OVERCOMING OBSTACLES TO NONHOMOLOGOUS END JOINING REPAIR OF CHROMOSOME DOUBLE STRAND BREAKS

Steven Andrew Roberts

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill 2008

Approved by:
Dale A. Ramsden
Jack D. Griffith
Jeff J. Sekelsky
Nikolay V. Dokholyan
Yi Zhang
ABSTRACT

STEVEN ANDREW ROBERTS: Overcoming Obstacles to Nonhomologous End Joining Repair of Chromosome Double Strand Breaks
(under the direction of Dale Ramsden)

Protein occlusions and oxidative nucleotide damages that flank chromosome double strand breaks are serious obstacles to nonhomologous end joining (NHEJ) since they each directly impact a specific step of this repair pathway. Because NHEJ requires the Ku heterodimer (a ring protein requiring 360° of access for binding) to recognize DNA ends, this pathway would seem particularly sensitive to DNA end-occlusions. End-associated nucleotide damages would also seemly present a significant obstacle to NHEJ as they specifically inhibit XRCC4-LigaseIV, and their proximity to DNA ends precludes their removal by traditional base excision repair enzymes.

Surprisingly, NHEJ is able to overcome these obstacles. I have shown that despite the potential steric conflicts, Ku readily binds protein-occluded DNA ends. Moreover, this very protein, predicted to be inhibited by end-associated protein obstructions, may specifically limit NHEJ’s need for ATP-dependent chromatin remodeling as Ku readily evicts weakly bound DNA-associated proteins and peels DNA from strongly positioned nucleosomes.

In the case of end-associated nucleotide damages, NHEJ employs a large repertoire of processing factors, some displaying unique activity, to specifically remove, repair, or bypass these complex lesions. My work identifies a new part of this repertoire:
Ku is novel AP lyase tailored to remove end-associated abasic sites and thereby facilitate joining of these complex breaks. Ku efficiently cleaves the phosphodiester backbone 3’ of end-associated abasic sites, precisely removing damage in a context unfavorable to other known lyases.

Finally, my work suggests NHEJ ensures difficult breaks containing incompatible ends are repaired. This involves complementing the above processing based ligation with an additional, low fidelity ligation mechanism. XLF/Cernunnos, the key factor establishing this subpathway, functions to clamp XRCC4-LigaseIV to DNA ends and thereby stimulates LigaseIV’s ability to ligate mismatched overhangs. Consequently, this activity allows NHEJ to join difficult DNA ends, potentially incorporating damaged nucleotides or mismatches into an intact chromosome, where they can be repaired without further risk of DNA translocation. Taken together, these unique activities highlight NHEJ’s flexibility and sophistication in overcoming obstacles to preserve genome integrity.
ACKNOWLEDGEMENTS

I would like to thank my research advisor, Dale, for his support during the last six years. He has taught me much, and I would not be the scientist or person I am today without his influence. I would also like to thank my dissertation committee, Nikolay Dokholyan, Jack Griffith, Jeff Sekelsky, and Yi Zhang for their time, suggestions, and support.

To Ramsden Lab members, current and past, thank you for your camaraderie. It has been a joy work in such a caring environment. A special thanks to Bryan and Martin, you could always make me laugh.

I am very grateful to my family for their love and guidance, particularly, my wife, Sarah, whose patience never ends. You inspire me everyday.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS AND SYMBOLS</td>
<td>xiii</td>
</tr>
</tbody>
</table>

## CHAPTER

1. INTRODUCTION .................................1
   
   Nonhomologous End Joining ...........................1
   
   Obstacles to NHEJ repair of chromosome double strand breaks ........2

2. LOADING OF THE NONHOMOLOGOUS END JOINING FACTOR, KU, ON PROTEIN-OCLUDED DNA ENDS ..................15
   
   Introduction ........................................................................15
NaBH₄ cross-linking

Results

NHEJ removes abasic sites before end ligation

Ku is an AP lyase

Ku is the dominant AP lyase in cell extracts

Discussion

Acknowledgements

4. XLF STIMULATES LOW FIDELITY LIGATION BY NONHOMOLOGOUS END JOINING

Introduction

Methods

Protein Purifications

DNA substrates

Electrophoretic Mobility Shift Assay

Ligation assay

Synapsis assay

Results

XLF clamps XRCC4-LigaseIV to DNA ends

XLF stimulates ligation under stringent conditions

XLF stabilizes the synaptic complex

XLF stimulates mismatch ligation

Discussion

Acknowledgements

5. DISCUSSION
LIST OF TABLES

Table 4.1. Sequence analysis of pol μ and XLF mediated junctions........................71
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1.</td>
<td>Steps of Nonhomologous End Joining</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1.2.</td>
<td>Oxidation of the C’4 carbon of nucleotides</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.3.</td>
<td>Generation of double strand breaks by ionizing radiation</td>
<td>12</td>
</tr>
<tr>
<td>Figure 1.4.</td>
<td>V(D)J Recombination</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.5.</td>
<td>Obstacles to repair by NHEJ</td>
<td>14</td>
</tr>
<tr>
<td>Figure 2.1.</td>
<td>Ku’s DNA binding activity at H1 bound DNA ends</td>
<td>31</td>
</tr>
<tr>
<td>Figure 2.2.</td>
<td>Ku’s DNA binding activity at nucleosome-associated ends</td>
<td>32</td>
</tr>
<tr>
<td>Figure 2.3.</td>
<td>Assessment of Ku’s ability to thread on nucleosome-associated DNA Ends</td>
<td>33</td>
</tr>
<tr>
<td>Figure 2.4.</td>
<td>Affinity of Ku for nucleosome-associated DNA ends</td>
<td>34</td>
</tr>
<tr>
<td>Figure 2.5.</td>
<td>Analysis of the protein composition of Ku-shifted species</td>
<td>35</td>
</tr>
<tr>
<td>Figure 2.6.</td>
<td>Hydroxyl radical footprinting of nucleosomes in the presence of Ku</td>
<td>36</td>
</tr>
<tr>
<td>Figure 2.7.</td>
<td>Ku’s DNA binding activity at RAG bound signal ends</td>
<td>37</td>
</tr>
<tr>
<td>Figure 2.8.</td>
<td>Model of Ku bound to nucleosome-associated DNA ends</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3.1.</td>
<td>NHEJ of ends with abasic sites</td>
<td>52</td>
</tr>
<tr>
<td>Figure 3.2.</td>
<td>Characterization of in vitro NHEJ junctions</td>
<td>53</td>
</tr>
<tr>
<td>Figure 3.3.</td>
<td>AP lyase activity of purified Ku</td>
<td>54</td>
</tr>
<tr>
<td>Figure 3.4.</td>
<td>Specific lyase activities of pol β, pol λ, and Ku</td>
<td>55</td>
</tr>
<tr>
<td>Figure 3.5.</td>
<td>Characterization of HeLa and CHO cell extracts</td>
<td>56</td>
</tr>
<tr>
<td>Figure 3.6.</td>
<td>AP lyase activity of cell extracts</td>
<td>57</td>
</tr>
<tr>
<td>Figure 3.7.</td>
<td>NaBH₄ cross-linking of cell extracts to Bleocin damaged DNA</td>
<td>58</td>
</tr>
<tr>
<td>Figure 4.1.</td>
<td>Interaction between XLF and other core NHEJ factors</td>
<td>72</td>
</tr>
</tbody>
</table>
Figure 4.2. In vitro NHEJ of complementary substrates ........................................73
Figure 4.3. Synaptic Maintenance ........................................................................74
Figure 4.4. Mismatch Ligation by NHEJ ...............................................................75
Figure 4.5. Model of XLF-mediated low fidelity ligation ....................................76
Figure 5.1. Removal of protein obstructions at DSBs in cells..............................86
Figure 5.2. Complexity of NHEJ ........................................................................87
LIST OF ABBREVIATIONS AND SYMBOLS

BSA, bovine serum albumin
DNA-PKcs, DNA dependent protein kinase catalytic subunit
DSB, double strand break
EMSA, electrophoretic mobility shift assay
HMG1, high mobility group protein 1
$K_d$, dissociation constant
MBP, maltose binding protein
NHEJ, nonhomologous end joining
PDB, Protein Data Bank
Pol, polymerase
RAG, recombination activation gene
RS, recombination signal
SA, streptavidin
SEC, paired signal end complex
XRCC4, X-ray cross-complementary gene 4
$\beta$, beta
$\delta$, delta
$\lambda$, lambda
$\mu$, mu
$\mu$M, micromolar
$\mu$L, microliter
CHAPTER 1
INTRODUCTION

Chromosome double strand breaks present a serious threat to genome stability and often result in cell death if left unrepaired. Nonhomologous end joining (NHEJ) is the primary pathway in mammalian cells responsible for the repair of these dangerous lesions. This pathway is active throughout the cell cycle. However, NHEJ features most prominently during G1 and G0 cell cycle phases when low levels of nucleotide precursors and the lack of a sister chromatid template reduce the efficiency of homologous recombination [1]. In contrast to homologous recombination, nonhomologous end joining is generally perceived as a simple repair pathway. But as I discuss throughout this dissertation, NHEJ is confronted with a variety of obstacles that increase the difficulty of repair. These obstacles (often determined by the source and location of the break) are overcome by specific activities associated with the NHEJ machinery. The implementation and coordination of this repertoire of factors ultimately highlight the flexibility and sophistication of this repair pathway.

**Nonhomologous end joining**

NHEJ proceeds through a four step mechanism involving 1) end recognition, 2) end alignment (termed synapsis), 3) end processing, and 4) ligation (Figure 1.1). The Ku heterodimer is responsible for end recognition and serves as a scaffold for recruitment of
other end joining factors to the break. Ku and a 460 kDa protein kinase, DNA-PKcs, bind to each end of a DSB and align them (term end synapsis) making use of complementary overhanging sequence [2, 3]. After synapsis, XRCC4-LigaseIV is recruited to the break [4-6] along with an associated factor, XLF/Cernunnos [7-9]. The five NHEJ core factors then serve as a platform for a set of processing factors that cleave non-complementary flaps (Artemis nuclease), excise damaged bases, and fill in gaps (polymerases μ and λ) to produce compatible ends amenable to ligation by LigaseIV [10-17]. Ligation is the critical step within NHEJ. Consequently, genetic deficiencies that prevent the recruitment of XRCC4-LigaseIV or limit its ability to covalently repair breaks are particularly troublesome. As a result, deficiency in any of the core NHEJ factors (Ku, DNA-PKcs, XLF, and XRCC4-LigaseIV) results in immunodeficiency and increased sensitivity to DNA damaging agents like ionizing radiation or radiomimetic drugs (term end Radio Sensitive-Severe Combined Immune Deficiency or RS-SCID) [18-27]. NHEJ is therefore designed to ensure breaks are efficiently discovered and repaired. In the following chapters, I discuss mechanisms NHEJ employs to overcome a variety of context specific obstacles to repair, specifically the protein-occlusion of DNA ends and the presence of incompatible overhangs due to nucleotide damages and mismatches.

