A), Nej1 (K291A, K293A, R295A) and Nej1 (K332A, R333A, K334A) had K_D values of 47, 14 and 75 μM respectively, compared to 9.5 μM for WT Nej1.

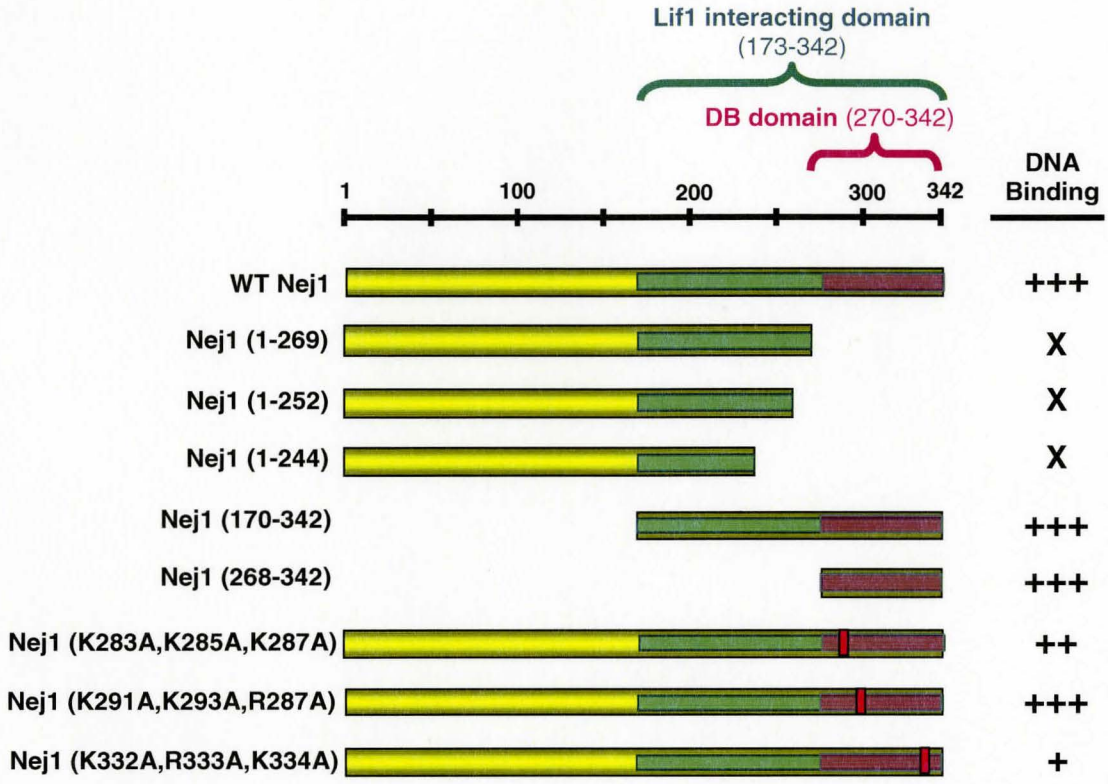

Figure 3.13 Summary of DNA binding data for all Nej1 deletion and point mutation constructs. Strength of DNA binding is stated in a qualitative manner where +++ is full DNA binding activity, ++ is slightly reduced DNA binding activity, + is moderately reduced DNA binding activity and X is no DNA binding activity. The Lif1 interaction domain (green) (Ooi *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007) and DNA binding (DB) domain (purple) are indicated.

3.6 Structural Characterization of Nej1

In order to investigate the nature of the observed co-operativity in DNA binding (Fig. 3.7, section 3.2), the quaternary structure of Nej1 was characterized. In a gel filtration experiment recombinant full-length Nej1 (monomer 39 kDa) eluted with an apparent molecular weight of 135 kDa, suggesting a Nej1 trimer (Fig. 3.14). However,

due to the inherent inaccuracy of gel filtration in predicting the molecular weight of irregularly shaped proteins, analytical ultracentrifugation was performed, under sedimentation equilibrium conditions, a technique in which the distribution of protein is entirely independent of the shape of the molecule. Analysis of the molecular weight of Nej1 by this method indicated that Nej1 exists as a dimer in solution, with a calculated molecular weight of 72 kDa and K_D of 3.9 μM (Fig. 3.15A).

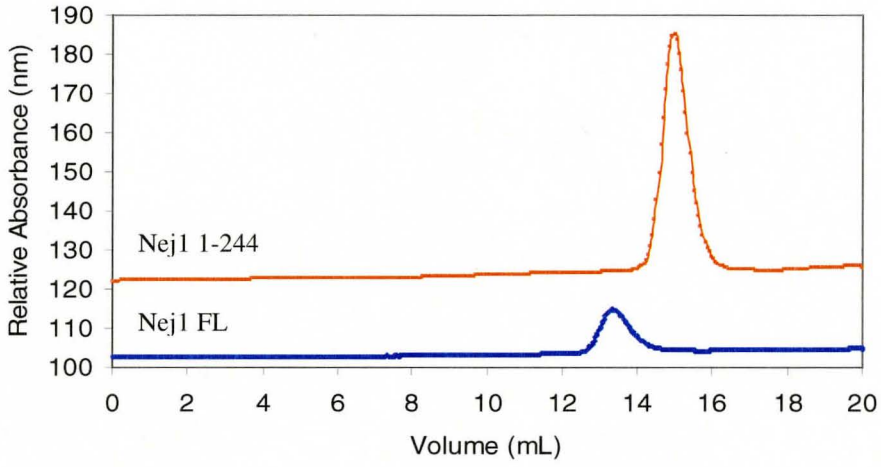

Figure 3.14 Nej1 exists as a dimer in solution. Gel filtration on full-length (FL) Nej1 resulted in the protein eluting with an estimated molecular weight of 135 kDa, while Nej1 1-244 eluted at 58 kDa, corresponding to a dimer. Volume = 0 represents point of injection and relative absorbance values are measured at 280 nm.

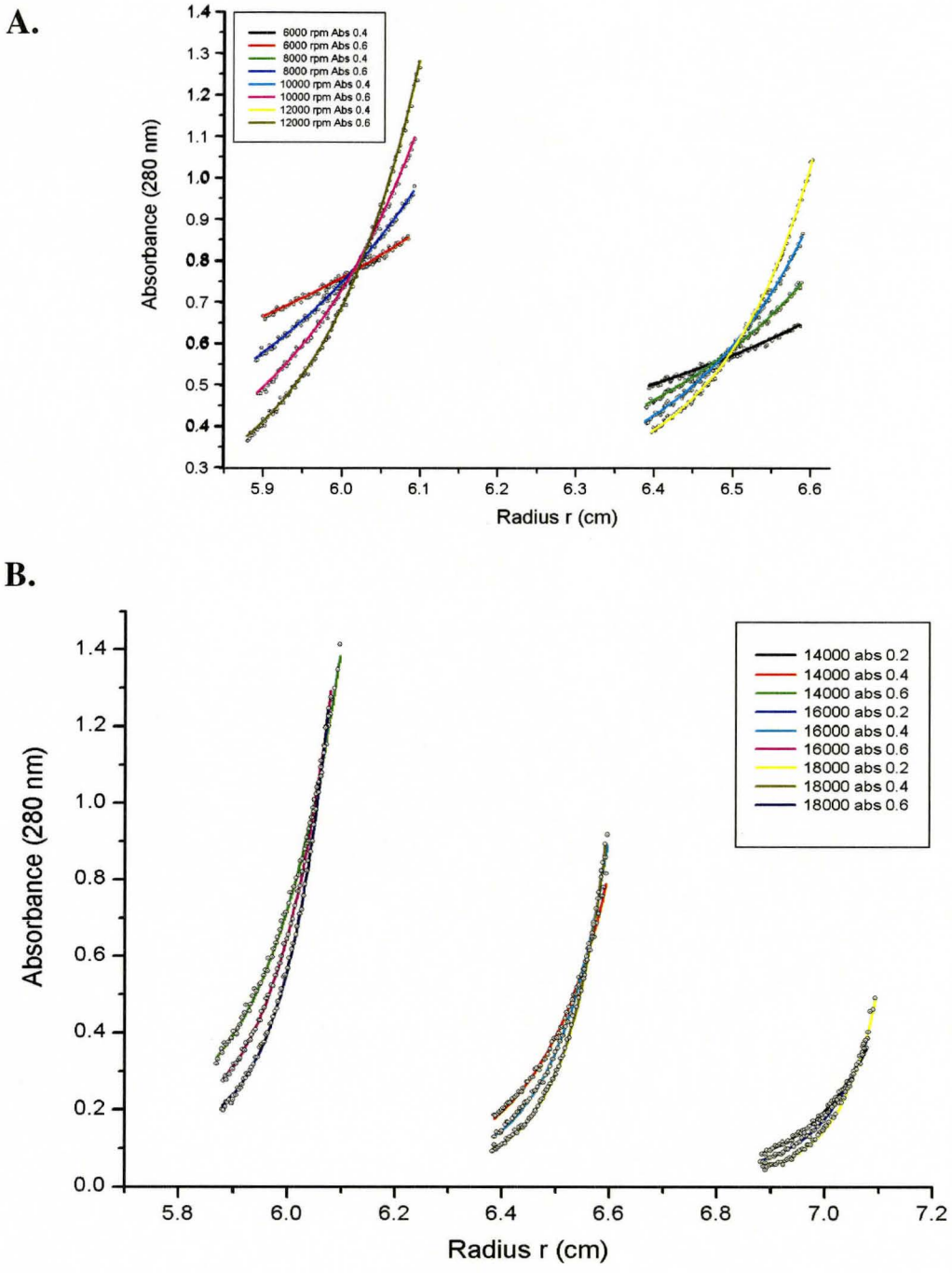

Figure 3.15 Analytical ultracentrifugation reveals that Nej1 is a dimer, and residues 1-244 are sufficient for dimerization. (A) Sedimentation equilibrium on full-length Nej1 resulted in a calculated molecular weight of 72 kDa, consistent with Nej1 existing as a dimer in solution. (B) Sedimentation equilibrium on Nej1 1-244 resulted in a molecular weight of 55 kDa, consistent with this region mediating Nej1-Nej1 contacts.

Gel filtration as well as analytical ultracentrifugation were also performed with Nej1 1-244 (Fig. 3.14 and 3.15B). Unlike full-length Nej1 which eluted as a trimer from gel filtration, Nej1 1-244 (monomer 28 kDa) eluted at a volume consistent with a molecular weight of 58 kDa, corresponding to a perfect globular dimer. The results from sedimentation equilibrium analysis of Nej1 1-244 indicated that this truncation was a dimer in solution, with a calculated molecular weight of 55 kDa and a K_D of 0.1 μM. These results indicate that the N-terminal 244 residues of Nej1 are sufficient for mediating homodimerization, and that this domain is not only stable (section 3.3) but also globular in nature. The relevance of the globular nature of Nej1 1-244 is discussed further in section 4.5. A summary of the functional and structural domains in Nej1 is provided in Fig. 3.16.

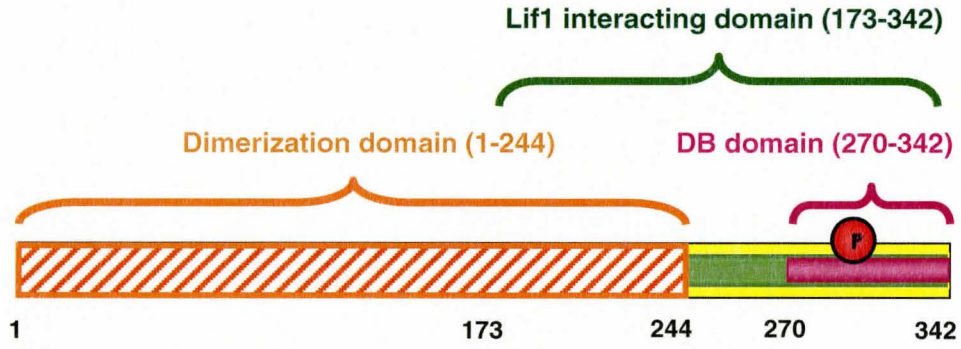

Figure 3.16 Summary of functional and structural domains in Nej1. The domains are colour coded as follows: Lif1 interacting domain spanning residues 173-342 (green) as previously determined (Ooi *et al.*, 2001; Frank-Vaillant *et al.*, 2001; Deshpande *et al.*, 2007), DNA binding (DB) domain containing residues 270-342 (purple), and dimerization domain containing residues 1-244 (orange stripes). Also indicated, is the phosphorylation site (red 'P') (Ser-297/8) (Ahnesorg *et al.*, 2006a). The full-length Nej1 protein is represented in yellow.

3.7 Studies of Nej1 in Complex with Lif1

Nej1 has been shown to interact with Lif1, both by yeast two hybrid analysis and immuno-precipitation (Ooi *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007); however, structural information regarding this interaction and its significance has not been revealed. To shed light on the functional significance of the Nej1-Lif1 interaction, a preliminary analysis of the structural and functional features of this complex was carried out.

Gel filtration was performed individually on purified samples of full-length Nej1 and Lif1, as well as following incubation of both proteins in equimolar amounts. On its own, Lif1 eluted close to the void volume of an S200 gel filtration column, suggesting the formation of a large complex and/or an unusually elongated shape with a large stokes radius (Fig. 3.17A). Based on the structure of Xrcc4 and its analysis by gel filtration (Modesti *et al.,* 1999), it seems reasonable that Lif1 would behave larger than expected for a globular protein with the same molecular weight. As previously described in section 3.6, Nej1 eluted as an apparent trimer. When Lif1 and Nej1 were applied simultaneously to a gel filtration column, there was no detectable change in the elution profile of either protein compared to their individual profiles. Thus, under these conditions, no interaction between Nej1 and Lif1 could be detected. Salt concentration had no effect on the observed elution profiles of either Lif1 or the mixture of Lif1 and Nej1 (data not shown). However, since the concentration of protein used in these studies is far above the dimer dissociation constant of Nej1, only dimeric Nej1 would have been available to interact with Lif1.

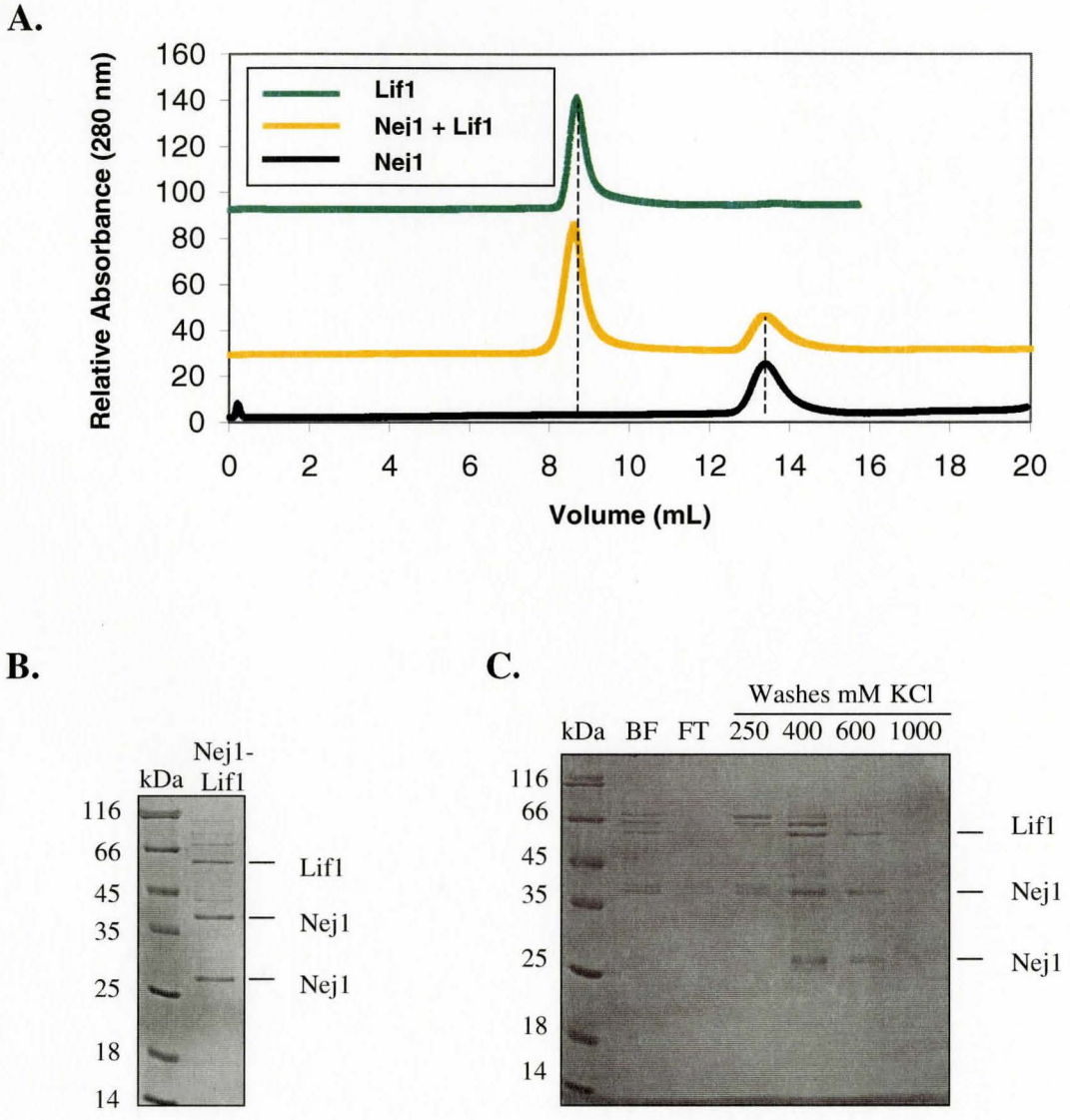

Figure 3.17 Characterization of Nej11-Lif1 full-length complex. (A) Gel filtration on Lif1 alone eluted close to the void volume of an S200, while Nej1 alone eluted with an apparent molecular weight of 139 kDa (trimer). Gel filtration performed on a mixture of Nej1 and Lif1 resulted in each protein eluting independently of the other, indicating no interaction. (B) A representative protein fraction from a Ni^{2+} charged Hi-Trap chelating column, shows co-purification of Nej1 and Lif1. (C) Q Sepharose protein fractions, illustrate the co-elution of Nej1-Lif1 complex in higher purity in the 600 mM KCl fraction. All protein fractions were analyzed via SDS-PAGE (12%, 150V, 1h) and stained with coomassie. Expected molecular weights are Lif1 His_6 – 51 kDa, Nej1 – 39 kDa, and cleaved Nej1 – 28 kD. BF- before filter, FT- unbound fraction.

Previous studies characterizing the interaction between Nej1 and Lif1 were performed under conditions where the proteins were co-expressed, rather than using individually purified recombinant proteins. To test whether a lack of interaction between Nej1 and Lif1 was a result of using independently expressed protein, full-length Nej1 and hexa-histidine tagged Lif1 were co-expressed. In-gel trypsin digestion followed by MS/MS was used to confirm the identity of both Nej1 and Lif1 as the major species observed in the Ni column elution fraction (Fig. 3.17B). Interestingly, in addition to full-length Nej1, a truncated Nej1 was also observed, possibly corresponding to the stable N-terminal domain generated during partial proteolysis (see section 3.3). This complex was stable to 1 M salt and co-purified stoichiometrically over a number of chromatographic steps (Fig. 3.17C). Therefore when co-expressed, Nej1 and Lif1 interact with the characteristics of a true, stable complex. When the purified Nej1-Lif1 complex was tested for DNA binding, interestingly the complex interacted with a 340 bp substrate in a co-operative manner (n = 2.0) and with higher affinity than each individual protein (Fig. 3.18).

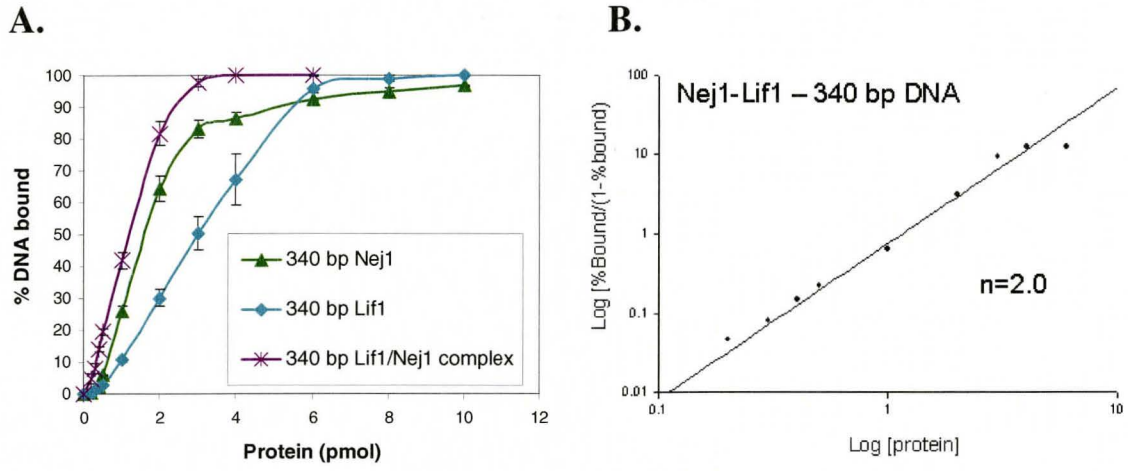

Figure 3.18 Nej1-Lif1 complex binds DNA with higher affinity than Nej1 or Lif1 and the interaction is co-operative. (A) DNA binding curve for Nej1-Lif1 complex showing higher affinity than Nej1 or Lif1 alone towards the 340 bp substrate. K_D of Nej1-Lif1 is 1.5 μM, compared to 1.8 μM for Nej1 and 3.0 μM for Lif1. (B) Hill-Plot generated for the DNA binding data for Nej1-Lif1 complex on 340 bp DNA, generating a Hill coefficient (n) of 2.0, and revealing a positive co-operative interaction. Error bars represent the standard deviation of 3 separate experiments.

Interestingly, when Lif1 was co-expressed with C-terminal truncations of Nej1 (1-244, 1-249 and 1-252), a stable complex could not be obtained, indicating that the N-terminal 244 residues are insufficient to promote formation of a stable Lif1-Nej1 complex (data not shown). Based on the analysis presented here it is difficult to say with certainty what quaternary structure is adopted by the co-purified Nej1-Lif1 complex. Although in principle the minimum quaternary structure of the complex may be a heterodimer, a heterotetramer (dimer of homodimers or dimer of heterodimers) and perhaps larger species may also exist (Fig. 3.19). The question regarding the quaternary structure of the XLF-Xrcc4 complex has not yet been answered, and similar models for their interaction have been proposed (Callebaut *et al.*, 2006).

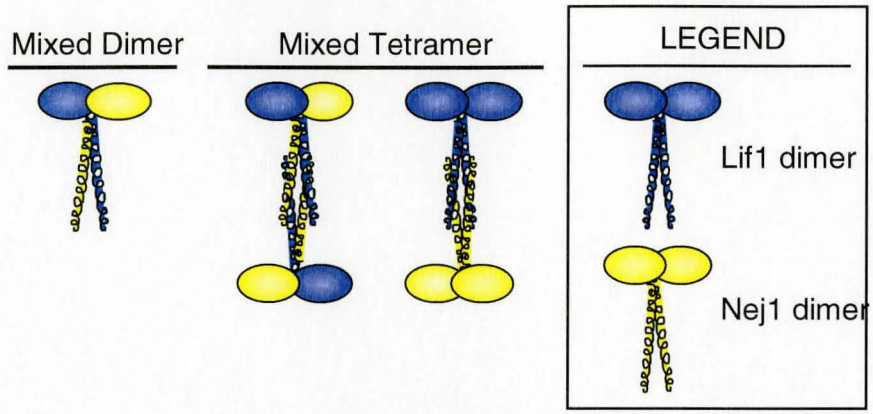

Figure 3.19 Models for Nej1-Lif1 interactions. Nej1-Lif1 complex could hypothetically form a number of different complexes, ranging from heterodimer, dimer of heterodimers, or dimer of dimers. These structural models are put forth based on the putative homology of Nej1 to XLF (Callebaut *et al.,* 2006), and on the proposed structural similarity of XLF, and thereby Nej1, to Xrcc4/Lif1 (Anhesorg *et al.,* 2006b). For structural information regarding XLF, Xrcc4 or Lif1 see Figs. 1.4, 1.5, 1.9 and 4.4.

3.8 Analysis of Nej1 Point Mutants and Deletion Mutants *In Vivo*

Having the ability to study the repair of DSBs in cells is a powerful tool. To study DSB repair *in vivo*, DSBs are usually induced by a variety of techniques, and their repair is subsequently monitored. The method of introducing DSBs can differ not only in the number and spacing of damage sites, but also in the type of DSBs that are produced and the extent to which other lesions and damages occur (reviewed in Wyman *et al.,* 2006).

In mammalian cells, DSB repair by NHEJ is commonly studied by inducing DSBs using ionizing radiation (Douglas *et al.,* 2005), microscope-coupled lasers (Uematsu *et al.,* 2007), or as a transiently transfected yeast I-*Sce*I endonuclease to generate site specific DSBs on a substrate plasmid (Willers *et al.,* 2006). Repair of DSBs

is monitored based on radioresistance of cells via cell survival assays, or by examining ligation products using PCR or sequencing methods.

Likewise, several biochemical assays have been developed to study NHEJ in *S. cerevisiae*. A common technique involves inducing site-specific DSBs using the yeast HO endonuclease (Lee *et al.,* 1998). This native *S. cerevisiae* endonuclease generates a single cut site at the mating type (MAT) locus that must be repaired in order for yeast to survive. This DSB is normally repaired by homologous recombination, using one of the two silent mating-type loci, HML and HMR. Thus, in order to prevent repair by HR, the yeast lack HML and HMR or are rad52 mutant strains, wherein the DSB can only be repaired by NHEJ. Similarly the yeast mitochondrial I-*Sce*I site specific endonuclease can also be used (Moore *et al.,* 1996). Importantly however, in these assays the expression of the endonucleases needs to be controlled and terminated in order to prevent re-cleavage of repaired sites (Karathanasis *et al.,* 2002).

Finally, a plasmid repair assay is also a popular approach to study end-joining *in vivo* (Boulton *et al.,* 1996). In this assay, a linearized plasmid, mimicking a DSB *in vivo*, containing a selectable marker is transformed into yeast and plated on appropriate selectable media. The linearized plasmid serves as a substrate for end-joining and must be circularized in order for yeast to survive under the selection pressure. Therefore, only yeast which successfully take up and repair the double-strand break (circularize the plasmid) will grow, and thus the transformation efficiency is a measure of NHEJ. Transformation efficiency is standardized by performing transformations with an un-cut plasmid.

Of these techniques, the most commonly used method of studying DSB repair in *S cerevisiae* is the plasmid-repair assay. Several groups have demonstrated that haploid *S. cerevisiae* lacking Nej1 are as defective in NHEJ as yeast containing deletions of Ku80, Lif1 or Dnl4 (Ooi *et al.,* 2001; Valencia *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001). Importantly, results from the plasmid repair assays correlate with end-joining assays in which DSBs were induced via endonucleases (Valencia *et al.,* 2001; Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001). Although these studies clearly demonstrate the importance of Nej1 in NHEJ, currently there is no evidence to suggest a specific mechanism by which Nej1 promotes double-strand break repair.

While it was valuable to analyze the various Nej1 deletion and point mutants *in vitro* (sections 3.3-3.5), it was also important to asses the effect of these mutants on NHEJ *in vivo*. Thus, to help elucidate the function of Nej1, Nej1 deletion and point mutants were investigated for their ability to perform NHEJ *in vivo*, via a plasmid repair assay. As expected, the *nej1* knockout strain was impaired in its ability to repair the double-strand break (circularize the linearized plasmid), resulting in a lower recovery of transformants, compared to the wild type strain (Fig. 3.20, bar 1 and 2). As previously reported (Ooi *et al.,* 2001; Valencia *et al.,* 2001), supplying a functional Nej1 via an expression plasmid (containing an endogenous Nej1 promoter) restored NHEJ back to wild type levels. In this assay Nej1 (K283A, K285A, K287A) and Nej1 (K291A, K293A, R295A), which exhibited only minor defects in DNA binding *in vitro*, had essentially no effect on the level of NHEJ *in vivo*. In contrast, Nej1 (K332A, R333A, K334A), which exhibited the strongest DNA binding defect *in vitro* (section 3.5), showed

a significant defect in NHEJ (for discussion see section 4.3). Consistent with their lack of *in vitro* DNA binding, all C-terminal truncation mutants were unable to support NHEJ *in vivo*. The fact that these same truncations were also unable to interact with Lif1 (section 3.7) makes it difficult to specifically attribute the NHEJ defect to the absence of DNA binding or Lif1 association. Nevertheless, these results clearly highlight the critical nature of the C-terminal domain *in vivo*.

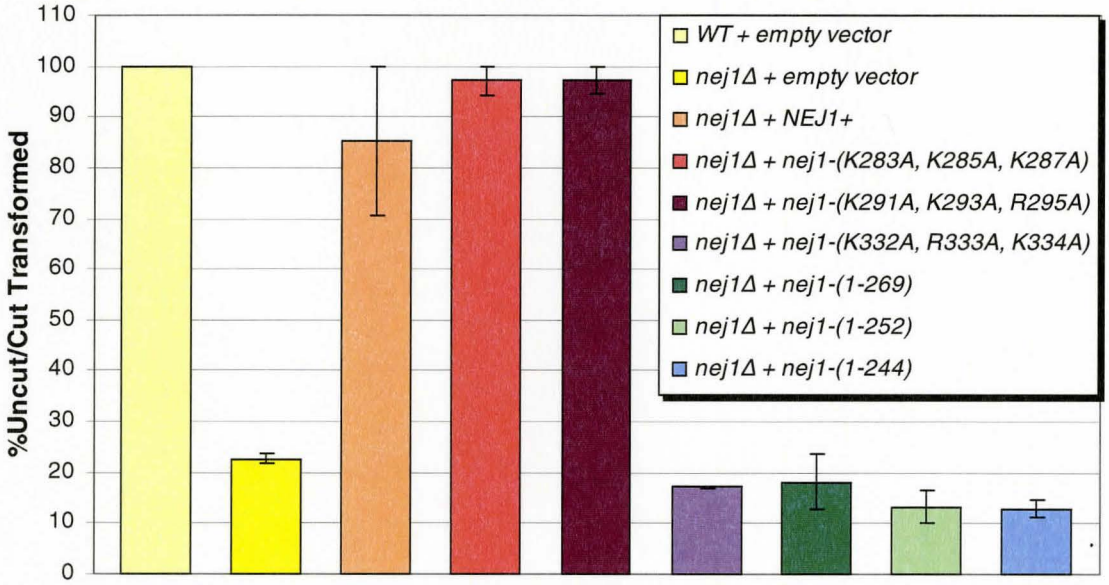

Figure 3.20 *In vivo* plasmid repair assay for Nej1 C-terminal deletion and point mutants. *nej1Δ* cells are unable to repair the double-strand break (bar 2) compared to wild type cells expressing Nej1 (bar 1). This defect is rescued by supplying a functional Nej1 on an expression vector (bar 3). Unlike Nej1 (K332A, R333A, K334A), both Nej1 (K283A, K285A, K287A) and Nej1 (K291A, K293A, R295A) were able to rescue the NHEJ defect in the *nej1Δ* cell line. All Nej1 C-terminal deletion constructs (1-269, 1-252 and 1-244) showed a drastic defect in end-joining.

To address the possibility that the complemented mutant strains have altered Nej1 protein expression compared to WT strain, and thus responsible for the functional defects seen in the *in vivo* assay, western blot analysis was carried out for each of the yeast strains used in the plasmid repair assay. Since an antibody towards Nej1 was not available commercially, and since the yeast expression vector, used to make each of the Nej1 yeast constructs, did not posses any tags that could be exploited with already available antibodies, a Nej1 specific antibody was required.

The Nej1 specific antibody was raised in a rabbit using purified recombinant Nej1 1-244 as the antigen, according to standard immunological guidelines (see materials and methods). The antibody generated was used to probe yeast lysates for the expression of Nej1 (wild type, deletion or point mutant) in each of the strains used previously for the *in vivo* plasmid repair assay. Unexpectedly, a band corresponding to the correct molecular weight of Nej1 could not be detected (Fig. 3.21, not all data shown). A number of variables, such as antibody titre and dilution, transfer efficiency, method of yeast lysate preparation, and lysate quantity, were addressed while troubleshooting this experiment (data not shown). However, despite numerous attempts, a conclusive result could not be obtained. To address the possibility that Nej1 may be in complex with DNA and remaining in the pellet during yeast lysate preparation, the pellet was analyzed by immunoblotting and still Nej1 was not identified (data not shown).