**Obstacles to NHEJ repair of chromosome double strand breaks**

The context of the chromosome double strand break is of utmost importance in determining the complexity of the repair process. DNA ends can be obscured by DNA binding proteins based upon the location of the break within the genome. Furthermore, the structure of the DNA ends themselves is highly variable, differing in sequence and
flanking nucleotide damages (Figure 1.5). Ultimately, the agent responsible for the break plays a pivotal role in these aspects of break context.

**Exogenous DSBs repaired by NHEJ**

**Ionizing radiation and radiometric drug induced breaks**

Ionizing radiation is a major source of chromosome double strand breaks repaired by NHEJ [28]. This DNA damaging agent and radiomimetic drugs (e.g. bleocin) generate breaks through a clustered series of oxidation events [29]. DNA can be oxidized at several positions within a nucleotide’s base or sugar [30]. Base oxidation results in the formation of modified nucleotide residues like 8-oxo-guanine, 5-hydroxyuracil, and thymine glycol, which base excision repair generally removes from the genome (reviewed in [31]). Alternatively, a nucleotide residue can be oxidized at the C’1, C’4, or C’5 sugar positions. Oxidation at the C’4 position (a major oxidation event seen with ionizing radiation and the radiometric drugs bleomycin and neocarzinostatin [32]) results in either direct formation of an abasic site or production of a single strand break with a 3’phosphoglycolate terminus [32, 33] (Figure 1.2). The clustering of these oxidation events leads to the formation of double strand breaks via two distinct mechanisms (Figure 1.3). First, close juxtaposition of 3’phosphoglycolate formation in opposing DNA strands directly results in a double strand break. Alternatively, clustering of damaged bases or abasic sites on opposing strands leads to DSB formation through aborted base excision repair. After removal of oxidized bases by specific glycosylases, AP endonuclease I cleaves 5’ of any abasic sites present in the DNA. If the abasic site is proximate to a strand break on the opposing strand, a DSB results. Because radiation induced double strand breaks result from clustered nucleotide oxidations, these breaks
contain flanking nucleotide damages (e.g. 8-oxoguanine, abasic sites, 3’phosphoglycolates, 5’ hydroxyls, and 3’phosphates) that represent an obstacle to NHEJ [29].

**Endogenous DSBs repaired by NHEJ**

In addition to exogenous sources (i.e. radiation), chromosome double strand breaks are also generated during programmed genome rearrangements such as meiotic recombination, V(D)J recombination, and class switch recombination. Each of these processes couples an enzymatic cleavage of the genomic DNA (in a site specific manner) to a specific double strand break repair pathway. This coordination allows controlled rearrangements that generate genomic diversity. V(D)J recombination and class switch recombination both require nonhomologous end joining for successful chromosome double strand break repair [34, 35].

**V(D)J Recombination**

V(D)J recombination (reviewed in [36]) is the process by which adaptive immune systems generate the diverse set of mature antibodies, B-cell receptors, and T-cell receptors needed to recognize and eliminate a changing array of pathogens. Immune diversity is accomplished by recombining the immunoglobulin locus to align one of 70 V genes with one of 25 D segments (for receptor heavy chains only) and one of 6 J coding segments to produce a single receptor gene. Recombination is initiated when the RAG1 and RAG2 proteins pair a 12 and 23 recombination signal sequence (12-RS and 23-RS) that flank V, D, and J immunoglobulin coding fragments (Figure 1.4). The RAGs cleave both DNA strands by a two step nucleophilic attack that generates hairpin coding ends and precisely excises an intervening chromosome fragment. After cleavage, the RAGs
remain stably bound to the paired signal sequences on the excised fragment (signal ends) [37-39]. This potentially aids in organization of the four newly generated DNA ends so that nonhomologous end joining ligates the proper DNA ends together (signal end to signal end and coding end to coding end). Once the RAGs are removed, NHEJ can simply ligate the blunt signal ends, however, coding end present a more significant obstacle. Opening of the coding end hairpins by the Artemis nuclease occurs at variable positions in the hairpin loop and frequently results in non-complementary DNA ends [40]. Furthermore, NHEJ purposely generates diversity at these DNA ends through the addition of random nucleotides by the lymphocyte specific NHEJ polymerase, Tdt (reviewed in [41]). Consequently, coding ends are often incompatible and may require either enzymatic processing (by nucleases and polymerases) or a low fidelity mismatch ligation event to overcome this obstacle.

**Class Switch Recombination**

Recombination of V, D, and J segments generates an exon of an antigen receptor that encodes the “variable domain” responsible for antigen recognition. This domain is in turn spliced to a series of constant domain exons that make up the rest of the receptor. In B cell IgH chains, the constant domain is changed during the immune response to give different “isotypes” and isotype specific effector functions. B cells initially use isotype IgD and switch to IgM by alternative splicing. Subsequent isotype class switching (e.g. to IgA, IgG) involves rearrangement of the locus (reviewed in [42]). This recombination event is initiated by the activity of AID, which deaminates cytosine residues to uracil in the switch regions of the immunoglobulin locus. Base excision repair removes these uracil bases, generating abasic sites that are recognized by AP endonuclease 1. This
enzyme cleaves the phosphodiester bond 5’ of the abasic site resulting in a strand break with a 5’ deoxyribose terminus. Thus, the double strand breaks generated during switch recombination mirror those generated by ionizing radiation in that the DNA ends are associated with neighboring base damage. Curiously, unlike both radiation induced breaks and the breaks generated during V(D)J recombination, deficiency in different core NHEJ factors results in different degrees of defect for class switch recombination. Where lack of XRCC4 or Ku generally results in an equivalent failure of V(D)J recombination and sensitization to ionizing radiation [21, 43-45], deficiency of Ku is more severe than deficiency in XRCC4 in the case of class switch recombination [46]. I discuss a possible explanation for this observation in Chapter 3.

**Recognition of breaks in Chromatin**

Within the nucleus, a cell’s genome is compacted into a series of hierarchal protein-DNA structures called chromatin. The base unit of chromatin is the nucleosome [41, 47, 48] which itself consists of 146 bp of DNA wrapped around a core of histone proteins. This histone core consists of two copies each of histones H2A, H2B, H3, and H4 [47]. Histones are positively charged and therefore bind strongly to the negatively charged DNA backbone [47]. This tight interaction between histones and DNA limits the ability of transcription factors, replication factors, or factors involved in DNA repair to access the cellular DNA. Consequently, chromatin structure inhibits these processes [49-51]. Further decreasing DNA access, genomic DNA is also compacted by additional chromatin associated proteins (i.e. linker histones) into higher order structures [48].

Chromatin structure is already known to inhibit certain DNA repair processes. Specifically, in vitro reconstituted base excision repair and nucleotide excision repair
pathways display a 3 to 10 fold decrease in the efficiency of repair of nucleosome-associated lesions [52-55]. Similarly, repair of UV lesions in yeast occurs less efficiently when the lesion occurs in the context of heterochromatic telomere regions as opposed to euchromatin [56]. This indicates that in a living nucleus chromatin structure impedes DNA repair. In the case of DNA double strand break repair, the location of chromosome double strand breaks is impacted by chromatin structure itself. Oxidative damage occurs most frequently in areas of open chromatin that lack associated protein or at solvent exposed nucleotide residues [57]. However, nucleosomes are only spaced by an average of 20 to 100 bp of protein-free DNA [58]. Therefore, DSBs likely occur within or near nucleosomes, and proteins likely obscure break sites. Obscured breaks could severely inhibit NHEJ as this repair pathway requires Ku’s DNA end recognition activity. Ku’s ring structure requires 360 degrees of access to the end for binding [59], and thus steric hindrance could block Ku’s ability to bind protein-associated ends. In chapter 2, I discuss how chromatin architecture impacts Ku’s ability to recognize DNA ends. I further propose a model suggesting Ku’s own structure is tailored for finding ends in protein-DNA complexes and potentially aides in generation of end access for the rest of the NHEJ machinery.

**Processing DNA end associated nucleotide damage**

As discussed earlier, the causative agent of a chromosome double strand break often determines the structure of a DNA end. The clustering of oxidative base and sugar damage carries an additional consequence beyond formation of double strand breaks. While two opposing damages are required to produce a DSB, the potential to oxidize nucleotide residues within the same strand makes it likely that the breaks formed are
complex and contain damaged nucleotides near the DNA ends [29]. These damages decrease the ability of XRCC4-LigaseIV to join DNA ends and thereby contribute significantly to the cytotoxicity of radiation induced breaks. As many forms of end-associated nucleotide damage exist, NHEJ employs a similarly large collection of processing factors to excise, repair, and bypass these damages. NHEJ processing factors have already been implicated in removal of 3’ phosphoglycolates [60, 61], repair of 5’hydroxyls and 3’phosphates [62], and bypass of 8-oxoguanine [63]. In chapter 3, I discuss a novel activity of the NHEJ core factor, Ku, and how it enables NHEJ to circumvent end-associated abasics sites, a flanking nucleotide damage common to breaks generated by ionizing radiation, aborted base excision repair, and during heavy chain class switch recombination.

**Low Fidelity Ligation**

Ligation is the pivotal step during nonhomologous end joining. As a result, complex double strand breaks and their inhibition of XRCC4-LigaseIV represent a significant hurdle for this repair pathway. As discussed above, NHEJ is well equipped to handle such breaks through the joint activity of a variety of processing factors that can remove damages, restore or remove phosphates, and fill gaps. Utilization of these processing factors generally limits the loss of DNA sequence. However, end processing requires re-organization of the synaptic complex as XRCC4-LigaseIV must release the aligned DNA ends to allow access to processing factors. This re-organization increases the chance for synaptic complex dissociation and consequently the potential for DNA translocations or chromosome loss [64]. Recent work describing the newly identified NHEJ factor, XLF/Cernunnos suggests that this protein may facilitate a low fidelity
NHEJ pathway [65, 66] that supersedes processing based ligation. This XLF-mediated pathway may allow ligation of difficult DNA ends without end processing and thereby reduce the chances for translocation or deletion.