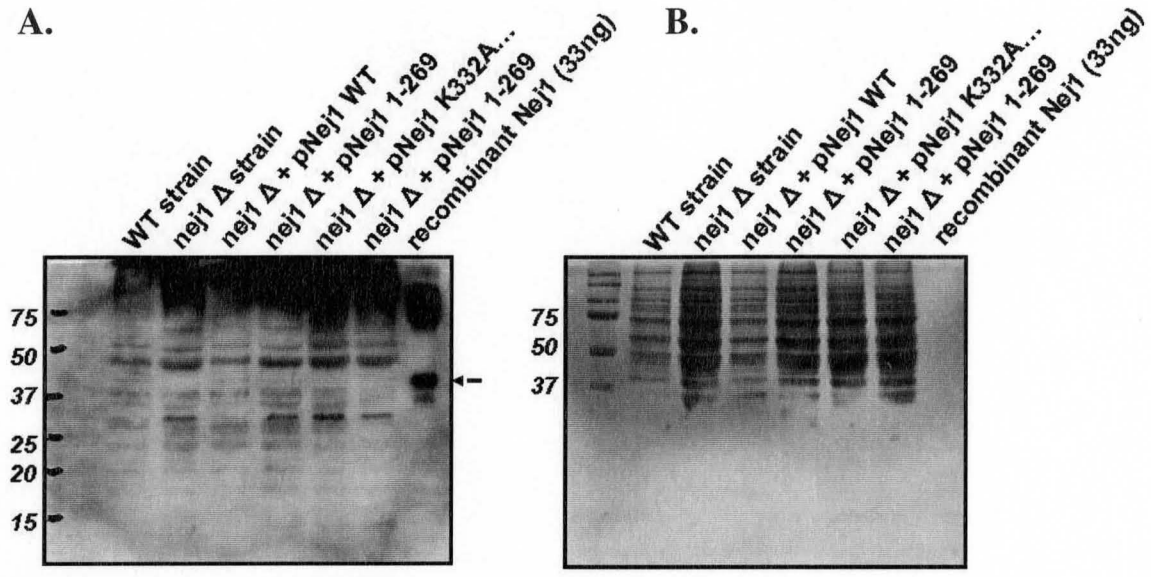

Figure 3.21 Immunoblot analysis of Nej1 protein expression. (A) Yeast lysates were prepared from the yeast strains used in the *in vivo* plasmid repair assay and probed for Nej1 expression using a polyclonal rabbit anti Nej1 antibody. Immunoblot fails to detect a band corresponding to Nej1 protein (39 kDa, arrow) despite the presence of various cross contaminating species. Corresponding coomassie stained gel after transfer is shown in (B). Immunoblot was performed with approximately 60 μg of yeast lystates.

The troubles encountered during the completion of the immunoblot may be understood by looking at previously published experimental data on Nej1. In the past, whenever Nej1 protein expression or mRNA levels were investigated, Nej1 was expressed using a non native promoter (MET25, ADH1, GDP1). Furthermore, for immunoblot of immunofluorescence assays, protein detection was aided with the use of a GFP or 13-Myc tagged Nej1 and their corresponding antibodies (Kegel *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Ahnesorg *et al.,* 2006). Conversely, in the immunoblot experiments described in this report, the approach relies on detecting endogenous levels of Nej1 using a polyclonal antibody against Nej1. Similar approaches, using non native

promoters or tagged protein, have been used to detect Lif1 in *S. cerevisiae*, and thus the problem of detection is not specific to Nej1. Additional studies will be required to troubleshoot this experiment and validate the results of the *in vivo* assay.

3.9 Nej1 X-Ray Crystallography Studies

Crystallization of Nej1-DNA complex

The ability to determine the precise molecular structure of biological macromolecules has revolutionized the study of their functions in many different fields. A widely utilized method for the determination of protein structure employs the crystallization of proteins, which are subsequently subjected to powerful X-rays to produce a diffraction pattern. The proteins' unique diffraction pattern is recorded and in turn analyzed using computer algorithms to generate a structural model of that protein (Rhodes 1993). An accurate, atomic-resolution structure can provide invaluable information regarding the mechanism of an enzyme, or offer clues about the functional relevance of a protein. Thus, to improve the current understanding of the role of Nej1 in end-joining, X-ray crystallography studies were pursued.

When full-length Nej1 was initially subjected to crystal trials by the hanging-drop vapor-diffusion method (McPherson *et al.*, 1995), this resulted in phase separation in 80% of the conditions tested. To overcome this unusual behaviour, a number of additives were tested and a pH titration was performed. No correlation between phase separation and pH was found, however, additives such as 10% glycerol and the addition of double-

stranded DNA had a marked effect in reducing the frequency of phase separation (data not shown). In subsequent purifications, 10% glycerol was included as a component of the exchange buffer for Nej1.

Having observed a positive effect with the addition of DNA in preliminary Nej1 crystal trials, suggested that Nej1 might be binding DNA and possibly creating favorable crystallization conditions. Since crystal growth is largely dependent on protein packing, it may be hindered by the presence of dynamic or unstructured molecule such as DNA (McPherson *et al.,* 1995). At the same time, the stability of the Nej1-DNA complex is dependent on the length of the DNA substrate, and may be compromised with shorter substrates (Tan *et al.,* 2000). Thus, to crystallize the Nej1-DNA complex that is not only stable but also optimally packs in a crystal lattice, co-crystallization experiments should be carried out using a range of DNA substrates (~ 15-21 bp).

In a preliminary screen, drops containing Nej1 or Nej1-DNA were set up side by side for 48 conditions using 15 bp DNA in a 1:1.3 molar ratio. From this initial screen, three conditions (Appendix I, Table A.IV, conditions #1-3) yielded crystals only in Nej1-DNA drops, and appeared ~10 days post set-up. The fact that the crystals grew in different conditions suggested that they may be protein crystals. Further, since crystal growth was dependent on the presence of DNA, it was promising that these crystals were Nej1-DNA complex crystals.

Crystals shown in Fig. 3.22A (grown in condition #31) were mounted and exposed to synchrotron radiation for 2 minutes, taking several exposures 0.5° wide, resulting in a diffraction pattern as seen in Fig. 3.22B. Lack of a high resolution

diffraction pattern suggests that the crystals are not salt (Howard *et al.,* 1992), while the presence of two diffraction spots at 30Å, in several images, is a promising preliminary result. Although subsequent crystal trials have yielded bigger and higher quality crystals (Appendix I, Table A.IV, conditions #4-11), further optimization is required to improve diffraction quality.

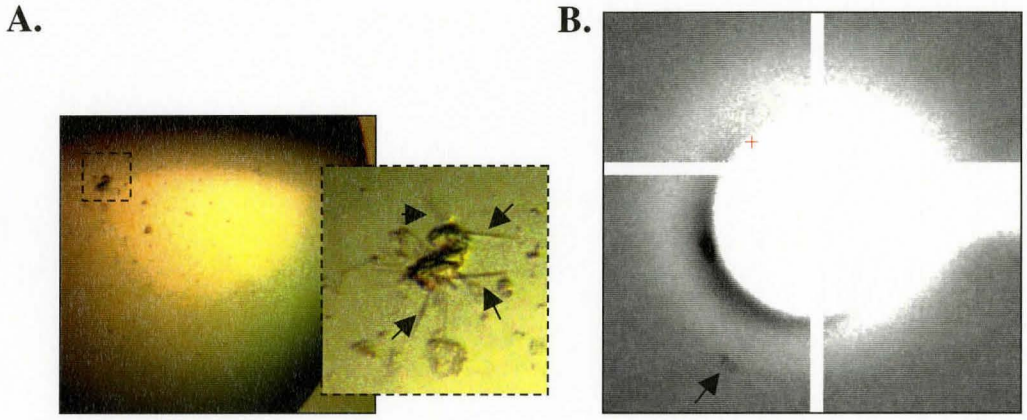

Figure 3.22 X-ray diffraction of Nej1-DNA complex crystals. (A) Nej1-15 bp DNA crystals grown in Sigma Membrane condition #31 (0.1 M $MgCl_2$, 0.1 M HEPES pH 7.5, 18% PEG 400). (B) Close up image of X-ray diffraction data, for a 2-minute, 0.5° wide exposure. Several images, representative of the one shown here, were recorded, revealing diffraction spots to 30Å resolution. Crystal Clear software (Rigaku) was used to display image in B.

Crystallization of Nej1 1-244

Since structural information of even a fragment of the protein can reveal valuable information, crystallization was attempted for the dimerization domain of Nej1. Nej1 1-244 was purified to near homogeneity, and following an initial 96-condition crystal screen (20°C) needle-shaped crystals grew in three different conditions (Appendix I,

Table A.V, conditions #1-3) 24 hours after set up (Fig. 3.23A). Following optimization of pH or precipitant concentrations (Howard *et al.,* 1992; McPherson *et al.,* 1995), and testing of ~72 different additives, the crystallization conditions were optimized leading to larger crystals grown in 0.2 M lithium sulfate, 0.1 M Tris pH7, 20 % PEG 4000 (Appendix I, Table A.V) in the presence of 0.1 M DTT (Fig. 3.23B and C).

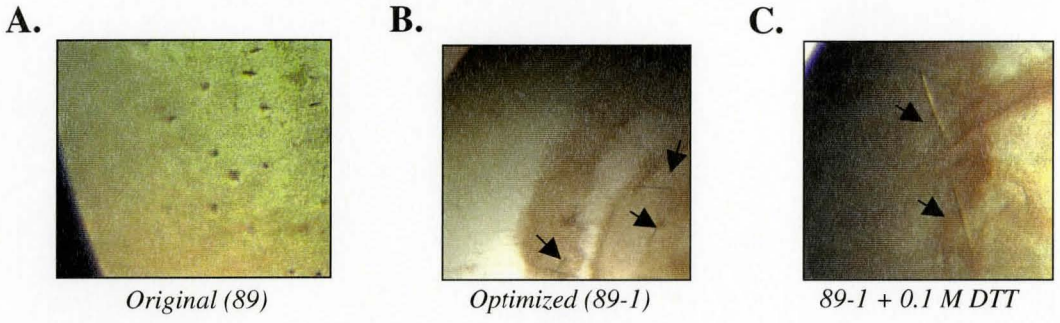

Figure 3.23 Crystals of Nej1 1-244. (A) Crystals observed for original condition #89 from Nextal's The Classics kit. These crystals are also representative of crystals obtained in condition 5. (B) Optimized crystals generated using a refined condition 89-1. (C) Crystals generated using condition 89-1 with 0.1 M DTT as an additive. (See Appendix I, Table A.V for condition information).

The crystal seen in Fig. 3.23C was subjected to home source X-rays, however the recorded images failed to reveal a diffraction pattern. In hopes of obtaining diffractable crystals, these crystals were also generated in similar conditions using seleno-methionine substituted Nej1 1-244. The largest crystals were exposed to synchrotron radiation in hopes of collecting a multiple-anomalous dispersion (MAD) data set, which would subsequently be used to obtain phase information, however no diffraction was detected. While lack of diffraction confirms that the crystals are not salt, it also indicates that the

crystals are poorly packed, and require further optimization to generate larger, better quality crystals.

Optimization of Nej1 1-244 crystals

In order to improve the likelihood of crystallizing Nej1 1-244, attempts were made to improve the purity of the protein, and thus the protein purification procedure was optimized. The new purification scheme resulted in a protein sample of superior purity (data not shown), and this protein was used in subsequent crystallization trials.

Broad screening of crystallization conditions was carried out, testing an initial 96-condition crystal screen, at 20°C and 4°C. Conditions that yielded promising leads are summarized in Appendix I, Table A.VI (conditions #1-3). A medium sized crystal obtained using 0.1 M MES pH 6.5, 1.6 M $MgSO_4$ (condition #3) (representative crystal shown in Figure 3.24A), was exposed to home source radiation for 15 minutes and subsequently for 30 minutes. Two images of diffraction data were collected, at 0° and 90°, 0.5° wide. The diffraction pattern in Figure 3.24B illustrates that the crystals diffracted to ~10Å resolution after both 15 minute and 30 minute exposures, indicating that the crystals' maximum resolution has been reached. Although this diffraction pattern is promising, a higher resolution diffraction (2-3Å) data set is required in order to determine the structure of Nej1 1-244. To achieve this, the initial crystallization condition which resulted in the growth of this crystal, needs to be optimized, to generate a larger crystal.

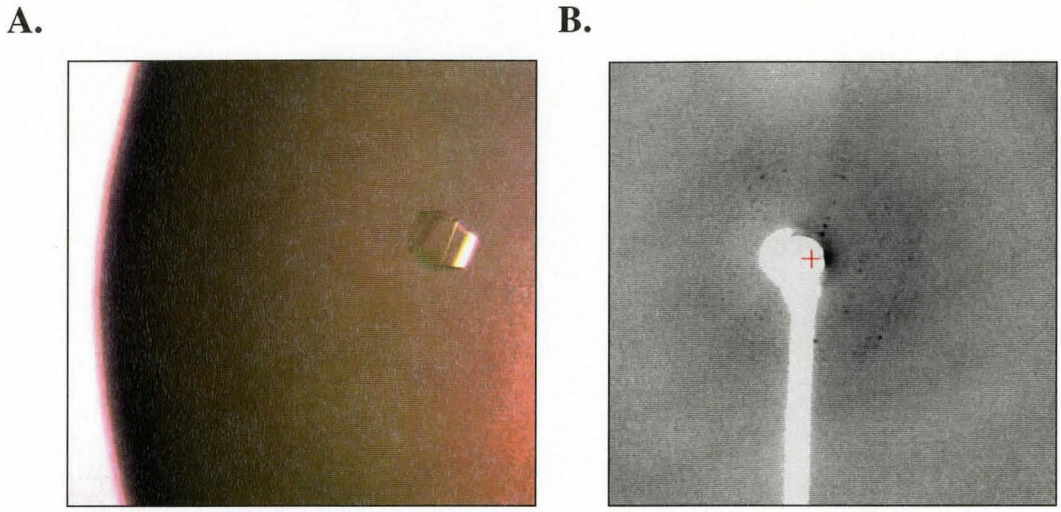

Figure 3.24 Optimized Nej1 1-244 crystals diffract to ~10Å resolution. (A) A small crystal of Nej1 1-244 representative of the one that was subjected to X-rays (crystal grown in 0.1 M MES pH 6.5, 1.6 M $MgSO_4$ at 4°C) (Appendix I, Table A.VI, condition #3). (B) Diffraction pattern recorded for a 30-minute exposure, illustrating ~10Å resolution. Crystal Clear software (Rigaku) was used to generate the image in B.

CHAPTER 4 – DISCUSSION

The experiments presented for Nej1 in this work will be discussed with reference to the binding partner of Nej1 (Lif1), the putative mammalian homologue called XLF, and the mammalian homologue of Lif1, Xrcc4. Furthermore, the results obtained will be highlighted in terms of the similarities and differences among the above proteins.

4.1 DNA binding properties of Nej1

The research communicated in this document serves as the first biochemical characterization of Nej1. Evidence is provided that the DNA binding activity of Nej1 is not only non-specific in terms of DNA sequence but also in terms of DNA structure. Further characterization of the DNA substrate length dependence of Nej1 revealed that although efficient binding was obtained on short double-strand DNA, the binding affinity towards larger DNA substrates increased (section 3.1). At this point, the mechanistic basis for this unusual preference for large DNA substrates is not clear. However, a similar DNA binding property was also found for Lif1 (section 3.2), as was previously reported for Xrcc4 (Modesti *et al.,* 1999) and more recently for XLF (Lu *et al.,* 2007). Collectively, these results suggest that this property of Nej1 may be real and also important. This preference may reflect the fact that a larger DNA fragment might promote more stable DNA-protein interactions due to structural feature common in each of these proteins (for discussion see section 4.5). Alternatively, this preference towards long DNA substrates may be linked to the idea that the DNA repair complex *in vivo*

encompasses a large section of DNA. In support of this hypothesis, are reports stating that in mammalian cells approximately 100 kb of DNA undergoes phosphorylation of the H2A histone surrounding the lesion (Burma *et al.,* 2006). In addition, phosphorylation of the H2A histone in *S. cerevisiae* has been linked to chromatin decondensation (Downs *et al.,* 2000), further implying that a large DNA segment is required for DSB repair.

Additional analysis of the DNA binding behaviour of Nej1 revealed that it is co-operative, suggesting that either Nej1 forms higher order complexes with varying DNA binding affinities, or that a Nej1 contains multiple DNA binding sites where the binding of DNA to the first site improves binding at the second site. The sedimentation equilibrium analysis reported here clearly indicates that Nej1 exists purely as a dimer in solution (section 3.6), suggesting that a single dimer contains at least two DNA binding regions that are allosterically regulated. The possibility of multiple sites on each monomer, however, cannot be excluded. This co-operative DNA binding behaviour is consistent with a protein that may function to bridge and possibly stabilize two DNA molecules during repair. Alternatively, or in addition to, this DNA binding activity may explain the requirement for Nej1 in preventing telomere fusions (Liti *et al.,* 2003), possibly through protein-protein and /or protein-DNA interactions.

4.2 XLF as the putative mammalian homologue of Nej1

Despite any obvious sequence similarity, using iterative BLAST and HCA manual alignment of multiple distantly related sequences, Callebaut *et al.* were able to provide compelling evidence to suggest that Nej1 may represent the *S. cerevisiae*

homologue of human XLF (Callebaut *et al.* 2006). Data presented here further supports and extends this proposed homology, illustrating that similar to XLF (Hentges *et al.*, 2006), Nej1 also possesses an ability to interact with DNA. Consistent with the behaviour of Nej1 towards DNA, a recent report by Lu and colleagues demonstrated that XLF also shares a preference for unusually long DNA (Lu *et al.* 2007). Indeed, initial studies of XLF DNA binding were conducted with 1000 bp DNA substrates. More recent data indicates that although 15 bp could theoretically suffice as a DNA binding substrate, a much larger DNA fragment (~80 bp) is required to form a stable complex. Thus although both Nej1 and XLF are similar in that longer DNA substrates are superior, Nej1 can interact stably with as little as 15 bp of DNA.

To date there have been no reports localizing the DNA binding site for any XLF family member. This study provides compelling evidence indicating that the C-terminus of Nej1 (aa 268-342) is responsible for interacting with DNA. Recent studies of XLF by Andres provide supporting data for the DNA binding domain being localized to the C-terminus (Andres 2007). Thus in terms of the location of the DNA binding domain, Nej1 and XLF appear to be similar.

Another similarity between Nej1 and XLF is that both proteins appear to interact with the same functional components, Nej1 with Lif1 in *S. cerevisiae* (Ooi, *et al.*, 2001; Kegel *et al.*, 2001; Frank-Vaillant *et al.*, 2001; Deshpande *et al.*, 2007) and XLF with Xrcc4-DNA Ligase IV in humans (Ahnesorg *et al.*, 2006b; Callebaut *et al.*, 2006). Surprisingly however, the interaction domains mapped out for these two pairs of interactions differ (Fig. 4.1). Whereas the C-terminus of Nej1 was demonstrated to be

necessary and sufficient for interacting with Lif1, the N-terminus of XLF was shown to suffice in interacting with Xrcc4 (Andres 2007; Deshpande *et al.,* 2007). Thus although at first glance the two proteins (Nej1 and XLF) share similarities in terms of DNA binding properties, there is a different structural organization in the interaction domains for their respective binding partners (Lif1 and Xrcc4). Although this argues for a difference in tertiary structure, it is possible that Nej1 and XLF adopt the same overall structure, and only differ in terms of the organization of functional domains.

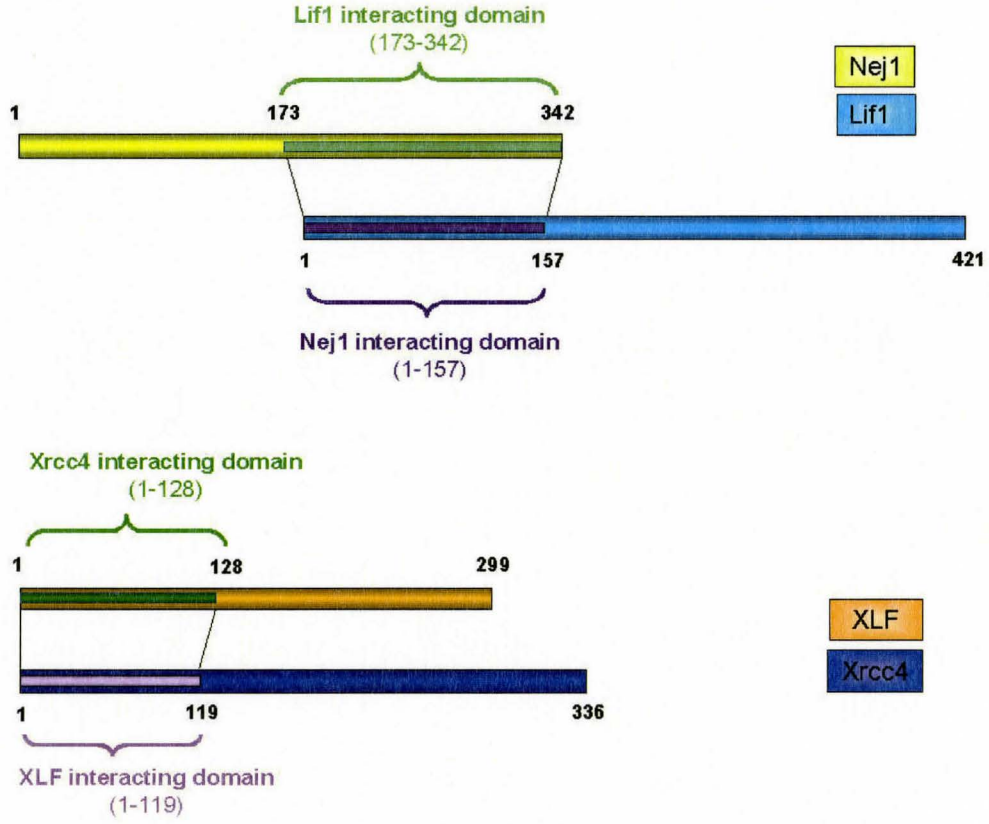

Figure 4.1 Minimal interaction domains for the contacts between Nej1-Lif1 and XLF-Xrcc4. The C-terminus of Nej1 interacts with the N-terminus of Lif1, whereas the N-terminus of XLF interacts with the N-terminus of Xrcc4 (Ooi *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007). Numbers represent amino acid residues.

4.3 The importance of Nej1 C-terminus for Lif1 association, DNA binding and non-homologous end-joining

Analysis of Nej1 structure via limited proteolysis studies indicated that the N-terminal 244 residues of Nej1 form a stable unit, whereas the C-terminal ~100 residues are susceptible to cleavage. Interestingly, the C-terminal ~70 amino acids were shown to be necessary and sufficient for mediating DNA interactions (section 3.3-3.5), raising the possibility that this region may become more structured upon DNA binding. However, when this possibility was tested by repeating the proteolytic treatment in the presence of DNA no change in the cleavage pattern was observed. This might indicate that Nej1 has a rapid on and off rate on DNA.

In addition to binding DNA, previous biochemical data illustrated that the C-terminus of Nej1 (aa 173-342) is necessary and sufficient for binding Lif1 (Ooi *et al.,* 2001; Frank-Vaillant *et al.,* 2001; Deshpande *et al.,* 2007). Co-expression and co-purification of Lif1 with various C-terminal truncations of Nej1 (aa 1-244, 1-249 and 1-252) did not result in the formation of a stable complex, although both proteins were expressed and available for interaction (data not shown). Thus, when the C-terminal region of Nej1 (aa 252-342) is absent, neither DNA binding nor stable Lif1 complex formation are observed. Although both DNA and Lif1 binding interactions are mediated through the same region of Nej1, it would appear that the binding of Lif1 does not preclude Nej1 DNA binding, since the DNA binding affinity of the Nej1-Lif1 complex was higher compared to either protein alone (section 3.7).

In vivo characterization of C-terminal Nej1 truncation mutants (Nej1 1-244, 1-252, and 1-269) incapable of both Lif1 interaction and DNA binding, confirmed the importance of this region in end-joining. Each of these mutants was unable to rescue the end-joining defect in *nej1* knockout yeast (section 3.8). Similarly, while full-length XLF was found to stimulate end-joining of non-compatible DNA ends, a truncation of XLF (1-224) was found inactive in this assay (Andres 2007). This finding provides support of the notion that Nej1 and XLF are genuine homologues and suggests that Nej1 may also be responsible for stimulating ligation by Lif1-Dnl4, a hypothesis that remains to be tested. In contrast to Nej1 and XLF, truncated human Xrcc4 (aa 1-200) not only retains DNA binding activity (Modesti *et al.,* 1999), but is able to fully complement the V(D)J defect in XR-1 cells *in vivo* (Mizuta *et al.,* 1997; Leber *et al.,* 1998). No information exists regarding the activity of full-length versus truncated forms of Lif1. While these structural-functional differences, between Nej1/XLF and Lif1/Xrcc4, argue against the proposed models for the structure of XLF (Ahnesorg *et al.,* 2006b), and in turn Nej1, it may be possible that the proteins are structurally related (see section 4.5).

Interestingly, the *in vivo* data revealed that although each of the point mutants had a modest decease in DNA binding *in vitro*, one mutant, specifically Nej1 (K332A, R333A, K334A), showed a drastic defect in NHEJ in the plasmid repair assay (section 3.8). A careful examination of the mutated residues and surrounding sequence in Nej1/XLF homologues of closely related species revealed that these residues are part of a conserved C-terminal motif (Fig. 4.2). This motif was recently identified among a number of XLF family members (Cavero *et al.,* 2007). This result provides further

evidence suggesting an essential role for this motif. The positively charged residues may be involved in mediating Nej1-DNA contacts (section 3.5), or may be significant for interacting with Lif1. It should be noted that the residues within this conserved motif resemble a NLS and therefore this *in vivo* analysis needs to be interpreted with this potential caveat. Also, the possibility that this protein was not expressed or is unstable cannot be excluded at this time since the immunoblot data did not provide evidence for the expression of this mutant (section 3.8). Nevertheless, the biochemical characterization presented here clearly demonstrates a requirement for this region of Nej1 in interactions with both DNA and Lif1.

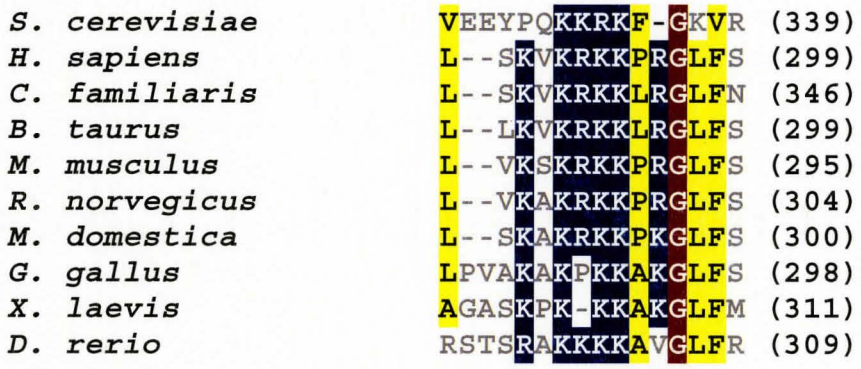

Figure 4.2 Alignment of the conserved C-terminal motif of the XLF family. Sequence corresponding to the absolute C-terminus of XLF proteins is shown. The residues in this motif may be involved in making contacts with DNA, and/or responsible for interacting with Lif1 and/or constitute an authentic NLS. Numbers in parentheses represent the last residue in each protein. Yellow-hydrophobic residues, Blue-positively charged residues, Maroon-conserved glycine.

Interestingly a recent study on XLF reported that a R57G mutation abolished the ability of XLF to translocate to the nucleus, resulting in a patient with NHEJ defects and

a SCID phenotype (Buck *et al.,* 2006; Lu *et al.,* 2007). Surprisingly, recombinant XLF (R57G) was able to stimulate Xrcc4-DNA Ligase IV ligation activity *in vitro* to the same extent as wild type XLF (Lu *et al.,* 2007). This new finding suggests that perhaps nuclear localization of XLF may be controlled by the putative C-terminal motif and/or R57 and surrounding residues. If this is true, this same mechanism may be in place in Nej1; thus without further experimentation, no conclusions can be made regarding what residues/region is responsible for nuclear localization of Nej1 and XLF.

4.4 Phosphorylation of Nej1 and its role in NHEJ

Ahnesorg and Jackson recently reported that Nej1 residues S297/8 are rapidly phosphorylated in response to DNA damage and that this phosphorylation is required for efficient NHEJ *in vivo* (Ahnesorg *et al.,* 2006a). Previous reports demonstrated that Xrcc4 is phosphorylated *in vivo* in a DNA-PK dependent manner (Modesti *et al.,* 1999). Surprisingly, phosphorylated Xrcc4 was able to stimulate end-joining by DNA Ligase IV to the same extent as unphosphorylated Xrcc4, even though phosphorylated Xrcc4 was unable to interact with DNA (Modesti *et al.,* 1999). Thus, while both Nej1 and Xrcc4 can be phosphorylated within their C-terminal regions, only phosphorylation of Nej1 appears to have a direct affect on NHEJ activity. Whether or not phosphorylation alters the DNA binding activity of Nej1 remains to be explored. The role of Nej1 phosphorylation in NHEJ requires further investigation. There is currently no data to regarding the phosphorylation XLF or Lif1.

There are a number of ways that phosphorylation within the C-terminal region of Nej1 could regulate NHEJ. However, given the functional importance of this region, an obvious possibility is that phosphorylation leads to altered Lif1 interaction (Ahnesorg *et al.,* 2006a), and/or altered DNA binding. Interestingly, the *in vitro* data presented in this report clearly suggests that phosphorylation is not required for either DNA binding or Lif1 association, since the proteins were generated recombinantly in *E. coli* and therefore were not phosphorylated. Although phosphorylation is not required for establishing these interactions, these contacts could potentially be stabilized upon phosphorylation, which in turn may be necessary for efficient end-joining (Fig. 4.3A and 4.3Bii). On the other hand, phosphorylation may function to promote dissociation of Nej1 from DNA (as is the case of Xrcc4) and/or the Lif1-Dnl4 complex, perhaps allowing for further access of downstream members to perform processing and/or ligation functions (Fig. 4.3Bi). Finally, phosphorylation of Nej1 may or may not be accompanied by a conformational change. Other outcomes as a result of Nej1 phosphorylation, besides the ones described above, are possible and cannot be excluded at this point. Additional experiments exploring the functional significance of Nej1 phosphorylation in end-joining are required to evaluate the mechanism of this regulation.

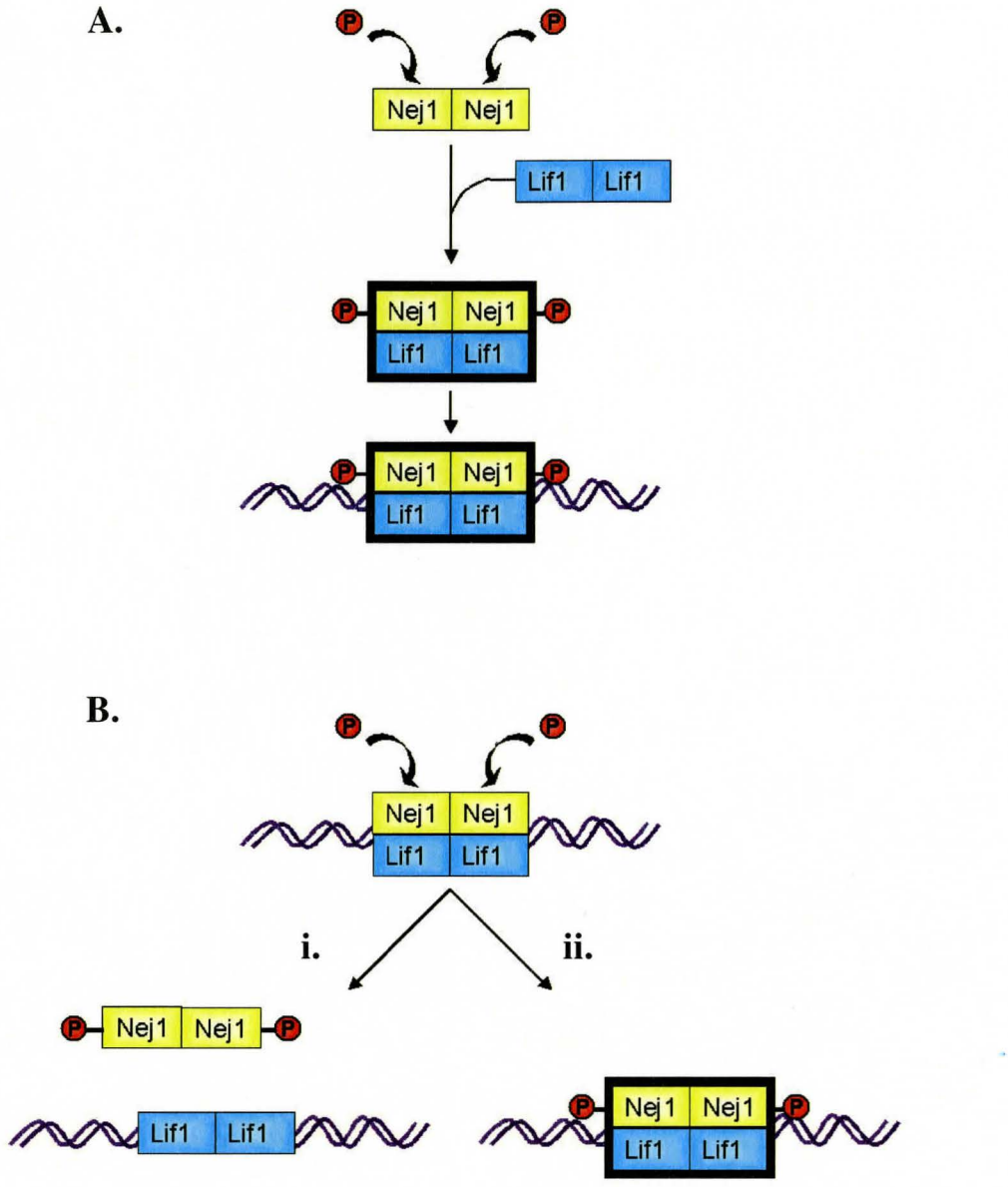

Figure 4.3 Phosphorylation of Nej1 in C-terminus and the possible functional outcomes. (A) If Nej1 is phosphorylated prior to binding DNA or Lif1, this modification may serve to enhance these interactions, since phosphorylation is required for efficient NHEJ (Ahnesorg *et al.,* 2006a). (B) Alternatively, if Nej1 is phosphorylated after coming in contact with either DNA and/or Lif1, phosphorylation may function to (i) dissociate Nej1 from DNA and/or Lif1-DNA Ligase IV complex, or (ii) stabilize these interactions so that DNA-Ligase IV can repair the double-strand break more efficiently. For simplicity DNA Ligase IV is omitted from the illustration. Phosphorylation of Nej1 may or may not be accompanied by a conformational change (not shown). Solid black lines represent interactions with greater stability.