XLF is a 33 kDa protein that associates with XRCC4-LigaseIV and DNA [7, 26]. Human patients deficient in XLF display the RS-SCID phenotype characteristic of deficiency in core end joining factors [26]. Presence of XLF in in vitro ligation reactions, however, fails to significantly stimulate ligation of complementary ends while ligation of mismatched overhangs is dramatically increased [65-67]. Thus, XLF may function less broadly than its phenotype suggests, but instead with a primary role in the mismatch ligation of incompatible DNA ends. In chapter 4, I present a mechanism by which XLF stimulates ligation of difficult DNA ends. I also propose a model where NHEJ balances end processing with XLF dependent low fidelity ligation to maximize the retention of DNA sequence while minimizing the risk of DNA translocation and deletion.
Figure 1.1: Steps of Nonhomologous End Joining. Upon generation of a double strand break, the Ku heterodimer binds the resulting DNA ends. Ku then recruits DNA-PKcs and this complex aligns the DNA ends in a process termed synapsis. XRCC4-LigaseIV, XLF, and associated processing factors then arrive at the break to restore the complementarity of the two overhangs. Once this occurs, XRCC4-LigaseIV ligates the ends.
Figure 1.2: Oxidation of the C'4 carbon of nucleotides. Ionizing radiation and radiomimetic drugs commonly oxidize the C'4 carbon of nucleotide residues. Two potential outcomes of this oxidation event exist, either 1) formation of an abasic site or 2) strand breakage and conversion of the damaged sugar to a 3'phosphoglycolate.
Figure 1.3: Generation of double strand breaks by ionizing radiation. Ionizing radiation creates clustered nucleotide damage that produces double strand breaks by two mechanisms involving C4 oxidation. 1) Double strand breaks can be generated directly by production of opposing strand breaks associated with 3'phosphoglycolate formation. 2) APE1 incises 5' of abasic sites opposite a strand break. Either mechanism generates DNA ends with flanking nucleotide damage.
Figure 1.4: V(D)J recombination. Antigen receptor diversity is accomplished by programmed rearrangement of the immunoglobulin locus. These rearrangements are initiated when RAG1/2 cleave precisely at paired 12 RS (black triangle) 23 RS (white triangle) sequences. The cleavage event generates hairpin coding ends (grey rectangles) and blunt signal ends. The RAGs remain bound to the excised signal fragment, occluding the DNA ends. Hairpin opening often generates mismatched overhangs that must be joined by NHEJ.
Figure 1.5: Obstacles to repair by NHEJ. Double strand breaks are often associated with 1) protein-occlusions (e.g. nucleosomes or RAGs), 2) flanking nucleotide damages (e.g. abasic sites or 3’phosphoglycolates), and 3) non-complementary ends.
CHAPTER 2
LOADING OF THE NONHOMOLOGOUS END JOINING FACTOR, KU, ON PROTEIN-OCCCLUDED DNA ENDS

INTRODUCTION

DNA double strand breaks (DSBs) pose a threat to genomic integrity. Nonhomologous end joining (NHEJ) is a major pathway for repair of exogenously introduced DSBs in mammals and is the only efficient way to repair DSB intermediates in V(D)J recombination [68]. NHEJ requires the DNA binding heterodimer, Ku, to first recognize broken DNA ends [69-72] and subsequently recruit the additional NHEJ factors necessary to complete repair (e.g. DNA-PKcs [73], the XRCC4-LigaseIV complex [5, 6], and polymerases [12, 74]). Consequently, deficiency in Ku cripples NHEJ and leads to severe immunodeficiency [75, 76] and cellular sensitivity to agents that cause DSBs (e.g. ionizing radiation) [25, 77]. Ku’s ability to recognize DNA ends requires the end to be inserted through a channel in its structure [59, 72, 78]. This manner of binding makes Ku highly specific for DNA ends but requires that the ends must be accessible through 360 degrees. In cells, DNA is generally coated with proteins (e.g. histones and other chromatin associated proteins) that might be expected to block Ku’s ability to load on DNA ends and thus impair NHEJ.
Here we show Ku displays a variety of responses to protein obstructions at DNA ends. Ku is unable to bind a class of RAG (recombination activating gene) protein-bound DNA ends generated during V(D)J recombination. In contrast, Ku can displace certain proteins (e.g. histone H1) from DNA, and retains the ability to load on nucleosome-associated ends by peeling up to 50 bp of DNA away from the nucleosome surface. We therefore suggest Ku may be specifically suited to loading on protein-obstructed DNA ends, ultimately facilitating the recruitment of the NHEJ machinery without requiring extensive remodeling of chromatin structure at broken ends.

METHODS

Purified proteins: Bulk histones were purified from 5 liters of HeLa cells as described previously [79]. Briefly, a nuclear pellet was isolated, homogenized, and genomic DNA sheared by sonication. The clarified supernatant was then applied to a hydroxyapatite column equilibrated in 0.6 M NaCl. The column was washed extensively with the equilibration buffer to deplete linker histones and non-histone proteins, and fractions primarily containing core histones were eluted with 2 M NaCl. This eluent was then applied to a S200 gel filtration column equilibrated in 2 M NaCl to further deplete linker histones and ensure the proper stoichiometry of the core histone octamer. Histone H1 was a gift of Dr. Yi Zhang. Recombinant core RAG1 and RAG2 maltose binding protein (MBP) fusions and recombinant Ku was obtained as previously described in [80] and [6], respectively.

DNA substrates: DNA duplexes used in the histone H1 and RAG electrophoretic mobility shift assays (EMSAs) (Figures 2.1 and 2.7) were generated by T4
Polynucleotide Kinase labeling an oligonucleotide with $^{32}$P-$\gamma$-ATP and annealing it to a complementary oligonucleotide. The 23 recombination signal (RS) duplex, 12-RS duplex, and histone H1 EMSA substrate consist of the following paired oligonucleotides respectively:

5’CACAGTGAGTACTTCCACTGTCTGGCTGTAACAAAAACCTCGGGACG and biotin-tetraethyleneglycol-

5’CGTCCCGAGGGTTTTTGTACAGCCAGACAGTGAGTACTACCACTGTG,

5’CACAGTGCTACAGACTGGGAACAAAACCTCGAGACG and biotin-tetraethyleneglycol-5’CGTCTGCAGGGTTTTTGTTCCAGTCTGTAGCAGTGTG, and finally 5’ATGGAAATTGTGAGCGGGGATAACAATTCATG and

5’CATTGAATTGTATCCGCTCACAATTTCAT. The 601.2 nucleosome positioning sequence [81, 82] was obtained from Dr. Jonathon Widom and inserted into the EcoR1 and Nhe1 sites of the litmus 38 (New England Biolabs) multiple cloning region. This plasmid served as a template for PCR with the primers

5’CTGCAGAAGCTTGGTCCCG and 5’ACAGGATGTATATATCTGACG to generate unlabeled DNA for core nucleosomes (Figure 2.2), or with the primers biotin-tetraethyleneglycol-5’GATATCTGGATCCACGAATTC and

5’ACAGGATGTATATATCTGACG in the presence of $^{32}$P $\alpha$-dCTP to generate the radiolabeled DNA for nucleosomes containing a 40 bp linker (Figures 2.3b, 2.4, 2.5a (without $^{32}$P $\alpha$-dCTP), and 2.6d). Nucleosomes used in footprinting experiments (Figures 2.6a-c) were labeled by substituting the biotinylated primer described above with the fluorescent biotinylated primer biotin-decaethyleneglycol-bodipy 630-5’GATATCTGGATCCACGAATTC (Integrated DNA Technologies).
**Nucleosomes:** As described previously [83], nucleosomes were reconstituted on positioning sequences by salt dialysis using a DNA to histone ratio determined empirically for each DNA preparation such that reconstitutions contained less than 5% free DNA. The quality of reconstitutions was monitored by EMSA. Nucleosomes used for footprinting were also purified by preparative electrophoresis using a Mini-prep cell (Biorad) [84]. Native mononucleosomes were obtained by resuspending HeLa nuclei in 10 mL of micrococcal nuclease digestion buffer (10 mM Tris pH 7.5, 5 mM MgCl₂, 5 mM CaCl₂, 0.1 mM PMSF, and 0.5 mM DTT) and digestion with 0.02 U/μL of micrococcal nuclease (USB) at 37°C for 8 minutes [79]. After addition of NaCl to 0.6 M, the suspension was dounced to extract soluble chromatin and subsequently pelleted. The resulting supernatant was separated on a 5% to 30% sucrose gradient. As mononucleosomes derived from this digestion are rich in linker histones [79], we excluded mononucleosome containing fractions and further digested pooled oligonucleosome containing fractions with 1.5 U/μL micrococcal nuclease at 0°C for 40 minutes to generate core mononucleosomes. Mononucleosomes were then further purified on a second sucrose gradient.

**EMSA:** DNA-protein complexes were assembled prior to EMSA by incubation of the DNA and proteins in a standard reaction buffer (10 mM TRIS pH 7.5, 145 mM NaCl, 0.5 mg/mL bovine serum albumin (BSA), 1 mM DTT, 0.1 mM EDTA), except for experiments using the signal end complex (SEC), which used the following buffer: 25 mM MOPS pH 7, 120 mM KCl, 5 mM CaCl₂, 0.1 mg/mL BSA, 2 mM DTT, and 0.125 ng/μL of supercoiled DNA plasmid. The resulting complexes were resolved by electrophoresis at 300 V for 1 hour through a 16 cm, 1/3x TBE, 3.5 % polyacrylamide gel
containing 100 μg/mL BSA. EMSAs containing antibody supershifts used the following antibodies: α-Ku (Ab-3, Neomarkers), α-MBP (New England Biolabs), and α-histone H1 (clone B419, Biomed) purified by batch adsorption with hydroxyapatite beads.

In Figure 2.4, the apparent dissociation constant (Kd) was calculated from quantified (using ImageQuant (Molecular Dynamics)) EMSA experiments according to the equation,

\[ K_d = \frac{[Ku_{\text{free}}][\text{DNA}_{\text{free}}]}{[Ku:DNA]} \]

When 50% of the substrate is bound, [Ku:DNA] = [DNA_{\text{free}}], thus Kd=[Ku_{\text{free}}]. Since the total amount of Ku was kept in large excess over total DNA (> 40 fold for experiments in Figure 2.4) and [Ku_{\text{free}}]=[Ku_{\text{total}}]-[Ku:DNA], [Ku_{\text{free}}] was further approximated as [Ku_{\text{total}}] (also see [85]). The fraction shifted species (determined from EMSAs) was plotted against the log of Ku’s concentration. The resulting binding curves were best fit with a sigmoidal dose-response (variable slope) regression line (GraphPad Prism version 4.03 (Trial) for Windows, GraphPad Software, San Diego California USA, www.graphpad.com) to determine apparent Kd's and associated error.

Analysis of the protein composition of Ku-nucleosome complexes: Large scale 50 μL EMSA reactions were generated by incubating 80 nM nucleosome or free DNA with either 60 nM, 120 nM, or 240 nM Ku in our standard EMSA buffer at 25°C for 10 minutes. Reactions were then incubated for 5 minutes at 37°C before separating the formed complexes as described above. Ku-nucleosome and Ku-DNA complexes were visualized by ethidium bromide staining and subsequently excised. Complexes were then electro-eluted from the gel and their protein components concentrated by trichloroacetic acid precipitation. The relative amounts of Ku and histone H3 in each of the excised
complexes was determined by semi-quantitative western analysis probing with a polyclonal rabbit antibody raised against native, recombinant human Ku and a polyclonal antibody against histone H3 (Ab1791, Abcam) using fluorescent detection and a Typhoon imager (GE biosciences). Western blots were quantified using ImageQuantTL (Molecular Dynamics) and the ratio of Ku70 (Ku 80 is overexposed) to H3 was determined for nucleosome containing complexes (Figure 2.5b, species IV, V, and VI).