4.5 Structure-function relationships of Nej1: similarities and differences to XLF, Lif1 and Xrcc4

In this work, evidence is presented indicating that the two families of proteins, Nej1/XLF and Lif1/Xrcc4, do share some functional similarities. However, this analysis also suggests that there appear to be significant differences surrounding the structure-function relationships, in particular within their C-terminal regions. The question remains then, to what extent are these two families of proteins structurally related?

XLF, and there by Nej1, have been proposed to be structurally similar to Lif1/Xrcc4 (Ahnesorg *et al.,* 2006b). The X-ray crystal structures of Xrcc4 (Junop *et al.,* 2000; Sibanda *et al.,* 2001; Meesala 2006) and Lif1 (Dore *et al.,* 2006) have been elucidated (Fig. 4.4). Xrcc4 exists as an unusually elongated molecule; overall dimensions of the Xrcc4 dimer are approximately 30 x 75 x 135 Å. The elongated, palm-tree like structure, is consistent with its large stokes radius observed in gel filtration experiments (Modesti *et al.,* 1999). Xrcc4 eluted from a gel filtration column with an apparent molecular weight six times greater than its true monomeric molecular weight. Recently, the structure of Lif1 was determined in complex with the tandem BRCT domains of Dnl4 (yeast homologue of mammalian DNA Ligase IV) (Dore *et al.,* 2006). The structure of Lif1 closely resembles Xrcc4, providing further evidence that these two proteins are genuine homologues.

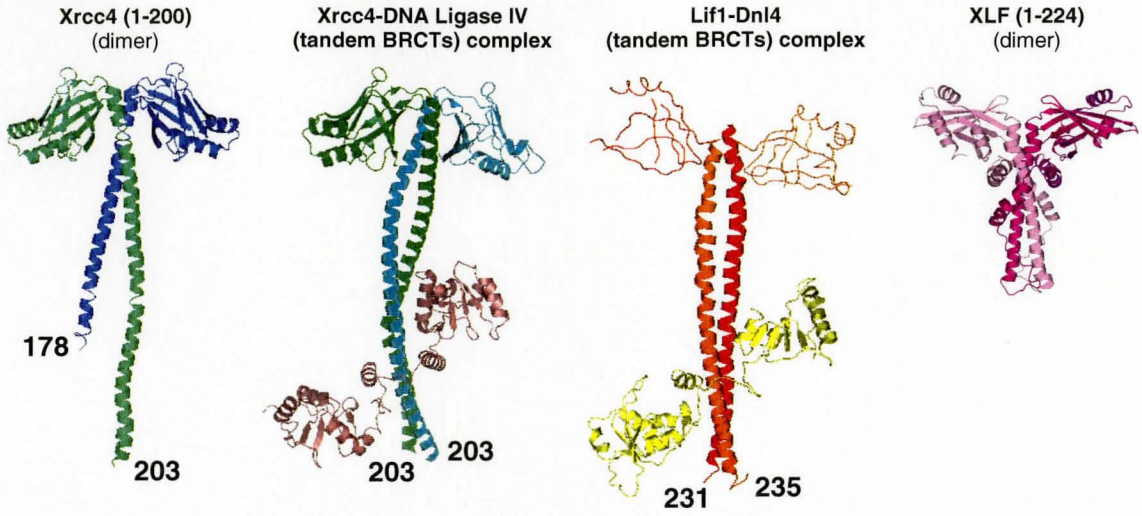

Figure 4.4 X-ray crystal structures of Xrcc4, Lif1 and XLF. Ribbon diagrams illustrating structural similarities between Xrcc4, Lif1 and XLF. Each is composed of a globular 'head' domain followed by an α-helical stalk. This stalk is interrupted in XLF and subsequently bent, causing it to wrap around underneath the head domain, resulting in a more globular structure. These structures further support the model of the structure of Nej1 as presented in Fig. 4.5 below. Furthermore, the X-ray crystal structure of XLF supports the initial model as predicted by Ahnesorg and colleagues (Introduction, Fig. 1.9) (Ahnesorg *et al.,* 2006b). PyMOL software was used to generate the pictures (De Lano *et al.,* 2002).

Nej1 is a dimer as determined by analytical ultracentrifugation, and analysis by gel filtration revealed that Nej1 1-244 behaved as a perfect globular dimer (section 3.6). Interestingly, both human and *S. pombe* XLF behave as globular, dimeric proteins in solution, exhibiting no unusual elution profile as judged by size-exclusion chromatography, unlike Xrcc4 which behaved as a hexamer (Hentges *et al.,* 2006; Modesti *et al.,* 1999). This is very surprising since the predicted structure of XLF family members is essentially identical to Xrcc4 and therefore one would instinctively presume that the behaviour of these proteins would be very similar on gel filtration

chromatography. It is difficult to reconcile how these proteins can behave as globular molecules and still maintain an Xrcc4-like structure. Thus if these two families of proteins do share similarity in terms of secondary structure it would appear that XLF family members (human XLF, *S. pombe* XLF and Nej1) do not share the same overall elongated structure observed for both Xrcc4 and Lif1. On the contrary, one would expect that the extended helical tail regions of Nej1 and XLF would fold back on themselves, resulting in a more globular structure (Fig. 4.5). In support of this foldback hypothesis, secondary structure predictions of Nej1 and XLF homologues reveal a short conserved coil region containing one or more glycines surrounded by several hydrophilic residues. This conserved, putative flexible region (~170-180) may function as a hinge allowing the protein to adopt a more globular structure. Further structural analysis of the XLF family will be needed to confirm the proposed foldback model suggested here and to provide greater insight into the function of these important NHEJ regulators, which now appear to function in part via direct contact of DNA and a NHEJ protein.

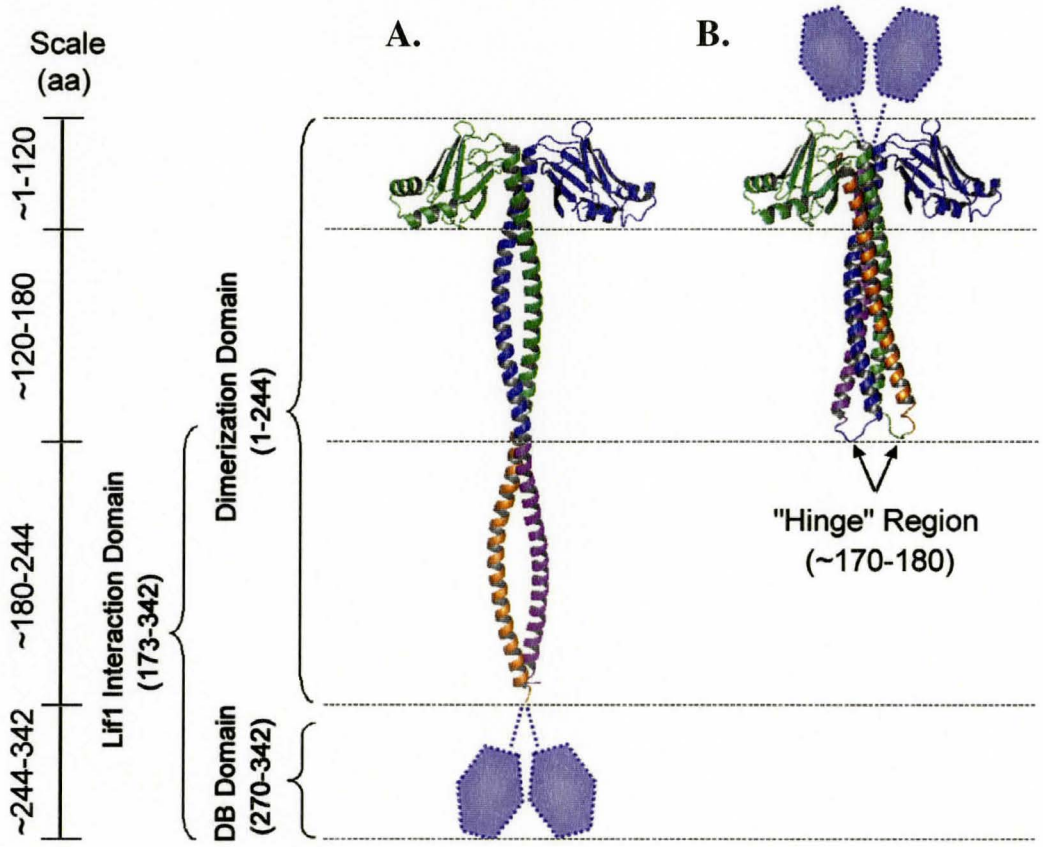

Figure 4.5 Model for the tertiary structure of Nej1. (A) Model of the "extended" version of Nej1 1-244. This "extended" model shows how the length of the C-terminal tails might be modulated by folding. In the form studied here, this C-terminal tail region cannot be fully extended as shown in model A, since the 1-244 truncation of Nej1 appears globular when analyzed by gel filtration chromatography. Instead, a proposed "folded" model is presented in (B), in which the C-terminal tails fold to create a more globular protein. The fold may be facilitated via a putative "hinge" in the ~170-180 aa region. Indeed the amino acid sequence of Nej1 contains several glycine and serine residues in this region (177-DGGS-180). The C-terminal domain (aa 244-342) is included in A to indicate that ~100 residues are missing from the full-length Nej1 protein. The location of this C-terminal domain, which is crucial for DNA binding, interaction with Lif1 and NHEJ *in vivo*, is unknown. Models were created based on the homology model previously reported (Ahnesorg *et al.,* 2006b). Further model manipulation was performed using the program WinCoot.

Supporting evidence for the model presented in Fig. 4.5 was recently obtained from structural studies on human XLF, which resulted in an X-ray crystal structure of this protein (Fig. 4.4) (Andres 2007). The structure of XLF illustrates similarities to the overall organization of Lif1/Xrcc4, in that it contains an N-terminal globular 'head' followed by an extended helical tail. This is in support of the XLF model initially predicted by Ahnesorg and colleagues (Introduction, Fig. 1.9) (Ahnesorg *et al.,* 2006b). Importantly however, the C-terminal stalk in XLF is interrupted and subsequently bent, causing it to wrap around underneath the 'head' domain, resulting in a more globular structure. This unique structural detail is in agreement with the biochemical data presented in this report and further supports the model presented in Fig.4.5.

CHAPTER 5 - SUMMARY AND FUTURE DIRECTIONS

5.1 Summary of Results

Initial published research on Nej1 indicated that this protein may function in regulating NHEJ in *S. cerevisiae*, and may be involved in shuttling Lif1 into the nucleus. Nevertheless, based on its nuclear localization, its interaction with Lif1 and its requirement for efficient NHEJ, it was hypothesized that Nej1 may be directly involved in the repair of DNA double-strand breaks. To address this possibility, Nej1 was purified from *E. coli* to near homogeneity and characterized functionally and structurally.

A functional characterization of Nej1 revealed that Nej1 exhibits a DNA binding activity, with an apparent dissociation constant of 1.8 μM *in vitro*. Additional characterization revealed that this DNA binding activity is both DNA-sequence and DNA-structure independent, displays co-operativity and increased affinity towards larger DNA substrates. To assess the similarities in the DNA binding properties between Nej1 and Lif1, similar studies were performed on Lif1. These studies suggested that Lif1 also exhibits a preference towards longer DNA substrates and like Nej1, also binds DNA in a co-operative manner. The discovery of co-operative DNA binding behaviour initiated an investigation into the quaternary structure of Nej1, using a combination of gel filtration and analytical ultracentrifugation. These studies showed that Nej1 exists as a dimer and that the N-terminus (aa 1-244) is responsible for dimerization. Finally, a preliminary

functional characterization of the Nej1 in complex with Lif1 indicated that this complex stably co-purifies, and binds DNA with higher affinity than each protein alone.

Structure-function studies using a number of Nej1 deletion constructs were able to locate the DNA binding domain to the C-terminal ~70 aa. Nej1 truncations lacking this C-terminal DNA binding domain were unable to restore NHEJ in *nej1* knockout yeast. These Nej1 truncations were also unable to form stable complexes with Lif1, which is consistent with previously published yeast-two hybrid data. Collectively, these results highlight the importance of the Nej1 C-terminus in end-joining. However, since the Lif1 interaction domain overlaps with the DNA binding domain, the effect observed *in vivo* cannot specifically be attributed to either loss of DNA binding or Lif1 interaction.

Point mutations were introduced into the C-terminal DNA binding domain, demonstrating a decrease in DNA binding affinity for each of the three Nej1 mutants tested. These same point mutants were also examined for their ability to perform NHEJ *in vivo*. Of the three point mutants tested, Nej1 (K332A, R333A, K334A) had both the strongest decrease in DNA binding activity *in vitro* and the strongest effect on end-joining *in vivo*. This mutant was unable to complement the *nej1* knockout, suggesting that these residues are critical for Nej1 function in end-joining. However, since these residues lie within the Lif1 interaction/DNA binding domain, they may not only be involved in promoting stable DNA contacts, but may also be contributing to stable Nej1-Lif1 interactions. Interestingly, these residues were found to be conserved among the XLF family of proteins and may represent an NLS. These issues should be addressed in

future studies, to be able to determine the precise role of these residues in end-joining and if the mutant proteins were indeed expressed.

The research presented in this report summarized an investigation into the mechanism of Nej1 in the NHEJ repair pathway. The results were evaluated in terms of the similarities and differences between Nej1 and members of the XLF and Xrcc4 protein families. Although the results support the proposed homology between Nej1 and XLF, they also highlight potential functional and structural differences between Nej1 and the XLF and Xrcc4 families of proteins.

5.2 Future Directions

Investigating conserved motif in C-terminus of Nej1

The strong effect observed for the Nej1 (K332A, R333A, K334A) *in vivo* suggests that either these residues are key for Nej1 function in end-joining, or that the mutant protein is not expressed. To determine the cause of this *in vivo* effect, it would be valuable to determine what role these residues play. Although mutating these residues decreased the DNA binding affinity of Nej1, the effect of these mutations on Lif1 interaction has not been investigated. Since the Lif1 interaction domain also resides in the C-terminus, it would be beneficial to asses the effect of these mutations on Lif1 binding. One of the ways that this could be investigated is to introduce these mutations in the Nej1-Lif1 co-expression plasmid, and attempt to purify the complex.

Another possibility worth investigating is the theory that the residues in this conserved C-terminal motif represent an authentic NLS. While it is reasonable that these residues, conserved among the XLF family members, may play a role in the nuclear localization of Nej1, a recent study on XLF found that XLF (R57G) was unable to localize to the nucleus although the purified, recombinant protein stimulated end-joining to the same extent as wild-type XLF. This result suggests that XLF, which also contains the conserved C-terminal motif, either contains additional NLS sites (~ aa 57) or means of regulating nuclear transport. To address this possibility, it would be beneficial to generate a Nej1-GFP fusion in yeast to monitor the localization of wild type Nej1 and Nej1 (K332A, R333A, K334A) using fluorescence microscopy.

Investigating the function of the C-terminus of Nej1

Since the C-terminus was found to be required for DNA binding, Lif1 interaction and end-joining *in vivo*, it would be interesting to see whether the C-terminus alone can rescue the NHEJ defect in *nej1* knockout yeast, using the already established plasmid-repair assay. Furthermore, since the C-terminal motif is a putative NLS, it would be beneficial to assay the activity of the C-terminus in the absence of this motif, in an effort to understand the role the conserved motif. Similarly, one could investigate constructs of the N-terminus of Nej1 (1-244) fused to the conserved C-terminal motif (aa 325-339), for their ability to perform end-joining, and follow up these studies by looking at the localization of these proteins in *S. cerevisiae*, via fluorescence microscopy.

Crystallization Studies

In order to improve the preliminary crystals that were generated in trials with Nej1 in the presence with 15 bp DNA, it should first be confirmed that these crystals are of the complex and not of either Nej1 or DNA. Once this is confirmed, optimization can proceed by altering the condition (pH, precipitant, salt), temperature and the presence of additives. In addition, a perhaps more critical aspect of generating better Nej1-DNA crystals is optimization of the DNA substrate itself. Trials with DNA substrates of various lengths and end-structure will explore a compromise between protein-DNA stability and optimal crystal packing.