**Hydroxyl Radical Footprinting:** Footprinting experiments utilize an asymmetrically positioned nucleosome substrate containing a single 40 bp linker DNA [86]. The DNA end of this linker was labeled with a bodipy 630 fluorophore (Integrated DNA technologies) and blocked with a biotin-streptavidin complex (Figure 2.6a). Footprinting reactions were conducted in 15 μL of 30 mM TRIS pH 8.0 and 0.1 mM EDTA by incubating 80 nM nucleosome with 120 nM or 240 nM Ku at 25°C for 10 minutes. Reactions were then incubated at 37°C for 5 min before being placed on ice. Nucleosomes were twice treated at 4°C with 3.5 μL 3% hydrogen peroxide, 3.5 μL of a 2 mM ammonium iron(II) sulfate and 4 mM EDTA mixture, and 3.5 μL 20 mM sodium ascorbate at 5 minute intervals before reactions were stopped by addition of 5 μL 400 mM thiourea [86]. Reactions were acidified by addition of 5 μL 3 M sodium acetate and nicked DNA was purified through a minElute reaction clean-up column (Qiagen). Eluted DNA was electrophoresed on an 8% Urea-PAGE sequencing gel at 1800 volts and 40 watts for either 1.5 (Figure 2.6b) or 4 hrs (Figure 2.6c).

**Protein Modeling:** The SPOCK modeling program [87] was used to dock the crystal structures of the core nucleosome (Protein Data Bank (PDB) accession 1AOI [47]) and Ku bound to DNA (PDB accession 1JEY [59]) by multiple alignments of the
phosphodiester backbone of residues 2 to 12 in chain C of 1JEY with the backbone of 10 consecutive residues in chain I of 1AOI. Iterative alignments were made, advancing target phosphate positions in the nucleosome in single nucleotide steps, until steric clashes were minimized to generate the docked structure model shown in Figure 2.8, where residues 2 through 12 in chain C of 1JEY are aligned with residues 10 through 20 in chain I of 1AOI.

RESULTS

Ku displaces histone H1 from DNA ends

 Binding of protein to DNA fragments has been shown in vitro to inhibit XRCC4-Ligase IV’s ability to join DNA ends [88, 89]. Interestingly, Ku and DNA-PKcs are required to relieve the specific inhibition caused by the presence of the linker histone H1 [89], suggesting Ku may play roles in recognizing H1-occluded DNA ends and making these ends accessible to DNA-PKcs and XRCC4-Ligase IV. To address this possibility, we incubated a radiolabeled 31 bp DNA duplex with a large excess (500 fold) of the linker histone. This generates H1-DNA complexes with heterogeneous mobility (Figure 2.1, lane 11), including aggregates that remain in the well (species III). Addition of Ku to these reactions results in progressive redistribution of the H1 bound DNA into two species (species V and VI). Species V is heterogenous, and its mobility can be reduced by antibodies to either H1 or Ku, indicative of the presence of at least one molecule of both. Species VI is more abundant at higher concentrations of Ku, and is consistent with DNA where Ku has mostly evicted H1: the mobility of a large proportion of this species resists addition of the α-H1 antibody, and its mobility is equivalent to DNA saturated by Ku.
only (compare lanes 8-10 to 15-17). Eviction of H1 occurs even when the Ku concentration is 1/50\textsuperscript{th} the concentration of H1. Ku is thus remarkably effective at clearing DNA of histone H1.

**Ku threads on nucleosome-associated DNA ends**

We next addressed whether Ku’s ability to load on H1-occluded ends could be generalized to other physiologically relevant protein-DNA complexes. The nucleosome is the protein occlusion that Ku is most likely to encounter near a DSB produced by exogenous DNA damaging agents (e.g. ionizing radiation). We therefore addressed whether Ku could load on DNA ends in the context of “linker-less” mononucleosomes purified from bulk cellular chromatin (species I, Figure 2.2). To obtain these core nucleosomes, oligonucleosome fragments lacking linker histones were first purified by sucrose gradient and subsequently subjected to exhaustive microccocal nuclease digestion generating a mononucleosome containing 147 bp of DNA [79] (unpublished data). We then assessed whether Ku could bind these mononucleosomes by EMSA. Consistent with previous reports [72, 90], incubation of Ku with either native or reconstituted mononucleosomes shifts the nucleosome’s mobility in a manner dependent on Ku’s concentration (species II and III, Figure 2.2), arguing Ku binds to nucleosome-associated ends.

Given the way Ku loads on naked DNA ends (threading of ends through a central channel), Ku must either have altered how it binds DNA ends, or Ku’s loading involves some form of nucleosome remodeling. We employed a series of homogeneously positioned mononucleosome substrates (e.g. Figure 2.2, Figure 2.3) and a previously described end-blocking strategy [78, 91] to address this issue in greater detail. Ku can
bind if one but not both ends of a naked DNA fragment are blocked by terminal biotin-streptavidin complexes (Figure 2.3a). We therefore reconstituted singly positioned mononucleosomes containing biotin appended to one or both DNA ends. Importantly, as with naked DNA, Ku requires at least one unblocked end before it can shift the equivalent nucleosome substrate (Figure 2.3b). Antibodies to Ku further reduced the mobility of candidate Ku-bound nucleosome species, confirming Ku is stably retained in these species (Figure 2.3c). Ku thus loads on nucleosome-associated DNA ends the same way it loads on naked DNA, by threading DNA ends through its channel.

The singly blocked nucleosome substrate was then used to determine the extent to which nucleosome association reduces Ku’s affinity for DNA ends. In accordance with previous estimates [69, 85], Ku displays an apparent $K_d$ of 0.34 +/- 0.02 nM for the first molecule binding a naked DNA end (Figure 2.4a). However, when confronted with ends on the surface of a nucleosome, Ku binds with an apparent $K_d$ of 6.0 +/- 0.4 nM for the first Ku bound (Figure 2.4b). Ku is thus somewhat less able to load on nucleosome-associated ends, but nevertheless retains an affinity for nucleosome-associated DNA ends that is comparable to other DNA binding proteins’ affinity for naked DNA (e.g. ~3 nM for HMG1 and ~1 $\mu$M for Rad51 [92-94]).

Once bound, Ku can translocate internally on linear DNA, allowing successive molecules of Ku to load on the same DNA end [95]. In EMSA experiments, this appears as “ladder” of distinct species of reduced mobility, each with an additional molecule of Ku (Figure 2.3a and [72, 95]). Multiple species are similarly observed upon addition of Ku to nucleosome-associated ends (lanes 10 and 11, Figure 2.3b; lane 10 in Figure 2.5a). We purified these species, and determined their protein complement by semi-quantitative
western analysis (Figure 2.5b). Ku and a representative histone (H3) are present in each species as appropriate. Additionally, the ratio of Ku to histone in species VI is 2.1 times that of species V, consistent with the presence of one and two molecules of Ku per nucleosome-associated end in the two respective species. Significantly, while each molecule of Ku loads on naked DNA ends essentially independently of prior molecules loaded (Figure 2.4a and [96]), the nucleosome strongly resists loading of a second molecule of Ku (Kd of second molecule >25 nM; Figure 2.4b).

**Ku peels DNA ends from the nucleosome surface**

Ku’s ability to bind nucleosome-associated DNA ends and even translocate internally to some extent indicates Ku must alter nucleosome structure in some manner. We considered three possibilities: Ku could reposition the histone octamer away from the DNA end (“pushing”), Ku could leave the nucleosome in its initial position and “peel” the DNA end away from octamer surface, or Ku could evict a subset of core histones near the DNA end. To address these possibilities, we generated a 187 bp substrate with the nucleosome positioned at one DNA end, leaving the other DNA end spaced 40 bps from the histone octamer (e.g. as in [86]). This end distal to the nucleosome was also blocked with a biotin-streptavidin complex, thereby forcing Ku to load from the nucleosome-associated end (Figure 2.6a). Hydroxyl radical footprinting of this substrate shows a 10 bp phasing indicative of a nucleosome positioned at the Ku-accessible DNA end (Figure 2.6b, lane 1; top of lane to Pst1 marker). If Ku’s loading on the open DNA end were to push the nucleosome onto the 40 bps of naked “linker” DNA, the phased hydroxyl radical sensitivity would either be shifted away from the end (if all nucleosomes in the sample were homogeneously “pushed”) or lost over the entire length of the substrate (if the
nucleosomes were pushed various distances). Alternatively, DNA peeling would be apparent as a loss of phasing limited to the region near the accessible DNA end.

Addition of Ku to our footprinting reactions has a general quenching effect, resulting in an overall reduction of hydroxyl radical cleavage (Figure 2.6b). To account for this, we included a control substrate where both DNA ends are blocked with biotin-streptavidin complexes (Figures 2.6b and 2.6c, lane 4). Our data shows that even at high concentrations of Ku, there are no major changes in hydroxyl radical sensitivity of the linker DNA (0 to 40 bps away from the label) or in the pattern of the phasing over most of the nucleosome (40 to approximately 140 bps from the label) (Figures 2.6b and 2.6c, compare lanes 3 and 4). We therefore conclude that in the majority of molecules, loading of Ku does not alter the translational location of the nucleosome on this DNA fragment.

However, increasing concentrations of Ku result in a correlating loss of phases near the accessible end, consistent with Ku peeling the DNA end away from the surface of the nucleosome. As expected, the phasing pattern of the doubly blocked nucleosome in this region is resistant to addition of Ku (Figure 2.6c, compare lanes 3 and 4). When Ku is present at three fold excess over the nucleosome, loss of phasing extends over the first 40 bp, with some protection observed as much as 50 bp (1/3rd of the length of the nucleosome) away from the end. Parallel experiments indicate the majority of nucleosome substrate possesses more than one molecule of Ku under these conditions (Figure 2.5, lane 10), arguing a molecule of Ku is able to load and translocate internally 40-50 bp, allowing a second molecule to load. Thus, Ku displays the unique ability to thread on nucleosome-associated DNA ends, gaining access to the DNA end by a peeling mechanism.
It is possible that the nucleosome accommodates loading of Ku through displacement of some subset of core histones. To address this possibility, we first loaded Ku onto nucleosome-associated ends, then added a large excess of linear competitor DNA to both remove Ku and ensure potentially displaced histones cannot be reincorporated into the nucleosome. Removal of Ku from the nucleosome restores the nucleosome to its original mobility (Figure 2.6d), arguing against eviction of histones.

**Ku fails to bind RAG-sequestered signal ends**

Ku must also recognize and promote the joining of protein-occluded double strand break intermediates during V(D)J recombination, a genome rearrangement required for assembly of the mammalian immune system’s mature antigen specific receptor genes. V(D)J recombination requires one each of the two types of targeting signals (12-RS and 23-RS), and is initiated when the RAG1 and RAG2 proteins bind this pair of signals and cleave the flanking chromosomal DNA. Importantly, RAG proteins remain in a complex with the paired signals after cleavage (paired signal end complex; SEC), and this complex is sufficient to block signal end joining in both extract- [37-39] and purified protein-based [88] in vitro assays. Nevertheless, signal end intermediates in V(D)J recombination are efficiently joined together by NHEJ in cells. We therefore utilized an EMSA to test if Ku could directly load on RAG-bound signal ends.