Aside from altering the crystallization conditions or DNA substrate, Nej1 itself can be modified in such a way as to improve crystallization potential. It is a well-documented fact that clusters of positively charged residues found at the surface of the protein (often in predicted loop regions) add to the entropy of the system and may hinder crystallization. It is therefore possible to lower the entropy by introducing alanine substitutions at such sites in hopes of engineering a protein that will crystallize more readily. This strategy can be adopted in order to pursue crystallization of Nej1 on its own or in complex with DNA. Ideally, these mutants would need to be characterized biochemically to ensure structure and function have not been altered dramatically, prior to proceeding with crystallization trials. Ideal starting candidates for such studies are the already existing mutants, Nej1 (K283A, K285A, K287A) and Nej1 (K291A, R193A, K295A), which demonstrated no significant changes in DNA binding activity *in vitro*, secondary structure and ability to perform end-joining *in vivo*. Furthermore, since the C-

terminus is sufficient to interact with DNA, it would also be beneficial to pursue crystallization of a complex of the C-terminus of Nej1 bound to DNA.

In addition to pursuing crystals of Nej1 in complex with DNA, it would be valuable to obtain insight into the structure of the Nej1-Lif1 complex. This can be accomplished by utilizing the existing co-expression vector and optimizing the purification procedure to obtain a more homogeneous protein sample. Alternatively, a new co-expression construct can be generated, containing the minimal interaction domains, which may crystallize more readily. Finally, since the structure of XLF has been elucidated, it would be beneficial to obtain the structure of the XLF-Xrcc4 complex. Structural information regarding any of the above interactions would greatly benefit the understanding of Nej1/XLF function in the NHEJ pathway.

APPENDIX I

Table A.I Catalogue of all constructs used in this study, for either protein expression or *in vivo* assay, organized by protein of interest. Construct numbers are given (MJ#), their method of preparation, cloning or site-directed mutagenesis (SDM) and their purpose, protein expression or *in vivo* assay. The expected molecular weight (MW) of the protein in kilodaltons (no tag) and the fusion, N or C-terminal (if present) is given.

Protein (aa)	MJ#	Parental Vector	Cloned or prepared by SDM	For protein expression or *in vivo* assay	MW -tag (kDa)	Fusion (N or C-term)
^Nej1 (1-342)	4108	pPROEX-HTb	Cloned	prot. express.	39.1	H_6 - N
Nej1 (1-269)	4356	pPROEX-HTb	SDM on 4108	prot. express.	31.0	H_6 - N
Nej1 (1-252)	4357	pPROEX-HTb	SDM on 4108	prot. express.	28.9	H_6 - N
Nej1 (1-244)	4358	pPROEX-HTb	SDM on 4108	prot. express.	27.9	H_6 - N
Nej1 (K283A, K285A, K287A)	4318	pPROEX-HTb	SDM on 4108	prot. express.	39.1	H_6 - N
Nej1 (K291A, K293A, R295A)	4323	pPROEX-HTb	SDM on 4108	prot. express.	39.1	H_6 - N
Nej1 (K332A, R333A, K334A)	4319	pPROEX-HTb	SDM on 4108	prot. express.	39.1	H_6 - N
^Nej1 (1-342)	4427	pRS413	Cloned	*in vivo* assay	39.1	-
Nej1 (1-269)	4432	pRS413	SDM on 4427	*in vivo* assay	31.0	-
Nej1 (1-252)	4461	pRS413	SDM on 4427	*in vivo* assay	28.9	-
Nej1 (1-244)	4433	pRS413	SDM on 4427	*in vivo* assay	27.9	-
Nej1 (K283A, K285A, K287A)	4462	pRS413	SDM on 4427	*in vivo* assay	39.1	-
Nej1 (K291A, K293A, R295A)	4434	pRS413	SDM on 4427	*in vivo* assay	39.1	-
Nej1 (K332A, R333A, K334A)	4435	pRS413	SDM on 4427	*in vivo* assay	39.1	-
Nej1 (170-342)	4509	pLIC-His	Cloned	prot. express.	20.0	H_6
Nej1 (214-342)	4525	pLIC-His	Cloned	prot. express.	15.1	H_6
Nej1 (268-342)	4510	pLIC-His	Cloned	prot. express.	8.6	H_6
Nej1 (170-342)	4521	pLIC-His-NusA	Cloned	prot. express.	20.0	H_6NusA- N
Nej1 (214-342)	4522	pLIC-His-NusA	Cloned	prot. express.	15.1	H_6NusA- N
Nej1 (245-342)	4523	pLIC-His-NusA	Cloned	prot. express.	11.0	H_6NusA- N
Nej1 (268-342)	4524	pLIC-His-NusA	Cloned	prot. express.	8.6	H_6NusA- N
^Lif1 (1-421)	4102	pPROEX-HTb	Cloned	prot. express.	48.3	H_6 - N
Lif1 (1-231)	4458	pPROEX-HTb	SDM on 4102	prot. express.	27.0	H_6 - N
Lif1 (1-246)	4459	pPROEX-HTb	SDM on 4102	prot. express.	28.6	H_6 - N
Lif1 (1-302)	4460	pPROEX-HTb	SDM on 4102	prot. express.	35.2	H_6 - N
Nej1-Lif1	4216	pET28a	Cloned	prot. express.	39.1-48.3	Lif1 H_6–N
Nej1 (1-244)-Lif1	4427	pET28a	SDM on 4216	prot. express.	27.9-48.3	Lif1 H_6–N
Nej1 (1-249)-Lif1	4571	pET28a	SDM on 4216	prot. express.	28.3-48.3	Lif1 H_6–N
Nej1 (1-252)-Lif1	4572	pET28a	SDM on 4216	prot. express.	28.9-48.3	Lif1 H_6–N
Nej1-Lif1 (1-231)	4471	pET28a	SDM on 4216	prot. express.	39.1-27.0	Lif1 H_6–N
Nej1-Lif1 (1-246)	4464	pET28a	SDM on 4216	prot. express.	39.1-28.6	Lif1 H_6–N
Nej1-Lif1 (1-302)	4465	pET28a	SDM on 4216	prot. express.	39.1-35.2	Lif1 H_6–N

^ These constructs were not prepared by Sulek, M.

Table A.II PCR program and reaction conditions for site-directed mutagenesis.

Reaction Conditions		**PCR Program**	
5 µl	Pfu Buffer (10x)	1. 95°C	5'
1 µl	dNTPs	2. 95°C	40''
1 µl	Primer 1	3. 56°C	1'
1 µl	Primer 2	4. 65°C	10'
0.3 µl	Template DNA	*Repeat steps 2-4 x16*	
0.25 µl	Taq Polymerase	5. 64°C	40'
0.25 µl	Pfu Polymerase	6. 15°C	hold
41.25 µl	sterile water	*' – min, '' - sec*	
50.00µl	TOTAL		

Table A.III All primer sequences used in this study, organized by construct number (MJ#).

MJ #	Purpose of Primer	Primer MJ#	Primer sequence (5'→3')
4216	PCR amplify Lif1	2118	gaccgaagatctcgatcccgcgaaatta
		2119	ggacgcgtcgaccagcaaaaaacccctcaagaccc
4318	K283A, K285A, K287A in Nej1	2235	cgaccaaatagccgcggt**gcc**att**gcc**cca**gcc**acagatttcaagc
		2236	ggcttgaaatctgt**ggc**tg**ggg**caat**ggc**accgcggctatttggtcg
4319	K332A, R333A, K334A in Nej1	2253	cgaagaatacccacagaaa**gccgccgcc**tttggaaaggtgagaataaaaaac
		2254	gttttttattctcacctttccaaa**ggcggcggc**tttctgtgggtattcttcg
4323	K291A, K293A, R295A in Nej1	2251	ccaaagacagatttc**gcc**cca**gcc**tct**gcc**gaaagtagtaccagttcg
		2252	cgaactggtactactttc**ggc**aga**ggc**tg**ggg**caaatctgtctttgg
4356	A270Stop in Nej1	2338	catgattttgaattgcaa**taa**gaccctacaaatgag
		2339	ctcatttgtagggtc**tta**ttgcaattcaaaatcatc
4357	V253Stop in Nej1	2340	cttaacaaaaaagagcgc**taa**agatttcctgcgg
		2341	ccgcaggaaatct**tta**gcgctctttttgttaag
4427	G245Stop in Nej1	2342	cggtttgaagaatct**taa**gaacttaacaaaaaag
		2343	ctttttgttaagttc**tta**agattcttcaaaccg
4432	A270Stop in Nej1	2338	catgattttgaattgcaa**taa**gaccctacaaatgag
		2339	ctcatttgtagggtc**tta**ttgcaattcaaaatcatc
4433	G245Stop in Nej1	2342	cggtttgaagaatct**taa**gaacttaacaaaaaag
		2343	ctttttgttaagttc**tta**agattcttcaaaccg
4434	K291A, K293A, R295A in Nej1	2251	ccaaagacagatttc**gcc**cca**gcc**tct**gcc**gaaagtagtaccagttcg
		2252	cgaactggtactactttc**ggc**aga**ggc**tg**ggg**caaatctgtctttgg
4435	K332A, R333A, K334A in Nej1	2253	cgaagaatacccacagaaa**gccgccgcc**tttggaaaggtgagaataaaaaac
		2254	gttttttattctcacctttccaaa**ggcggcggc**tttctgtgggtattcttcg
4458	Q232Stop in Lif1	2404	gagcttcatgagattctgcga**taa**aataatatcaaattgtctgac
		2405	gtcagacaatttgatattat**ttta**tcgcagaatctcatgaagctc
4459	L247Stop in Lif1	2406	gatgttttagattctgcg**taa**attaatacggaagtgc
		2407	gcacttccgtattaat**tta**cgcagaatctaaaacatc
4460	E303Stop in Lif1	2408	ccgtaatcaaaatg**taa**gacgacgattttgatgac
		2409	gtcatcaaaatcgtcgtc**tta**cattttgatcacgg
4461	V253Stop in Nej1	2340	cttaacaaaaaagagcgc**taa**agatttcctgcgg
		2341	ccgcaggaaatct**tta**gcgctctttttgttaag
4462	K283A, K285A, K287A in Nej1	2235	cgaccaaatagccgcggt**gcc**att**gcc**cca**gcc**acagatttcaagc
		2236	ggcttgaaatctgt**ggc**tg**ggg**caat**ggc**accgcggctatttggtcg
4464	L247Stop in Lif1	2406	gatgttttagattctgcg**taa**attaatacggaagtgc
		2407	gcacttccgtattaat**tta**cgcagaatctaaaacatc
4465	E303Stop in Lif1	2408	ccgtaatcaaaatg**taa**gacgacgattttgatgac
		2409	gtcatcaaaatcgtcgtc**tta**cattttgatcacgg
4471	Q232Stop in Lif1	2404	gagcttcatgagattctgcga**taa**aataatatcaaattgtctgac
		2405	gtcagacaatttgatattat**ttta**tcgcagaatctcatgaagctc
4509	L170M in Nej1	2554	ccagggagcagcctcg**atg**atttctcttcgtgatttag
		2558	gcaaagcaccggcctcg**tta**gttttttattctcacctttcc
4510	L268M in Nej1	2557	ccagggagcagcctcg**atg**caagcagaccctacaaatg
		2558	gcaaagcaccggcctcg**tta**gttttttattctcacctttcc
4521	L170M in Nej1	2554	ccagggagcagcctcg**atg**atttctcttcgtgatttag
		2558	gcaaagcaccggcctcg**tta**gttttttattctcacctttcc
4522	K214M in Nej1	2555	ccagggagcagcctcg**atg**ttccagcatcaagaatttc
		2558	gcaaagcaccggcctcg**tta**gttttttattctcacctttcc

4523	G245M in Nej1	2556 2558	ccagggagcagcctcg**atg**gaacttaacaaaaaagagc gcaaagcaccggcctcg**tta**gttttttattctcacctttcc
4524	L268M in Nej1	2557 2558	ccagggagcagcctcg**atg**caagcagaccctacaaatg gcaaagcaccggcctcg**tta**gttttttattctcacctttcc
4525	K214M in Nej1	2555 2558	ccagggagcagcctcg**atg**ttccagcatcaagaatttc gcaaagcaccggcctcg**tta**gttttttattctcacctttcc

Constructs are colour coded as in Table A.I. Nej1 constructs- yellow, Lif1 constructs- orange, Nej1-Lif1 constructs – green. Bolded nucleotides represent location of changes made to the original sequence.

Table A.IV Crystallization conditions that resulted in crystals of the Nej1-DNA co-crystals. Top three are initial hits (#1-3), with conditions that yielded bigger crystals than initial hits listed below (#4-11). All leads were obtained using the hanging-drop vapor-diffusion method at 20°C.

#	Salt	Buffer	Precipitant	Source/code
1	0.1 M Magnesium chloride	0.1 M HEPES sodium salt pH 7.5	18 %(v/v) PEG 400	Sigma Membrane (31) (same as Nextal MbClass I (34))
2	0.1 M Magnesium sulfate, 0.1 M Potassium, Sodium-tartrate	0.1 M TRIS.HCl pH 8.5	-	Sigma Membrane (47)
3	1M Magnesium sulfate	0.1 M Sodium citrate pH 5.6	-	Sigma Membrane (16)
4	0.05 M Magnesium acetate	0.05 M Sodium acetate pH 4.6	25 %(v/v) PEG 400	Nextal MbClass I (36)
5	0.1 M Magnesium chloride	0.1 M Sodium acetate pH 4.6	18 %(v/v) PEG 400	Nextal MbClass I (32)
6	0.1 M Magnesium chloride	0.05 M TRIS.HCl pH 8.5	30 %(w/v) PEG 550 MME	Nextal MbClass I (40)
7	0.1 M Magnesium chloride, 0.1 M Sodium chloride	-	10 %(w/v) PEG 1500, 5 %(w/v) Ethanol	Nextal MbClass I (43)
8	0.1 M Magnesium chloride	0.05 M HEPES sodium salt pH 7.5	15 %(w/v) PEG 2000	Nextal MbClass I (49)
9	-	0.02 M Bis-TRIS propane pH 7.0	15 %(w/v) PEG 2000	Nextal MbClass I (48)
10	0.3 M Magnesium chloride	0.1 M BICINE pH 9.0	25 %(w/v) PEG 2000, 15 %(w/v) Glycerol	Nextal MbClass I (53)
11	0.3 M Magnesium nitrate	0.1 M TRIS.HCl pH 8.0	20 %(w/v) PEG 2000, 2 %(w/v) MPD	Nextal MbClass I (52)

Numbers in parentheses correspond to the number of the condition in the kit.

Table A.V Crystal conditions for Nej1 1-244. Initial conditions are #1-3, optimized conditions are 4-7, and leads from Hauptman Woodsworth automated screening are #8-11. All leads were obtained using the hanging-drop vapor-diffusion method at 20°C.

#	Salt	Buffer	Precipitant	Source/code
1	-	0.1 M HEPES pH 7.5	10% isopropanol 20 %(v/v) PEG 4000	Nextal The Classics (5)
2	0.2 M Lithium sulfate	0.1 M TRIS pH 8.5	30% PEG 4000	Nextal The Classics (89)
3	-	0.1 M HEPES pH 7.5	20% PEG 10000, 8% Ethylene glycol	Nextal The Classics (95)
4	-	0.1 M HEPES pH 7.5	6% isopropanol 20 %(v/v) PEG 4000	Optimized (5-4)
5	0.2 M Lithium sulfate	0.1 M TRIS pH 7.0	20% PEG 4000	Optimized (89-1)
6	0.2 M Lithium sulfate	0.1 M TRIS pH 7.5	20% PEG 4000	Optimized (89-4)
7	0.2 M Lithium sulfate	0.1 M TRIS pH 7.5	30% PEG 4000	Optimized (89-4-6)
8	0.2 M Sodium iodide	-	20% PEG 3350	Hauptman Woodsworth (HW) #1138
9	0.2 M Potassium iodide	-	20% PEG 3350	HW #1139
10	0.2 M di-ammonium hydrogen phosphate	-	20% PEG 3350	HW #1172
11	0.8 M lithium sulfate monohydrate	0.1 M bis-tris propane ph 7	-	HW #1508

Numbers in parentheses correspond to the number of the condition in the kit.

Table A.VI Promising conditions for Nej1 1-244 that yielded crystals. The crystal generated in condition #3 was big enough to be suitable for analysis by X-ray diffraction. Conditions #1 and 2 correspond to Nextal/Qiagen's Classics kit conditions #20 and 62 respectively. Condition 4 corresponds to Crystal Magic II kit condition #8.