We generated an SEC in vitro (see Figure 2.7a) by incubating purified RAG proteins and HMG1 (high mobility group protein 1) with recombination signal-containing oligonucleotide duplexes. A stable SEC (species IV) requires RAG1, RAG2, HMG1, and both 12-RS and 23-RS to be present: this species is inefficiently formed when one of the signal sequences is omitted (Figure 2.7b; compare species III and IV, lanes 3 and 4) or if
one of the signals is substituted with an oligonucleotide duplex composed of irrelevant sequence (unpublished data). We used a pair of RS-containing oligonucleotide duplexes to generate the SEC instead of a continuous DNA fragment terminating in signals ends (as is typically generated in vivo) since this keeps the complex sufficiently small to be resolved by EMSA. However, these duplexes possess an end distal to the site of cleavage that is anomalously accessible. We therefore selectively blocked access to these ends as described above, by appending biotin to the appropriate ends of the oligonucleotides when synthesized, and including streptavidin (SA) in these reactions (Figure 2.7a and Figure 2.7b, species V).

Addition of Ku to SEC mixtures generates a new species (species VI) with increased mobility relative to the streptavidin-blocked SEC (species V). However, this probably reflects loading of Ku onto a minor population of incompletely formed SEC (evident as species with heterogeneous mobility in lanes 4 through 7) because Ku changes neither the mobility nor the intensity of accurately formed SEC (Figure 2.7b, compare species V in lane 5 to species V in lane 8). Furthermore, the mobility of species V was significantly reduced by an antibody to the MBP tag on the RAG proteins, but not an antibody directed against Ku, excluding the possibility that Ku was present in the SEC, but our EMSA was unable to resolve this species. Accurately formed SEC similarly resisted loading of Ku when incubated for longer times (1 hour), at higher temperature (37°C), with increased salt (150 mM KCl), using either Mg$^{2+}$ or Ca$^{2+}$, with higher concentrations of Ku (250 nM), or when accompanied by XRCC4-ligase IV, DNA-PKcs, and ATP (unpublished data). Therefore, we conclude that Ku, either alone
or together with core NHEJ factors, is probably insufficient to recognize and promote joining DNA ends in the context of the SEC.

DISCUSSION

In vitro, at least 50 bps of naked DNA is required to assemble a functional ligation complex at a DNA end [97], but DNA double strand breaks in cells are typically occluded by nucleosomes, linker histones, and other proteins involved in chromatin structure. Therefore, the ability of Ku to load onto such ends is likely a major determinant of the efficiency of cellular NHEJ given Ku’s critical place in break recognition and subsequent nuclease action of the NHEJ complex.

Ku’s unusual structure, which requires DNA ends to be threaded through a circular protein channel, might argue it is poorly suited for loading on protein-occluded ends. Consistent with this argument, we show the RAG proteins bound to signal end intermediates in V(D)J recombination strongly resist Ku’s ability to recognize and load on these DNA ends. Nevertheless, NHEJ must resolve these intermediates in cells to maintain intact receptor loci in certain contexts, and it may be critical in reducing RAG-mediated transposition activity [98]. The prior disassembly of the SEC by factors extrinsic to the core NHEJ machinery and the SEC (e.g. by proteolysis; [99-101]) may thus be the limiting step in resolution of this class of protein-occluded ends.

In contrast, Ku appears readily able to displace the more weakly bound linker histone, H1. Since H1 occlusion is sufficient to block activity of the NHEJ ligase (Ligase IV) in vitro [89], this argues Ku’s ability to recognize broken ends occluded by H1 may be the important first step in allowing NHEJ to act on breaks generated in linker regions
of chromatin. However, we note linker histones may be bound more tightly in the context of higher order chromatin, thus this question should be re-addressed when it becomes possible to accurately recapitulate higher-order chromatin structures in vitro.

Surprisingly, Ku can also recognize DNA ends on the surface of nucleosomes and does so by the same mechanism it uses for naked DNA: threading DNA ends through its central channel. Loading of Ku can be accompanied by peeling as much as 1/3rd of the nucleosomal DNA away from the histone octamer. However, unlike traditional chromatin remodeling, the nucleosome structure (i.e. octamer composition or translational position) can remain otherwise largely unperturbed, and energy from ATP hydrolysis is not required. Instead, Ku presumably takes advantage of transient dissociations of DNA (“breathing”) as it enters and leaves the nucleosome [102]. The crystal structures of Ku bound to a DNA end [59] and the nucleosome core particle [47] also suggest Ku may be particularly well-suited for loading on nucleosome-associated DNA ends as the two structures can be docked with only slight alteration of the DNA path (Figure 2.8). We propose that Ku’s narrow β-strand bridge portion of its DNA binding channel can act as a wedge between the DNA end and the histone octamer surface, allowing Ku to pass DNA ends through its channel while minimally disrupting histone-DNA interactions. Once bound, Ku may then translocate internally, most likely by being pushed by DNA-PKcs [91], XRCC4-Ligase IV [97], or a second molecule of Ku.

Ku’s ability to load on chromatinized ends and translocate inwards may provide cells with several benefits. As suggested above, Ku alone may be sufficient for recognition of chromatinized ends and activation of NHEJ, allowing this pathway to
proceed without always requiring chromatin remodeling. Alternatively, loading of Ku may be a critical first step in directing other factors to perform a limited remodeling of chromatin at ends. In yeast, Ku interacts with the SWI/SNF family remodeling complex, RSC [103, 104], and Ku’s ability to peel DNA ends from the nucleosome surface could help orient the direction of this complex’s activity such that nucleosomes are pushed away from break sites.

NHEJ thus may repair double strand breaks with minimal or no remodeling of flanking chromatin. In contrast, the other major double strand break repair pathway, homologous recombination, is associated with the removal of nucleosomes within several kilobases of the break site [105]. This difference may help rationalize why NHEJ is the preferred repair pathway in differentiated cells [44, 106], where significant disruption of chromatin state could lead to inappropriate gene activation.

ACKNOWLEDGEMENTS
This chapter is adapted from a published manuscript [107]. Steven Roberts was responsible for all experiments and wrote the manuscript with Dr. Dale Ramsden. We thank Brenda Temple (UNC) for assistance in protein modeling, the laboratory of Dr. Yi Zhang (UNC) for the gift of purified histone H1, Dr. Jonathon Widom (Northwestern University) for providing the 601.2 nucleosome positioning sequence, and the members of the Ramsden lab for helpful discussion. This work was supported by US Public Health Service Grant CA-84442 to Dale Ramsden, a Leukemia and Lymphoma Society Scholar.
**Figure 2.1: Ku's DNA binding activity at H1 bound DNA ends.** Increasing concentrations of Ku (0, 2, 10 nM) were incubated for 10 minutes at 25°C with 1 nM of a 31 bp DNA duplex (lanes 1 and 5-10), or the same duplex pre-bound for 10 minutes at 25°C with 100 nM histone H1 (lanes 2-4) or 500 nM H1 (lanes 11-17). The presence of Ku or H1 in various species was confirmed by supershift with an α-Ku antibody (lanes 4, 7, 10, 14, 17) or α-H1 antibody (lanes 3, 6, 9, 13, 16). The compositions of DNA protein complexes were inferred from relative mobilities and antibody supershifts as noted at the sides of the panel; species I, naked DNA duplex; species II, heterogeneous complexes of H1 bound to DNA; species III, H1 saturated DNA, species IV, DNA bound by one molecule of Ku; species V, DNA bound by Ku and H1; species VI, DNA bound by two molecules of Ku. Antibody supershifts are denoted as the upper arrows within a bracket.
Figure 2.2: Ku's DNA binding activity at nucleosome-associated ends. 10 nM HeLa-derived core mononucleosomes (lanes 1-4), reconstituted core mononucleosomes (lanes 5-8), or 147 bp naked DNA (lanes marked D) were incubated with 50 nM Ku (lanes 2, 6, and D), 100 nM Ku (lanes 3 and 7), or 200 nM Ku (lanes 4 and 8) for 10 minutes at 25°C. The resulting complexes were separated by gel electrophoresis and visualized by SYBR Green staining. Roman numerals indicate various Ku-bound species as described in the results.
Figure 2.3: Assessment of Ku's ability to thread on nucleosome-associated DNA ends. DNA substrates containing a biotin at one or both DNA ends (as indicated) were incubated with 5 μM streptavidin (except as noted) at 25°C for 5 minutes. **(a)** 10 pM naked DNA substrate was incubated at 25°C for 10 minutes without Ku (lanes 1-2 and 11-12) or with 0.5 nM (lanes 3 and 10), 1 nM (lanes 4 and 9), 2 nM (lanes 5 and 8), or 4 nM Ku (lanes 6 and 7). **(b)** As in panel (a), except using reconstituted nucleosomes incubated with 10 nM (lanes 4 and 9), 20 nM (lanes 5 and 10), or 30 nM Ku (lanes 6 and 11) and containing an additional 20 minute 37°C incubation. **(c)** As in panel (b) except nucleosomes were incubated with 50 nM Ku (lanes 3 and 4) and an α-Ku antibody (lanes 2 and 4). Lanes 5-7 contain the corresponding naked DNA substrate incubated with (lane 5) or without (lanes 6 and 7) 1 nM Ku. The inferred compositions of various species are noted at the side of each panel. In panel (a), H indicates DNA duplexes bound by more than three molecules of Ku.
Figure 2.4: Affinity of Ku for nucleosome-associated DNA ends. The fraction of species with one Ku bound (squares) or more than one Ku bound (triangles; see also Figure 2.5) were determined for various concentrations of Ku (50 pM to 25 nM) incubated at 25°C for 10 minutes and then 37°C for 20 minutes with 10 pM (a) naked DNA or (b) reconstituted nucleosome. As described in materials and methods, the noted $K_d$s were approximated, after curve fitting, as the amount of Ku required to bind 50% of substrate.
**Figure 2.5: Analysis of the protein composition of Ku-shifted species.**

(a) 80 nM singly blocked naked DNA (lanes 1-5) or nucleosome substrates (lanes 6-10) were incubated with 60 nM (lanes 3 and 8), 120 nM (lanes 4 and 9), or 240 nM (lanes 5 and 10) Ku at 25°C for ten minutes. Reactions were then incubated at 37°C for 5 minutes before separating the formed complexes by polyacrylamide gel electrophoresis. Complexes were visualized by ethidium bromide staining and the species of interest were excised as indicated with boxes. (b) In lanes 4-9, the protein complements of species gel purified in panel (a) (labeled I-VI) were analyzed by western blotting using α-Ku and α-histone H3 antibodies (as indicated to the right of the panel) after electroelution, precipitation, and SDS-PAGE. In lanes 1-3 (Standards), 80, 40, and 20 pmol of Ku or nucleosome (e.g. H3 dimer) were loaded directly to verify the western is semi-quantitative.
Figure 2.6: Hydroxyl radical footprinting of nucleosomes in the presence of Ku. (a) The footprinting substrate consists of a 187 bp DNA fragment where the previously described positioning sequence locates the nucleosome at one end, leaving a 40 bp linker of naked DNA whose end is fluorescently labeled and blocked by a biotin-streptavidin complex. Pst1 digestion produces a 40 bp fragment that marks the boundary between the nucleosome and naked DNA (panel b, lane M). (b and c) 80 nM singly (lanes 1-3) or doubly (lane 4) blocked nucleosomes were incubated with 0 nM (lane 1), 120 nM (lane 2), or 240 nM Ku (lanes 3-4) at 25°C for ten minutes followed by a 5 minute 37°C incubation. Reactions were placed on ice, treated as described in the methods, and electrophoresed for either 1.5 hrs (b) or 4 hrs (c). Lane L represents the mobility of a Cy5-labeled 10 bp ladder. (d) 30 nM Ku was incubated with 10 pM nucleosome (lanes 1-4) and 1 nM Ku incubated with naked DNA (lanes 5-7) for ten minutes at 25°C followed by a 20 minute 37°C incubation. 500 nM of naked 25 bp duplex competitor was added either before (B) or after (A) addition of Ku. Inferred complex composition is indicated to the left of the panel; species I, nucleosome; species II, nucleosome bound by one molecule Ku; species III, nucleosome bound by two molecules of Ku.
Figure 2.7: Ku's DNA binding activity at RAG bound signal ends. (a) A signal end complex (SEC) is defined here as RAG1, RAG2, and HMG1 bound to two oligonucleotide DNA duplexes, one containing the 12-RS and the other containing the 23-RS. A streptavidin-biotin complex was used to block the ends of the DNA duplexes distal to the site of cleavage. (b) Approximately 10 ng/μL of purified RAG1 and RAG2 and 425 nM HMG1 were incubated with 0.4 nM radiolabeled 23-RS containing and 10 nM 12-RS containing DNA fragments at 37°C for 10 minutes. 5 μM streptavidin was added at 25°C for 5 minutes prior to addition of 25 nM Ku. Antibody supershifts required an additional 10 minute room temperature incubation step with either 0.2 μg of the monoclonal antibody to Ku or 1 μL of a polyclonal antisera recognizing the maltose binding domain fused to recombinant RAG proteins. The inferred composition of each of the generated species is indicated to the side of the panel: species I, 23-RS; species II, HMG1 bound 23-RS; species III, RAGs and HMG1 bound to the 23-RS; species IV, SEC; species V, streptavidin blocked SEC (upper arrow indicates α-MBP supershift); species VI, Ku bound to incompletely formed SEC (upper arrow indicates α-Ku supershift).
Figure 2.8: Model of Ku bound to nucleosome-associated DNA ends. The core nucleosome structure (PDB accession 1AOI [47]) and the structure of Ku bound to DNA (PDB accession 1JEY [59]) were docked by aligning the DNA within these structures as described in methods.
CHAPTER 3

KU, AN AP LYASE, HAS A DIRECT ROLE IN EXCISION OF NUCLEOTIDE DAMAGE NEAR BROKEN ENDS

INTRODUCTION

Mammalian cells require nonhomologous end joining (NHEJ) for efficient repair of chromosome double strand breaks, including those made by ionizing radiation [28] and during immunoglobulin class switch recombination [34, 35]. A defining feature of these breaks is the presence of damage to flanking nucleotides, including sites of base loss (abasic or AP sites) [29]. At single strand breaks, 5’ terminal abasic sites are excised by pol β’s 5’dRP lyase activity [108-111], but it is unclear how NHEJ copes with this damage near double strand breaks. We show here both in vitro and in cells that accurate and efficient repair of such breaks by NHEJ also requires excision of the abasic site by an AP lyase. Surprisingly, the heterodimeric NHEJ factor Ku can be implicated as the source of this activity. Ku nicks DNA 3’ of abasic sites, it is inactive on the AP site analog tetrahydrofuran, and its 70kD subunit forms a Schiff base covalent intermediate with abasic sites which is characteristic of lyases like pol β [112]. Importantly, the capacity of cell extracts to excise AP sites near double strand breaks was reduced approximately 5-fold when extracts were deficient in Ku, indicating it is the primary source of AP lyase activity on substrates relevant to its biological role. Ku had previously
been presumed only to recognize ends and recruit other factors that processed ends: our data supports an unexpected direct role for Ku in end processing steps as well.

METHODS

Protein Preparations: Recombinant purified Ku, XRCC4-LigaseIV, Pol β and pol λ were prepared as previously described[113]. An XLF cDNA (the gift of K. Meek, MSU) was introduced into pFASTBAC1 with a C terminal hexahistidine tag, and purified by successive chromatography on HisTrap and MonoQ columns (GE biosciences). Ku (untagged) and DNA-PKcs were purified from HeLa cells as described in [114] and [115] respectively. HeLa and CHO cells were extracted in 10 mM TRIS-HCl pH8.0, 600 mM KCl, 0.1% NP-40 substitute (Fluka), 20% glycerol, 1 mM EDTA, 1mM DTT, and protease inhibitors (Sigma, P8849). DNA was removed from extracts by precipitation with 0.1% polyethylene imine (Sigma), followed by sequential adsorption of the supernatant to phosphocellulose (Sigma), and hydroxyapatite (Biorad). Extracts were then dialysed or diluted until equivalent to 10 mM TRIS pH 8.0, 250 mM KCl, 0.1% NP40, 10% glycerol, and 1 mM EDTA. HeLa cell extracts were further immunodepleted by two sequential adsorptions to protein A Sepharose beads pre-loaded with either pre-immune serum (Mock depleted) or serum from rabbits immunized against human Ku (Ku-depleted). HeLa and CHO cell extracts were then analyzed by Western blotting using antibodies against Ku (raised against purified human Ku heterodimer), pol β (ab26343, Abcam), and actin (A2066, Sigma).

DNA substrates: Plasmid templates for 250 bp substrates were generated by first appending the restriction sites needed to produce desired end structures to mouse Jκ1
sequence using polymerase chain reaction (PCR) and cloning the PCR product. The primers used to make template for 5’dRP-EJ were

5’TGATGCCGAGCTCTTCTATTTAACTACCCCTGCTTCTTTGAG and

5’CGTTCAAGTGCTCTTCTATATGGGAATCAACGTAAGTAG, and for AP-EJ they were 5’TGATGCCGAGGTGCTCAAATATGGCTACCCCTGCTTCT and

5’CGTTCAAGTGCTCTTCCATATGGGAATCAACGTAAGTAG. Once cloned (TOPO-TA, Invitrogen), the sequence verified plasmid templates were again amplified, but using 5’TGATGCCGAGCTCTTCAUTTAATTAC and

5’5’CGTTCAAGTGCTCTTCCUAATGG for 5’dRP-EJ and

5’TGATGCCGAGGTGCTCAAATATGG for AP-EJ. Substrates for in vitro reactions were labeled at this step by inclusion of $^{32}\text{P-}\alpha$-dATP during PCR. 5’dRP-EJ was then digested with Bsp QI and AP-EJ with Bsa I, followed by purification of DNA using a Qiaquick PCR clean up kit (Qiagen).

Oligonucleotides for substrates were obtained from Integrated DNA technologies and were labeled at the 3’ end with $\alpha^{32}\text{P-cordycepin}$ and additionally “cold” 5’ phosphorylated with T4 polynucleotide kinase before annealing with complementary strands to form duplex substrates. The labeled strand for 5’dRP-DSB and 5’dRP-SSB was 5’UGGAAATCAACGTAAGTAG, while the labeled strand for AP-DSB was 5’GUGGAAATCAACGTAAGTAG. For both DSB substrates, the complementary strand was 5’TCTACTTACGTTTGTCTTTTC, while the SSB substrate required annealing of the labeled strand noted above to both 5’biotin-TEG-TCTACTTACGTTTGTCTTTTC and 5’biotin-TEG-TGGAGGCACCAAGC. Ends of the 5’dRP-SSB substrate were additionally blocked by
pre-incubation of the substrate with 1 μM streptavidin (Pierce) for 5 minutes. The
tetrahydrofuran containing substrate was generated by substituting dU in the labeled
strand for 5’dRP-DSB with dSpacer. Bleocin damaged substrates were generated by
annealing the oligonucleotide
5’TCTACTTACGTTTGATTTCAGCTTGGGTGCCTCCA to
5’TGGAGGCACCAAGCTGGAAATCAAACGTAAGTAG, and incubation of 100 fmol
of the resulting duplex with 100 pmols Bleocin (EMD Biosciences) for 10 minutes at
37°C. For all dU-containing substrates, abasic sites were made by incubation with 0.02
units uracil DNA glycosylase (NEB) per fmol substrate for 5 minutes at 37°C. Reduced
AP site substrates were made by treating these glycosylated substrates with 50 mM
NaBH₄ for 20 minutes on ice before an additional purification (Qiaquick PCR cleanup,
Qiagen). The concentrations of all substrates were determined by Qubit (Invitrogen)
using high sensitivity dsDNA (EJ substrates) or ssDNA (oligonucleotide substrates)
stains.

In vitro end joining reactions: 5 nM of 250 bp substrates with noted end structures were
pre-incubated with 20 nM recombinant Ku, 10 nM DNA-PKcs, 40 nM XRCC4-
LigaseIV, and 80 nM XLF in a standard reaction buffer (25 mM NaPO₄ pH 7.4, 125 mM
KCl, 0.1 mM EDTA, and 1 mM DTT) supplemented with 10% polyethylene glycol for 5
minutes at 25°C. Reactions were started by addition of 2 mM MgCl₂ and 100 μM ATP
and incubated at 37°C, stopped by deproteinization, and analyzed by native 5%
polyacrylamide gel electrophoresis (PAGE).

Cellular NHEJ assays: CHO cells grown to approximately 70% of confluency in 6 well
dishes were transfected with a mixture of 5 ng of linear NHEJ substrates and 1.5 μg of
pMaxGFP (Amaza) using Fugene6 (Roche). Transfected cells were harvested 5 hrs later and small molecular weight DNA recovered in a Hirt supernatant [116]. Head-to-tail junctions were detected by amplification of 100 ng of the Hirt supernatant with primers 5’CCTTGGAGAGTGCCAGAATC and 5’AAAGCAAAGCTGGGAATAGGC for 28 cycles. Successful transfection was verified by amplification of the 100 ng of the Hirt supernatant for 20 cycles with primers 5’TGATGGGGCTACGGCTTCTAC and 5’GCTGCCATCCAGATCGTTAT. PCR amplifications were then analyzed by electrophoresis on a 6% polyacrylamide gel under native conditions and southern analysis performed using 5’CCGCATCGAGAAGTACGAGGA (5’ labeled with ³²P with T4 polynucleotide kinase) as a probe. For key samples, an aliquot was cloned using a TOPO-TA kit (Invitrogen), and 20 cloned junctions analyzed by restriction digestion.