#	Condition	Additive	Temp	Crystal Morphology
1	0.5 M NH_4SO4, 0.1 M Hepes pH 7.5, 30% MPD	-	20°C	Small needles
2	0.1 M MES pH 6.5, 1.6 M $MgSO_4$	-	20°C	small 3D crystal
3	"	-	4°C	medium 3D crystal
4	0.1 M Tris pH 8, 40% MPD	20 % ethylene glycol	4°C	Tiny needles
5	"	0.2 M MgFormate	4°C	Tiny needles
6	"	0.5 M NaCl	4°C	Tiny needles
7	"	10 mM DTT	4°C	Tiny needles

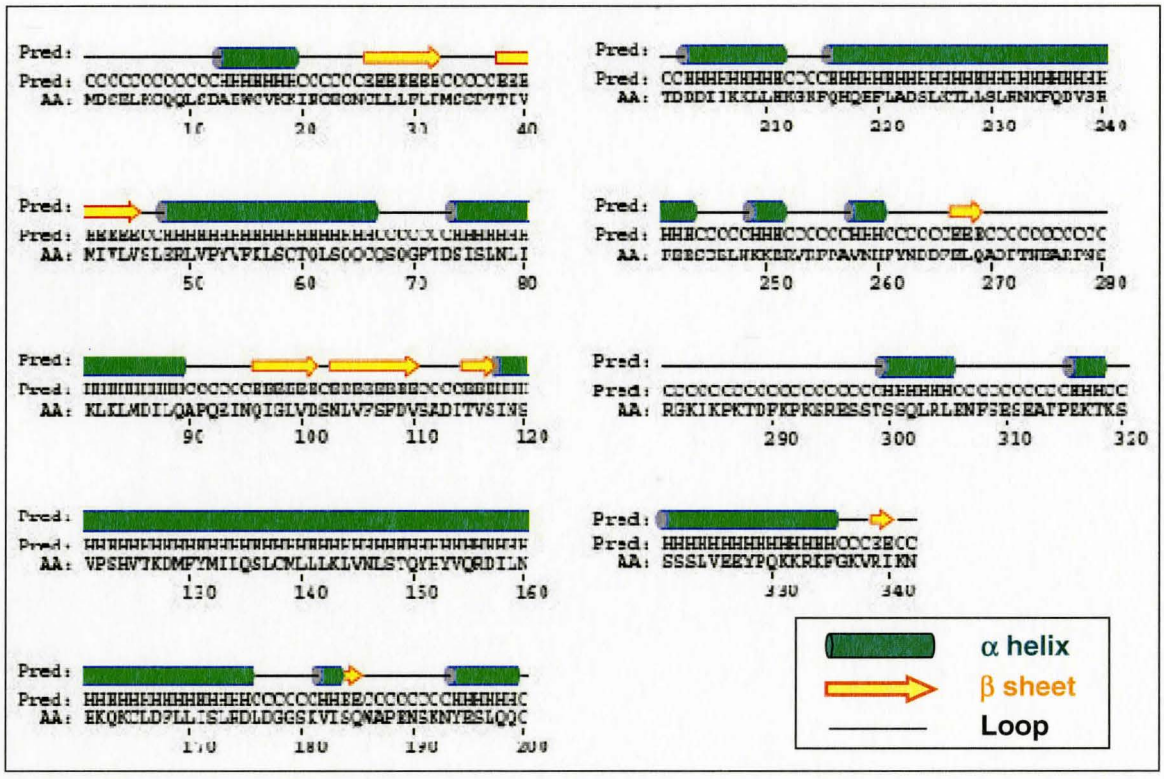

Figure A.I Secondary structure prediction for Nej1. Nej1 C-terminus (~ aa 244-342) is predicted to be the least structured, containing mostly loops. Displayed are regions predicted to exist either as an α–helix ('H', green rods), β-sheet ('E', yellow arrows) or random coil ('C', black line). Secondary structure prediction obtained using the program PsiPred, which was modified to create this image.

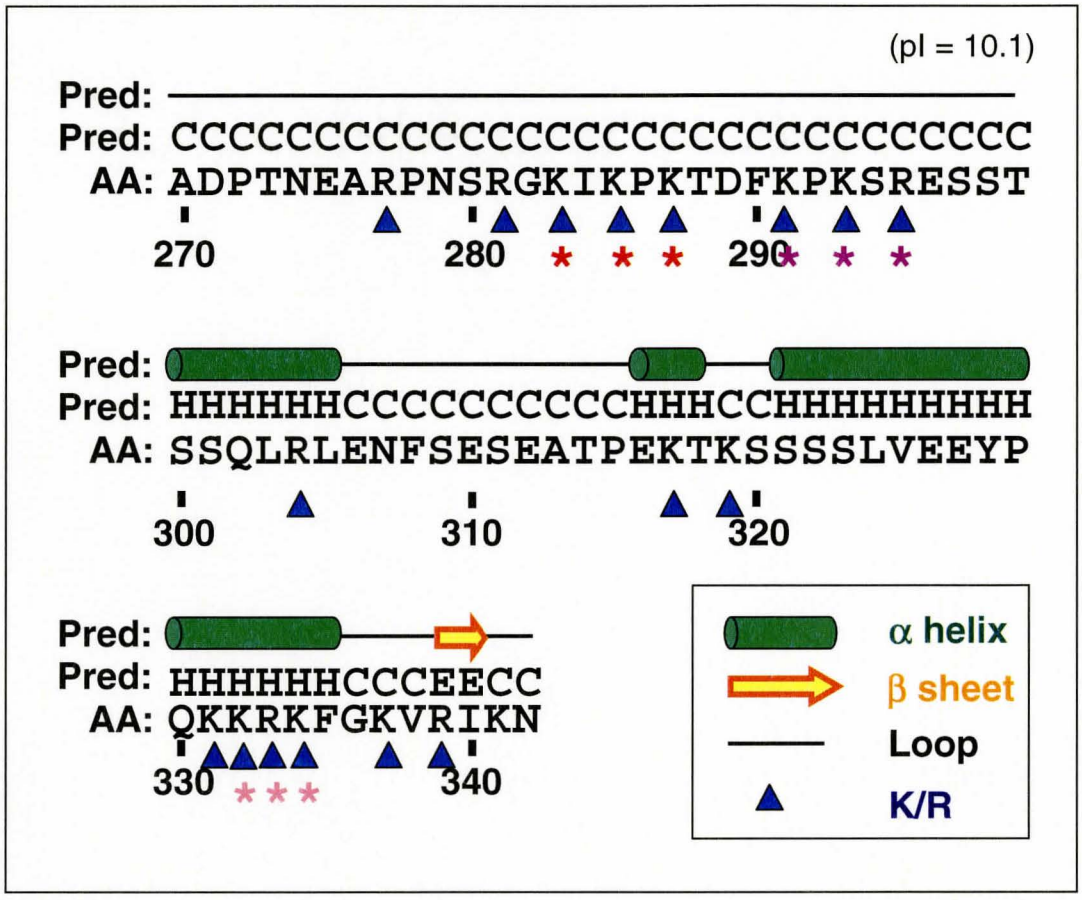

Figure A.II C-terminus of Nej1 is abundant in positively charged residues. The pI of this fragment is 10.1 due to the many positive residues located in this region. Within this region, three clusters of positively charged residues were targeted for mutagenesis to yield Nej1 (K283A, K285A, K287A) (red stars), Nej1 (K291A, K293A, R295A) (purple stars) and Nej1 (K332A, R332A, K334A) (pink stars). Positively charged residues (K/R) are highlighted with blue triangles. Location of predicted α helices ('H', green rods), β sheets ('E', yellow arrows) and loops ('C', black lines) is indicated. The secondary structure prediction was obtained using the program PsiPred, and subsequently modified to generate this illustration.

APPENDIX II

This Appendix contains supplementary information regarding the purification of Pfu and Taq Polumerases, and TEV protease.

Purification of Pfu Polymerase

Pfu Polymerase was over-expressed from pfu-pET15b in BL21 (DE3) pLysS cells. Protein expression was induced at OD_{600} of 0.5, for 3 hours in the presence of 1 mM IPTG, at 37°C. For purification, cells were resuspended in the presence of 20 mM Tris pH 8.5, 0.03% LDAO, 500 mM KCl, 10% glycerol and 10 mM imidazole. Following the addition of protease inhibitors the cells were lysed by four passages through a French Press cell at 12 000 psi. The lysate was cleared by centrifugation (15 000 g, 40 min) and the supernatant was heated at 75°C for 15 minutes. Subsequently, the supernatant was mixed by inversion and placed on ice for 15 minutes. After this, the sample was centrifuged (15 000 g, 20 min), the protease inhibitors were added again and the sample was applied to a Nickel column equilibrated in the above buffer. The column was washed with four column volumes of the above buffer, and the protein was subsequently eluted using an 18 column volume gradient to 20 mM Tris pH 8.5, 0.03% LDAO, 500 mM KCl, 10% glycerol and 300 mM imidazole. The eluted sample was diluted to a final salt concentration of 100 mM KCl, in the presence of 50 mM MES pH 5.6, 1 mM EDTA, and 10 mM DTT. TEV was eluted using a gradient from 100 mM KCl to 500 mM KCl over a 150 mL volume. The eluted fractions were analyzed by SDS-

PAGE, the appropriate fractions were pooled, concentrated to 0.175 mg/mL and frozen at -20°C in glycerol.

Purification of Taq Polymerase

Taq was over-expressed from pET28a-Taq in BL21 (DE3) cells. Cells were grown at 37°C to OD_{600} of 0.5, after which protein induction was induced with 1 mM IPTG for 3 hours at 37°C. For purification, cells were resuspended in a buffer containing 20 mM Tris pH 8.5, 10 mM imidazole, 500 mM KCl, 0.06 % LDAO and 10 % glycerol, in the presence of 1 mM of each protease inhibitor (pepstatin, PMSF, leupeptin and benzamidine). Cell lysis was achieved using four consecutive passes through a French Press cell at 12 000 psi. Following the addition of all protease inhibitors, the lysate was cleared by centrifugation (15 000 g, 20 min) and the supernatant treated with 1 mM of benzamidine and leupeptin. The sample was filtered and applied to a Nickel charged column. The column was washed with the above buffer and subsequently a gradient, from 10 mM to 300 mM imidazole (60 minutes, 1.5 mL/min), was used to elute Taq. Collected fractions were assessed for purity using SDS-PAGE. The fractions of greatest purity were pooled, the salt was reduced to 100 mM KCl and the sample was applied to a SP Sepharose column in the presence of 50 mM MES pH 6.5, 1 mM EDTA, 5 mM DTT and 100 mM KCl. A gradient from 100 mM to 500 mM KCl (100 min 1.5 mL/min) was used to elute Taq and the fractions collected were analyzed for purity by SDS-PAGE. Appropriate fractions were pooled, concentrated to 1.5 mg/mL and stored at –80°C.

Purification of TEV protease

TEV was over-expressed from pRK793 containing His-TEV(S219V)-Arg in BL21-RIL cells. Protein expression was induced at OD_{600} of 0.5, for 4 hours in the presence of 1 mM IPTG at 30°C. For protein purification, the cells were resuspended in 50 mM Phosphate pH 8.0, 100 mM NaCl, 10% glycerol and 25 mM imidazole. Following the addition of protease inhibitors the cells were lysed by four passages through a French Press cell at 12 000 psi. The volume of the lysate was determined and 5% of polyetheleneimine was added to a final concentration of 0.1%. The lysate was cleared by centrifugation (15 000 g, 40 min), and the supernatant was applied to a Nickel column equilibrated in the above buffer. The column was washed with 7 column volumes of the above buffer, after which the TEV protease was eluted using a 10 column volume gradient to 50 mM phosphate pH 8.0, 100 mM NaCl, 10% glycerol and 200 mM imidazole. Fractions were analyzed by SDS-PAGE and pooled. EDTA and DTT were subsequently added to the sample to a final concentration of 1 mM. Next, the sample was concentrated and loaded onto a SP-sepharose column in the presence of 25 mM phosphate pH 8.0, 200 mM NaCl, 10% glycerol, 2 mM EDTA and 10 mM DTT using 3% of the column volume. Purity of the fractions was determined using SDS-PAGE and the appropriate fractions were pooled. The protein was concentrated to 1.0 mg/mL and frozen at -80°C in 50% glycerol.

REFERENCES

Ahnesorg, P., Jackson, S.P. The non-homologous end-joining protein Nej1p is a target of the DNA damage checkpoint. *DNA Rep.* 6, 190-201 (2006a).

Ahnesorg, P., Smith, P., Jackson, S.P. XLF interacts with the Xrcc4-DNA Ligase IV complex to promote DNA nonhomologous end-joining. *Cell.* 124, 301-313 (2006b).

Andres, S. Personal Communication, (2007).

Ataian, Y., Krebs, J.E. Five repair pathways in one context: chromatin modification during DNA repair. *Biochem. Cell Biol.* 84, 490-504 (2006).

Aylon, Y., Kupiec, M. DSB repair: the yeast paradigm. *DNA Rep.* 3, 797-815 (2004).

Bailey, S.M., Meyne, J., Chen, D.J., Kurimasa, A., Li, G.C., Lehnert, B.E., Goodwin, E.H. DNA double-strand break repair proteins are required to cap the ends of mammalian chromosomes. *Proc. Natl. Acad. Sci.* 96, 14899-14904 (1999).

Bassing, C.H., Swat, W. Alt. F.W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell.* 109, S45–S55 (2002).

Blier, P.R. Griffith, A.J. Craft, J. Hardin, J.A. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J Biol. Chem.* 10, 7594-7601 (1993).

Bosma, G.C., Davisson, M.T., Ruetsch, N.R., Sweet, H.O., Shultz, L.D., Bosma, M.J. A severe combined immunodeficiency mutation in the mouse. *Nature.* 301, 527-530 (1983).

Boulton, S.J. Jackson, S.P. Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double-strand break rejoining and in telomeric maintenance. *Nucleic Acids Res.* 24, 4639-4648 (1996).

Bradford, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254 (1976).

Buck, D., Malivert, L., de Chasseval, R., Baraud, A., Fondaneche, M.C., Sanal, O., Plebani, A., Stephan, J.L., Hufnagel, M., Le Deist, F., Fischer, A., Durandy, A., de Villartay, J.P., Revy, P. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell.* 124, 287-299 (2006).

Burma, S. Chen, B.P.C. Chen, D.J. Role of non-homologous end-joining (NHEJ) in maintaining genomic integrity. *DNA Rep.* 5, 1042-1048 (2006).

Busch, S.J., P.Sassone-Corsi. Dimers, leucine zippers and DNA-binding domains. *Trends Genet.* 6, 36-40 (1990).

Cabrita, L.D., Dai, W., Bottomley, S.P. A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production. *BMC Biotechnol.* 6, 12 (2006).

Callebaut, I. Malivert, L. Fischer, A. Mornon, J.P. Revy, P. de Villartay, J.P. Cernunnos interacts with the Xrcc4/DNA-Ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. *J. Biol Chem.* 281, 3857-13860 (2006).

Carey, J. Gel retardation. *Methods Enzymol.* 208, 103–117 (1991).

Cantagrel, V., Lossi, A.M., Lisgo, S., Missirian, C., Borges, A., Philip, N., Fernandez, C., Cardoso, C., Figarella-Branger, D., Moncla, A., Lindsay, S., Dobyns, W.B., Villard, L. Truncation of NHEJ1 in a patient with polymicrogyria. *Hum Mutat.* 28, 356-364 (2007).

Cavero, S., Chahwan, C., Russell, P. Xfl1 is required for DNA repair by nonhomologous end joining in Schizosaccharomyces pombe. *Genetics* 175, 963-967 (2006).

Chappell, C., Hanakahi, L.A., Karimi-Busheri, F., Weinfeld, M., West, S.C. Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. *EMBO J.* 21, 2827-2832 (2002).

Chaudhuri, J., Alt, F.W. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nature Rev. Immunol.* 4, 541–552 (2004).

Chen, L., Trujillo, K., Ramos, W., Sung, P., Tomkinson, A.E. Promotion of Dnl4-catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Mol. Cell.* 8, 1105-1115 (2001).

Chen, L., Huang, S., Lee, L., Davalos, A., Schiestl, R.H., Campisi, J., Oshima, J. WRN, the protein deficient in Werner syndrome, plays a critical structural role in optimizing DNA repair. *Aging Cell.* 2, 191-199 (2003).

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A., Johnston, M. Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* 301, 71-76 (2003).

Collis, S.J., DeWeese, T.L., Jeggo, P.A., Parker, A.R. The life and death of DNA-PK. *Oncogene*. 24, 949-961 (2005).

CrystalClear: An integrated program for the collection and processing of area detector data. Rigaku Corporation (1997-2002).

Dai, Y., Kysela, B., Hanakahi, L.A., Manolis, K., Riballo, E., Stumm, M., Harveille, T.O., West, S.C., Oettinger, M.A., Jeggo, P.A. Nonhomologous end joining and V(D)J recombination require and additional factor. *Proc. Natl. Acad. Sci.* 100, 2462-2467 (2003).

Daley, J. M., Palmobs, P. L., Wu, D., Wilson, T.E. Nonhomologous end joining in yeast. *Annu Rev. Genet.* 39, 431-451 (2005a).

Daley, J. M., Laan, R. L., Suresh, A. and Wilson, T. E. DNA joint dependence of pol X family polymerase action in nonhomologous end-joining. *J Biol Chem.* 280, 29030-29037 (2005b).