Lyase reactions: 1 nM oligonucleotide substrate was incubated at 37°C in standard reaction buffer, stopped by addition of 50 mM NaBH₄, incubated for 20 minutes on ice, and analyzed by denaturing 15% PAGE.

NaBH₄ cross-linking: Analyses of Schiff base intermediates were performed by incubation with 10 nM substrate for 20 minutes at 37°C in standard reaction buffer supplemented with 5 mM NaBH₄. Reactions were then subjected to SDS-PAGE, total protein detected by staining with SYPRO orange (Invitrogen), and radioactive species detected by phosphorimage analysis of the dried gel.
RESULTS

NHEJ removes abasic sites before end ligation

We first reconstituted NHEJ of ends with damaged termini in vitro using purified proteins. 250 bp linear DNA substrates with abasic sites located either at the extreme 5’ terminus (5’dRP-EJ, Figure 3.1a) or embedded one nucleotide internal to the 5’ end (AP-EJ, Figure 3.1a) were incubated with the core end joining machinery (purified human Ku, DNA-PKcs, XLF/Cernunnos, and XRCC4-LigaseIV). Surprisingly, the core factors were sufficient to join abasic site associated ends without addition of a known 5’dRP/AP lyase (hereafter termed “AP lyase”) (lanes 2 and 5, Figure 3.1a). Nevertheless, joining activity was very low if the substrates were pre-treated with NaBH4 (Figure 3.1a, lanes 3 and 6), which reduces the abasic site and specifically blocks the ability of AP lyases to act on this lesion [108]. This result implies both that joining requires removal of abasic sites by an AP lyase before ligation can occur, and that AP lyase activity was present in our purified core factor preparations. We further addressed this possibility by characterizing ligation products from reactions with the “standard” (non-reduced) AP site substrate (Figure 3.2a). Ligation products were not detectably sensitive to alkali treatment (Figure 3.2b), indicating junctions with embedded abasic sites are rare. Additionally, our substrates were designed such that joining after precise excision of the abasic site (e.g. by an AP lyase) generates a new unique restriction site (Ase1 for the 5’dRP-EJ and Nco1 for AP-EJ) at the junction. Only trace levels of junctions were insensitive to the diagnostic restriction digest (Figure 3.2c), confirming that the abasic sites within these substrates were typically precisely excised prior to joining.
We next asked if an AP lyase was also important in cellular repair of such substrates. We introduced into Chinese hamster ovary (CHO) cell lines two versions of the AP-EJ substrate, containing either standard or reduced AP sites, as well as an undamaged version of this substrate, where a T was substituted for the abasic site. Head-to-tail joining was assessed by semi-quantitative polymerase chain reaction. Wild type CHO cells (K1) were similarly proficient in joining the undamaged substrate when compared to the substrate with an embedded normal AP site (Figure 3.1b, lanes 1 and 2). As with in vitro generated junctions, we used diagnostic restriction enzyme sites (Figure 3.1c) to probe the accuracy of joining. For the undamaged substrate, 60% of junctions were sensitive to Nde I, indicative of frequent joining without deletion (Figure 3.1d). A similar proportion (45%) of junctions using the AP site containing substrate involved precise excision of the embedded AP site, but these junctions were otherwise not deleted (Nco I sensitive). An embedded standard AP site thus had little impact on either the efficiency or the gross accuracy of joining. In contrast, reduction of the AP site (and consequent blocking of AP lyase activity) resulted in much less efficient joining, and junctions were typically accompanied by extensive deletion of flanking DNA (25-100 bp; Figure 3.1b, lane 3). We conclude that precise excision of the AP site by an AP lyase is critical for efficient and accurate resolution of this substrate in cells, as previously shown in vitro (Figure 3.1a).

Joining of these substrates was next evaluated in xrs6, a CHO line derived from K1 that is deficient in Ku. Notably, deficiency in Ku had little impact on joining of the undamaged end (T; Figure 3.1b, compare lanes 1 and 5), consistent with prior studies arguing that undamaged ends can still be efficiently resolved in the absence of classical
NHEJ [43, 117]. However, the presence of a standard AP site more clearly impacted joining in Ku deficient cells as compared to wild type cells, since both efficiency and accuracy were reduced relative to the undamaged substrate (Figure 3.1b, compare lanes 5 and 6). Of 20 cloned junctions made in Ku-deficient cells using the standard AP site substrate, only 2 (10%) involved precise excision of the AP sites (i.e. were Nco I sensitive). NHEJ is thus uniquely effective in coupling AP lyase activity to joining.

**Ku is an AP lyase**

The in vitro experiments described in Figure 3.1a indicated that a candidate for this lyase activity is one of the known NHEJ core factors. We tested each NHEJ core factor preparation individually for AP lyase activity using abasic site-containing oligonucleotide substrates (5’dRP-DSB, AP-DSB) analogous to the 250 bp NHEJ substrates previously described in Figure 3.1. Of the four core factors in our prior in vitro assay (Ku, XRCC4-ligase IV, XLF and DNA-PKcs), only our preparation of the end recognition and scaffold protein, Ku, had significant activity. Both 5’dRP and embedded AP sites were efficiently excised (Figure 3.3a, lanes 3 and 9), resulting in a species that co-migrates with alkali-cleaved samples (lanes 4 and 10). This is consistent with cleavage 3’ of the abasic site and production of a 5’ phosphorylated terminus. Ku was also inactive on a substrate where the abasic site was substituted with tetrahydrofuran (THF), an abasic site analog resistant to lyase activity [108] (lane 6).

AP lyases cleave 3’ of abasic sites by forming a Schiff base between a catalytic lysine and the 1’ carbon of the abasic site [112]. Adding NaBH₄ to the reaction can stably trap this covalent protein-DNA intermediate: if the DNA is radiolabeled, the trapped intermediate can then be visualized by denaturing SDS-PAGE and phosphorimage
analysis. We performed this reaction with the 5’dRP-DSB substrate and preparations of Ku heterodimer, purified either with a hexahistidine tagged 70kD subunit (Figure 3.3b, lanes 1-3), or with an untagged version of the 70kD subunit (Figure 3.3b, lane 4). We generate one major radiolabeled species dependent on both the presence of NaBH₄ and an abasic site (Figure 3.3c). This radiolabeled protein has reduced mobility relative to that of the free Ku70 kDa subunit, consistent with covalent linkage of a 6 kDa DNA to Ku70. Additionally, the crosslinked species using hexahistidine tagged recombinant Ku 70 has slightly reduced mobility relative to the cross linked species detected using untagged Ku70 (Figure 3.3c, compare lanes 3 and 4). We conclude Ku is an AP lyase, and its primary active site is in the 70 kD subunit.

How does Ku’s AP lyase activity compare to previously described AP lyases? Pol β is the primary source of 5’dRP lyase activity at single strand breaks [108-111], and we detect no significant activity for Ku in this context, as expected. In contrast, Ku possessed between 2 and 4 fold greater specific activity than DNA polymerase β when abasic sites were located near a double strand break (5’dRP-DSB, AP-DSB; Figure 3.4 a,b). Pol λ, like pol β, is a member of the pol X family of polymerases and possesses lyase activity at single strand breaks [14]. Additionally, pol λ specifically associates with NHEJ core factors at DNA ends [13, 74, 113]. However, pol λ’s intrinsic 5'dRP lyase activity is even more restricted to single strand breaks than pol β, and consequently addition of pol λ to reactions with NHEJ core factors, including Ku, did not significantly augment activity seen with Ku alone (Figure 3.4c). These biochemical data are consistent with recent genetic studies indicating that specific deficiency in the lyase activity of pol4, a polymerase implicated in NHEJ and the only pol X member in S. cerevisiae, had no
impact on cellular repair of ends with nearby abasic sites [118]. We suggest that contacts with template and primer terminus help properly orient pol β and pol λ’s lyase active site on AP sites associated with single strand breaks, and the absence of these contacts at termini of double strand breaks makes them less active in this context. Similar loss of activity near double strand break termini has also been observed for other base excision repair enzymes [119].

**Ku is the dominant AP lyase in cell extracts**

We further addressed the significance of Ku’s AP lyase activity by comparing the relative cumulative AP lyase capacity of normal and Ku-deficient whole cell extracts. We generated two pairs of whole cell extracts: in the first pair, a HeLa extract was either mock depleted or immunodepleted of Ku, while the second pair of extracts was made either from the wild type or Ku-deficient CHO lines used previously (K1, xrs6). Using Western analysis (Figure 3.5) and total protein stain (Figure 3.6b) we verified that these extract pairs possessed Ku (Mock, K1) or not (α-Ku, xrs6), but were otherwise indistinguishable. The AP lyase activity for each extract pair was then compared for AP sites in three contexts: at the end of a single strand break (5’dRP-SSB), at the end of a double strand break (5’dRP-DSB), and embedded one nucleotide from the 5’ end (AP-DSB). As expected, deficiency in Ku did not significantly impact lyase activity on the 5’dRP-SSB substrate (Figure 3.6a, top panel). In contrast, extracts deficient for Ku display 2 to 7 fold reduced activity when abasic sites are near 5’ termini of double strand breaks (Figure 3.6a, bottom two panels). As another means of addressing Ku’s contribution to lyase activity in these extracts, we assessed covalent protein-DNA intermediates trapped by NaBH₄ (e.g. as in Figures 3.3b and 3.3c). We detected a species
that co-migrates with cross-linked recombinant Ku70 in normal but not the matched Ku deficient extracts (Figure 3.6b, 3.6c). Strikingly, the Schiff base formed with Ku70 is by far the most abundant species observed in normal HeLa cell extracts, and is still the primary species observed in normal CHO cell extracts. There is also little evidence for significant amounts of other species in Ku deficient extracts, arguing for a lack of redundancy in this function. Ku was similarly the primary species cross-linked to DNA duplexes damaged with bleocin, a radiomimetic drug (Figure 3.7), confirming that Ku recognizes abasic sites made by oxidative damage in addition to glycolytic abasic sites. We conclude that Ku accounts for the majority of AP lyase activity in cell extracts when AP sites are near 5’ termini of double strand breaks.

DISCUSSION

Ku has been characterized as a factor that recognizes double strand breaks, and then recruits other factors to process and join broken ends. However, we show here that an AP lyase must act during NHEJ for accurate repair of ends with nearby abasic sites (Figure 3.1), and Ku in principle could fulfill this role (Figure 3.3). However, there are many proteins with detectable AP lyase activity and, with the exception of pol β [110], it is unclear if these activities are high enough to be biologically significant. Critically, we show here that Ku is the primary source of AP lyase activity in contexts relevant to its biological role (Figure 3.6). Therefore, we argue Ku has an important and unexpected direct role in processing abasic site damage at broken chromosome ends.