De Lano, W.L. The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, USA (2002).

Deshpande, R.A., Wilson, T.E. Modes of interaction among yeast Nej1, Lif1 and Dnl4 proteins and comparison to human XLF, XRCC4 and Lig4. *DNA Rep.* (2007).

Dore, A.S., Furnham, N., Davies, O.R., Sibanda, B.L., Chirgadze, D.Y., Jackson, S.P., Pellegrini, L., Blundell, T.L. Structure of an Xrcc4-DNA ligase IV yeast ortholog complex reveals a novel BRCT interaction mode. *DNA Rep.* 7, 362-368 (2006).

Douglas, P., Shikha, G., Morrice, N., Meek, K., Lees-Miller, S.P. DNA-PK-dependent phosphorylation of Ku70/80 is not required for non-homologous end-joining. *DNA Rep.* 4, 1006-1018 (2005).

Downs, J.A. Lowndes, N.F. Jackson, S.P. A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* 408, 1001–1004 (2000).

Downs, J.A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S.J., Jackson, S.P., Côté, J. Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol. Cell.* 16, 979-990 (2004).

Drenckhahn, D., Jons, T., Schmitz, F. Production of polyclonal antibodies against proteins and peptides. *Methods Cell Biol.* 37, 7-56 (1993).

Dudas, A., Chovanec, M. DNA double-strand break repair by homologous recombination. *Mutat. Res. Fundam. Mol. Mech. Mut.* 566, 131-167 (2004).

Dvir, A., Peterson, S.R., Knuth, M.W., Lu, H., Dynan, W.S. Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II. *Proc. Natl. Acad. Sci.* 24, 11920-11924 (1992).

Elliott, B., Jasin, M. Double-strand breaks and translocations in cancer. *Cell Mol Life Sci.* 59, 373-375 (2002).

Enders, A., Fisch, P., Schwarz, K., Duffner, U., Pannicke, U., Nikolopoulos, E., Peters, A., Orlowska-Volk, M., Schindler, D., Friedrich, W., Selle, B., Niemeyer, C., Ehl, S. A severe form of human combined immunodeficiency due to mutations in DNA ligase IV. *J Immunol.* 176, 5060-5068 (2006).

Fisher, T.S. Zakian,V. Ku: a multifunctional protein involved in telomere maintenance. *DNA Rep.* 4, 1215-1226 (2005).

Foster, S.S., Zubko, M.K., Guillard, S., Lydall, D. MRX protects telomeric DNA at uncapped telomeres of budding yeast cdc13-1 mutants. *DNA Rep.* 5, 840-851. (2006).

Frank-Valliant, M. and Marcand, S. NHEJ regulation by mating type is exercised through a novel protein, Lif2p, essential to the Ligase IV pathway. *Genes Dev.* 15, 3005-3012 (2001).

Frobisher, M. Fundamentals of Microbiology. W.B. Saunders Company. Pennsylvania. 1968.

Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T.E., Mann, M., Lieber, M.R. Activity of DNA Ligase IV stimulated by complex formation with Xrcc4 protein in mammalian cells. *Nature.* 388, 492-495.

Harlow, E., Lane, D. Immunoprecipitation: Lysing Yeast Cells Using Glass Beads. *Cold Spring Harbor Protocols.* 2006.

Hentges, P. Ahnesorg, P. Pitcher, R.S. Bruce, C.K. Kysela, B. Green, A.J. Bianchi, J. Wilson, T.E. Jackson, S.P. Doherty, A.J. Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos. *J. Biol. Chem.* 281, 37517-37526 (2006).

Howard, A.J., Poulos, T.L. Methods in macromolecular crystallography. *Adv. Biophys Chem* 2, 1-36 (1992).

Johnson, W.C.Jr. Secondary structure of proteins through circular dichroism spectroscopy. *Ann. Rev. Biophys. Chem.* 17 (1988) 145–166.

Junop, M.S., Modesti, M., Guarne, A., Ghirlando, R., Gellert, M., Yang, W. Crystal structure of the Xrcc4 DNA repair protein and implications for end-joining. *EMBO J.* 19, 5962-70 (2000).

Kamiguchi, Y., Tateno, H., Iizawa, Y., Mikamo, K. Chromosome analysis of human spermatozoa exposed to antineoplastic agents in vitro. *Mutat. Res.* 326, 185-192 (1995).

Karathanasis, E., Wilson T.E. Enhancement of *Saccharomyces cerevisiae* end-joining efficiency by cell growth stage but not by impairment of recombination. *Genetics.* 161, 1051-1027 (2002).

Karimi-Busheri, F., Daly, G., Robins, P., Canas, B., Pappin, D.J., Sqouros, J., Miller, G.G., Fakhari, H., Davis, E.M., Le Beau, M.M., Weinfeld, M. Molecular characterization of a human DNA kinase. *J. Biol. Chem.* 274, 24187-24194 (1999).

Keeney, S., Giroux, C.N., Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell.* 88, 375-384 (1997).

Kegel, A., Sjostrand, J.O.O., Astrom, S. U. Nej1, a cell type-specific regulator of nonhomologous end-joining in yeast. *Curr. Biol.* 11, 1611-1617 (2001).

Kim, J.S., Krasieva, T.B., Kurumizaka, H., Chen, D.J., Taylor, A.M.R., Yokomori, K. Independent and sequential recruitment of NHEJ and HR factors to DNA damage sites in mammalian cells. *J. Cell. Biol.* 170, 341-347 (2005).

Kneale, G. G. DNA-Protein Interactions: Principles and Protocols, Volume 30, Methods in Molecular Biology. Humana Press Inc. Totowa N.J. (1994).

Krogan, N.J., Lam, M.H., Fillingahm, J., Keogh, M.C., Gebbia, M, Li, J., Datta, N., Cagney, G., Buratowski, S., Emili, A., Greenblatt, J.F. Proteasome involvement in the repair of DNA double-strand breaks. *Mol. Cell.* 16, 1027-1034 (2004).

Kuhne, M., Rothkamm, K., Lobrich, M. Physical and biological parameters affecting DNA double strand break misrejoining in mammalian cells. *Radiat. Prot. Dosimetry* 99, 129-32 (2002).

Kushnirov, V.V. Yeast functional analysis report: rapid and reliable protein extraction from yeast. *Yeast.* 2000, 16, 857-860.

Langland, G., Elliott, J.L.Y., Creaney, J., Dixon, K., Groden, J. The BLM helicase is necessary for normal DNA double-strand break repair. *Cancer Res.* 62, 2766-2770 (2002).

Leber, R. Wise, T.W. Mizuta, R. Meek, K. The Xrcc4 gene product is a target for and interacts with the DNA-dependent protein kinase. *J. Biol. Chem* 273, 1794-1801 (1998).

Lee, S.E., Moore, A., Holmes, J.K., Umezu, K., Kolodner, R.D., Haber J.E. *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell.* 94, 399-409 (1998).

Lee, Y., McKinnon, P. J. DNA Ligase IV suppresses medulloblastoma formation. *Cancer Res.* 62, 6395–6399 (2002).

Lewis, S.M. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.* 56, 27–150 (1994).

Lindahl, T. Instability and decay of the primary structure of DNA. *Nature.* 362, 709 (1993).

Liti, G. and Louis, E.J. NEJ1 prevents NHEJ-dependent telomere fusions in yeast without telomerase. *Mol. Cell.* 11, 1373-1378 (2003).

Lobrich, M., Jeggo, P. A. Harmonising the response to DSBs: a new string in the ATM bow. *DNA Rep.* 4, 749-759 (2005).

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C.A., Krieger, M., Scott, M.P., Zipursky, S.L., Darnell, J. Molecular Biology of the Cell, 5th ed. W.H. Freeman. New York (2004).

Lu, H., Pannicke, U., Schwarz, K., Lieber, M.R. Length-dependent binding of human XLF to DNA and stimulation of Xrcc4:DNA Ligase IV activity. *J. Biol. Chem.* 282, 11155-11162 (2007).

Ma, Y., Pannicke, U., Schwarz, K., Lieber, M.R. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell.* 108, 781-794 (2002).

Ma, Y., Ku. H., Tippin, B., Goodman, M.F., Shimazaki, N., Koiwai, O., Hsieh, C.L., Schwarz, K., Lieber, M.R. A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol. Cell.* 16, 701-713 (2004).

Ma, Y., Lu, H., Schwarz, K., Lieber, M.R. Repair of double-strand DNA breaks by the human nonhomologous DNA end joining pathway: the iterative processing model. *Cell Cycle.* 4, 1193-1200 (2005a).

Ma, Y., Schwarz, K., Lieber, M.R. The Artemis: DNA-PKcs endonuclease cleaves DNA loops, flaps and gaps. *DNA rep.* 4, 845-51 (2005b).

Mahajan, K.N. McElhinny, S.A.N., Mitchell, B.S., 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.* 22, 5194-5202 (2002).

McPherson, A. Crystallization of proteins by variation of pH or temperature. *Methods Enzymol.* 114, 125-127 (1985).

McPherson, A., Malkin, A.J., Kuznetsov, Y.G. The science of macromolecular crystallization. *Structure* 3, 759-768 (1995).

Meesala, S. Structural characterization of the C-terminal domain of human DNA Ligase IV bound to Xrcc4. Master of Science Thesis (2006).

Michel, B., Ehrlich, S.D. and Uzest, M. DNA double-strand breaks caused by replication arrest. *EMBO J.* 16, 430-438. (1997).

Mizuta, R. Cheng, H.L. Gao, Y. Alt, F.W. Molecular genetic characterization of Xrcc4 function. *Int. Immunol.* 9, 1607-1613 (1997).

Modesti, M. Hesse, J.H. 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.* 7, 2008-2018 (1999).

Modesti, M., Junop, M.S., Ghirlando, R., van de Ra, M., Gellert, M., Yang, W., Kanaar R. Tetramerization and DNA ligase IV interaction of the DNA double-strand break repair protein XRCC4 are mutually exclusive. *J Mol Biol.* 334, 215-228 (2003).

Moore, J.K., Haber, J.E. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 16, 2164-2173 (1996).

Morgan, W.F., Day, J.P., Kaplan, M.I., McGhee, E.M., Limoli, C.L. Genomic instability induced by ionizing radiation. *Radiat. Res.* 146, 247-258 (1996).

Moshous, D., Li, L., Chasseval, R., Philippe, N., Jabado, N., Cowan, M.J., Fischer, A., Villartay, J.P. A new gene involved in DNA double strand break repair and V(D)J recombination is located on human chromosome 10p. *Hum. Mol. Genet.* 9, 583-588 (2000).

Nonoyama, S., Ochs, HD. Immune deficiency in SCID mice. *Int. Rev. Immunol.* 13, 289-300 (1996).

Ooi, S.L., Shoemaker, D.D., Boeke, J.D. A DNA microarray-based genetic screen for nonhomologous end-joining mutants in *Saccharomyces cerevisiae*, *Science* 294, 2552-2556 (2001).

Palmobos, P.L., Daley, J.M., and Wilson, T.E. Mutations in the Yku80 C-terminus and Xrs2 FHA domain specifically block yeast nonhomologous end-joining. *Mol and Cell Biol* 25, 10782-10790 (2005).

Rhodes, G. Crystallography Made Crystal Clear. Academic Press. San Diego (1993).

Richardson, C., Jasin, M. Frequent chromosomal translocations induced by DNA double strand breaks. *Nature.* 405, 697-700 (2000).

Rogakou, E.P., Boon, C., Redon, C., Bonner, W.M. Megabase chromatin domains involved in DNA double-strand breaks *in vivo*. *J. Cell Biol.* 146, 905-916 (1999).

Rooney, S., Sekiguchi, J., Whitlow, S., Eckersdorff, M., Manis, J. P., Lee, C., Ferguson, D. O., Alt, F. W. Artemis and p53 cooperate to suppress oncogenic *N-myc* amplification in progenitor B cells. *Proc. Natl. Acad. Sci.* 101, 2410–2415 (2004).

Rothkamm, K., Kuhne, M., Jeggo, P.A., Lobrich M. Radiation-induced genomic rearrangements formed by nonhomologous end-joining of DNA double-strand breaks. *Cancer Res.* 61, 3886-3993 (2001).

Rothkamm, K., Lobrich, M. Misrepair of radiation-induced DNA double-strand breaks and its relevance for tumorigenesis and cancer treatment. *Int J Oncol.* 2, 433-440 (2002).

Schiestl, R,H,, Wintersberger, U. DNA damage induced mating type switching in Saccharomyces cerevisiae. *Mutat Res.* 284, 111-23 (1992).

Sibanda, B., Critchlow, S., Begun, Jake., Pei, X., Jackson, S.P., Blundell, T., Pellegrini, L. Crystal structure of an Xrcc4-DNA ligase IV complex. *Nature Struct. Biol.* 8, 1015-1019 (2001).

Spagnolo. L., Rivera-Calzada, A., Pearl, L.H., Llorca, O. Three dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. *Mol Cell.* 22, 511-519 (2006).

Taccioli, G.E., Amatucci, A.G., Beamish, H.J., Gell, D., Torrez Arzayus, M.I., Priestley, A., Jackson, S.P., Rothstein, A.M., Jeggo, P.A., Herrera, V.L.M., Xiang, X.H. Targeted disruption of the catalytic subunit of the DNA-PK gene in mice confers severe combined immunodeficiency and radiosensitivity. *Immunity.* 9 335-366 (1998).

Tan, S., Hunziker Y., Pellegrini, L., Richmond T.J. Crystallization of the yeast Matα2/MCM1/DNA ternary complex: general methods and principles for Protein/DNA crystallization. *J. Mol. Biol.* 297, 947-959 (2000).

Teo, S.H. Jackson, S.P. Lif1p targets the DNA ligase Lig4p to sites of DNA double-strand breaks. *Curr. Biol.* 10, 165-168 (2000).

Tsai, C.J., Kim, S.A., Chu, G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proc. Natl. Acad. Sci.* 104, 7851-7856 (2007).

Tseng, H.M., Tomkinson, A.E. Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1. *J. Biol. Chem.* 279, 47580-47588 (2004).

Uematsu, N., Weterings, E., Yano, K., Morotomi-Yano, K., Jakob, B., Taucher-Scholz G., Mari, P.O., van Gent, D.C., Chen, B.P.C., Chen, D.J. Autophosphorylation of DNA-PKcs requlates its dynamics at DNA double-strand breaks. *J. Cell. Biol.* 177, 219-229 (2007).

Valencia, M., Bentele, M., Vaze, M.B., Herrmann, G., Kraus, E., Lee, S.E., Schär, P., Haber, J.E. NEJ1 controls non-homologous end-joining in Saccharomyces cerecisiae. *Nature.* 414, 666-669 (2001).

Varga, T., Aplan, P.D. Chromosomal aberrations induced by double strand DNA breaks. *DNA Rep.* 4, 1038-1046 (2005).

Walker, J.R., Corpina, R.A., Goldberg, J. Structure of Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature.* 412, 607-614 (2001).

Wang, H., Perrault, A.R., Takeda, Y., Qin, W., Wang, H., Iliakis, G. Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res.* 31, 5377-5388 (2003).

Watanabe, S., Kamiguchi, Y. Clastogenic effects of mitomycin C (MMC) on human sperm chromosomes. *Chromosome Sci.* 3, 157 (1999).

Willers, H., Husson, J., Lee, L. W., Hubbe, P., Gazemeier, F., Powell, S.N., Dahm-Daphi, J. Distinct mechanisms of nonhomologous end joining in the repair of site-directed chromosomal breaks with noncomplementary and complementary ends. *Radiation Research.* 166, 567-574 (2006).

Wilson, T.E., Lieber, M.R. Efficient processing of DNA ends during yeast non-homologous end joining. Evidence for a DNA polymerase beta (Pol4)-dependent pathway. *J. Biol. Chem.* 274, 23599-23609 (1999).

Wu, X., Wilson. T.E. A role for FEN-1 in non-homologous DNA end-joining: the order of strand annealing and nucleolytic processing events. *Proc. Natl. Acad. Sci.* 96, 556-566 (1999).

Wyman, C., Kanaar, R. DNA double-strand break repair: All's well that ends well. *Annu. Rev. Genet.* 400, 363-383 (2006).

Yan, C.T., Kaushal, D., Murphy, M., Zhang, Y., Datta, A., Chen, C., Monroe, B., Mostoslavsky, G., Coakley, K., Gao, Y., Mills, K.D., Fazeli, A.P., Tepsuporn, S., Hall, G., Mulligan, R., Fox, E., Bronson, R., De Girolami, U., Lee, C., Alt, F.W. XRCC4 suppresses medulloblastomas with recurrent translocations in p53-deficient mice. *Proc. Natl. Acad. Sci.*103, 7378-7383 (2006).

Zha, S., Alt, F.W., Cheng, H.L., Brush, J.W., Li, G. Defective DNA repair and increased genomic instability in Cernunnos-XLF-deficient murine ES cells. *Proc. Natl. Acad. Sci.* 104, 4518-4523 (2007).