In addition to helping resolve chromosome breaks made directly by ionizing radiation and radiomimetic drugs, Ku’s AP lyase should be important for resolving those
breaks made during aborted base excision repair of clustered damage sites and during immunoglobulin class switch recombination. For example, a role for Ku in processing abasic sites could explain the surprising greater severity of defect in immunoglobulin class switch recombination when comparing Ku-deficient mice to XRCC4- or Ligase IV-deficient mice [46]. This phenotypic difference is atypical: when there is one, Ku deficiency usually results in the less severe phenotype [21, 43-45]. Since abasic sites are the primary trigger for class switch recombination [120], we suggest Ku’s capacity to excise abasic sites near switch recombination intermediates is an important step in allowing their resolution by the non-classical NHEJ pathway that is active in XRCC4- or Ligase IV-deficient cells [46]. In contrast, deficiency in Ku means less efficient AP site removal, reducing the effectiveness of even the alternate end joining pathway in Ku-deficient mice.

ACKNOWLEDGEMENTS

The work discussed in this chapter is adapted from a manuscript written by Steven Roberts and Dale Ramsden. Steven Roberts conducted in vitro ligation experiments, cell transfection experiments, lyase activity assays, and pilot NaBH₄ cross-linking experiments. Jody Havener performed experiments described in Figure 3.2b, c and 3.3b, c. Martin Burkhalter aided with experimental design and conducted additional unpublished cellular experiments. Dale Ramsden conducted cross-linking experiments with bleocin damaged DNA. We would like to thank Christina Strom for her assistance and Drs. Martin Gellert, Thomas Kunkel, Kathryn Meek and Miguel Garcia-Diaz for
their discussions and for reading the manuscript. This work is supported by PHS grant CA 84442 and a Leukemia and Lymphoma Society scholar award to Dale Ramsden.
Figure 3.1: NHEJ of ends with abasic sites. (a) Substrate cartoons show position of abasic site (open circle) relative to intact nucleotides (blocks). 5 nM substrates with normal abasic sites (lanes 1, 2, 4, and 5) or reduced abasic sites (lanes 3 and 6) were incubated with purified core factors at 37°C for 20 minutes (5'dRP-EJ) or 60 minutes (AP-EJ) as described in Methods. (b) AP-EJ variants with an undamaged nucleotide (T), a standard abasic site (AP), a reduced abasic site (rAP), or transfection control DNA only (-) were introduced into wild type (K1) or Ku-deficient (xrs6) CHO cells. Recoveries of head-to-tail junctions (top panel) or the control plasmid (bottom panel) were evaluated by semi-quantitative PCR. Mobility of molecular weight marker DNAs (in 25 bp increments) indicated to right of top panel. (c) Diagnostic restriction sites define accurate junctions. (d) Reactions from (b) were digested with Nde I for undamaged substrate or Nco I for standard AP site substrate transfections.
Figure 3.2: Characterization of in vitro NHEJ junctions. (a) Ligation of substrates without excision of abasic sites will generate junctions sensitive to alkali treatment; alternatively, joining of ends after precise excision of abasic sites will generate junctions sensitive to a diagnostic unique restriction enzyme (Ase I for 5'dRP-EJ, and Nco I for AP-EJ). (b) In vitro ligation products were digested with a restriction site internal to ends (Hinf I) to resolve concatemers, and either mock treated or treated with 80mM NaOH for 20 minutes at 95°C. Shown are head-to-tail junctions (P) and an internal Hinf I fragment (S) after 5% denaturing PAGE. Denaturing 15% PAGE confirmed alkali treatment was sufficient to completely cleave abasic sites present at un-reacted substrate ends (unpublished data). (c) Head-to-tail junctions (P) were PCR amplified as in Figure 3.1b. Junction products from 5'dRP-EJ reactions were then digested with Ase I, while junction products from AP-EJ were digested with Nco I.
Figure 3.3: AP lyase activity of purified Ku. (a) 1 nM radiolabeled substrates with abasic site or lyase-resistant analog (THF) were incubated with 2 nM Ku as indicated (+) at 37°C for 10 minutes (5’dRP-DSB substrate, lanes 1-6) or 40 minutes (AP-DSB substrate, lanes 7-10). Reactions were then treated by addition of 50 mM NaBH₄ or 80 mM NaOH (lanes 4 and 10 only) and analyzed on a denaturing 15% polyacrylamide gel. (b and c) 10 nM 5’dRP oligonucleotide duplex was incubated with 20 nM recombinant Ku either with hexahistidine tagged Ku70 (R; lanes 1-3) or untagged Ku70 (N; lane 4) in the presence of 5 mM NaBH₄ at 37°C for 20 minutes. Reactions were subjected to SDS-PAGE and total protein detected by Sypro Orange (b) before drying and Schiff base intermediates were detected by phosphorimage analysis (c).
Figure 3.4: Specific lyase activities of pol β, pol λ, and Ku. (a and b) 2 nM Ku or pol β was mixed with 1 nM (a) 5' dRP-DSB, or (b) AP-DSB at 37°C. Aliquots were removed at specific time intervals and treated with NaBH₄ to terminate the lyase reaction. Reaction products were separated by electrophoresis and quantified using ImageQuantTL (GE Biosciences). Background loss present in a no protein control was subtracted from experimentally produced product for each time point. (c) The contribution of pol λ’s lyase in the context of other NHEJ core factors was assessed with 1 nM 5'dRP-DSB and either 2 nM Ku, 2 nM Ku + 2 nM XRCC4-LigaseIV, or 2 nM Ku + 2 nM XRCC4-LigaseIV + 2 nM pol λ. Reactions were terminated and processed as in a, b.
Figure 3.5: Characterization of HeLa and CHO cell extracts. 1 μg HeLa based extract (mock and Ku depleted) and 50 μg CHO based extract (K1 and xrs6) was probed with an antibody raised against human Ku. Alternatively, 10 μg of HeLa or CHO extract was probed with antibodies against pol β and actin.
Figure 3.6: AP lyase activity of cell extracts. (a) 10 μg mock depleted or Ku-depleted (α-Ku) HeLa cell extracts and 50 μg wild type (K1) or Ku-deficient (xrs6) CHO cell extracts were incubated at 37°C with 1 nM of either 5'dRP-SSB for 2 minutes (top panel), 5'dRP-DSB for 10 minutes (middle panel) or AP-DSB for 30 minutes (bottom panel). Samples were processed as in Figure 3.2a. Noted below the panels is the reduction of lyase product attributable to Ku deficiency for each matched pair of extracts. (b and c) 10 nM of the radiolabeled 5'dRP-DSB substrate was incubated with 3 μg HeLa extracts or 10 μg CHO extracts (normal or Ku deficient as in panel a) as in 3.2b, c. A reaction with substrate cross-linked to 2 nM recombinant Ku and 20 nM recombinant pol β is included for comparison. Total protein (b) or phosphorimage (c) as in Figure 3.2b,c.
Figure 3.7: \textbf{NaBH}_4 \text{ cross-linking of cell extracts to Bleocin damaged DNA.}
10 nM Bleocin damaged duplex was incubated with 10 μg HeLa based extracts or 50 μg of CHO based extracts (K1 and xrs6) and 5 mM NaBH$_4$. These samples were incubated for 20 minutes at 37°C and subsequently processed as in Figure 3.2b,c.
CHAPTER 4

XLF STIMULATES LOW FIDELITY LIGATION BY NONHOMOLOGOUS END JOINING

INTRODUCTION

Chromosome double stand breaks are potentially lethal to cells as failure to re-establish the broken phosphodiester bonds can result in loss of large regions of genetic information. DNA end ligation is thus the pivotal step of nonhomologous end joining repair of these breaks. However, DSBs often contain ends that are not readily ligatable. Breaks generated by ionizing radiation, radiomimetic drugs, or during class switch recombination all are associated with flanking nucleotide damages (i.e. modified bases or abasic sites) [29, 32], and the synapsed overhangs of the programmed breaks generated during V(D)J recombination often contain mismatches or gaps [121, 122]. Damaged nucleotides, mismatches, and gaps disrupt alignment of the DNA ends and thereby inhibit XRCC4-LigaseIV activity. Previously, NHEJ has been shown to circumvent the presence of abasic sites and gaps by employing processing factors (Ku as a lyase and polymerases λ and μ) to precisely excise lesions or fill gaps. These processing events however require XRCC4-LigaseIV to momentarily release the aligned DNA ends while processing factors access the ends. This re-organization of the synaptic complex could result is dissociation of the complex and increase the potential for DNA translocation [64]. Thus, incorporation of damaged nucleotides or mismatches into the repair product
may be a beneficial means to avoid genomic instability. We propose the recently characterized NHEJ factor XLF/Cernunnos is a key player within a low fidelity NHEJ pathway as it clamps XRCC4-LigaseIV to broken DNA ends and increases ligation of mismatched overhangs.

XLF/Cernunnos is a 33 kDa nuclear protein displaying structural homology to another NHEJ factor, XRCC4 [7, 123]. Deficiency in XLF has been shown to cause a RS-SCID phenotype (characteristic of deficiency in a core NHEJ factor) in several human patients. Generally, the XLF dysfunction in these individuals results from mutations that cause inappropriate splicing or truncation of the protein. However, two point mutations of XLF, R57G and C123R, also cause the RS-SCID phenotype [26]. Further cellular studies indicate XLF-deficient maturing B-cells undergo normal V(D)J recombination but heavy chain class switch recombination is defective [124]. This suggests that XLF’s role during NHEJ is either more limited than the typical core NHEJ factor or that there may be compensation for its absence during V(D)J recombination. Studies looking at XLF function in vitro indicate that XLF exists as a homodimer [123] and associates with XRCC4-LigaseIV as well as DNA [7, 67]. Furthermore, inclusion of XLF in in vitro NHEJ reactions increases the ligation of mismatched ends [65, 66], suggesting this protein may be involved in the resolution of complex DNA ends that are difficult to join.

We show here that in vitro XLF functions as a clamp, stabilizing XRCC4-LigaseIV at Ku bound DNA ends. Under stringent reaction conditions, XLF stimulates ligation of compatible overhangs by increasing the half life on the synaptic complex in a manner complementary to DNA-PKcs. Where DNA-PKcs acts to bring DNA ends
together, XLF maintains the synapsed ends. Ultimately, the presence of XLF and the consequent increased residence of XRCC4-LigaseIV at synapsed DNA ends correlates with decreased ligation fidelity as measured through the ability to ligate across mismatches. This low fidelity ligation may represent an alternative pathway to processing based ligation focused upon avoidance of DNA translocation and deletion instead of accurate, high fidelity repair.

METHODS

Protein Preparations: Recombinant Ku, XRCC4-LigaseIV, and pol μ were expressed and purified as described in [12]. A cDNA encoding FLAG-tagged XLF (the gift of K. Meek, MSU) was inserted into pFASTBAC1, expressed through baculovirus, and purified by successive chromatography on α-FLAG and MonoQ columns (GE biosciences). The R57G mutant of XLF was generated by site-directed mutagenesis using the primers 5’GTGGTCAGCCAGGGAGCCAAGGAGC and 5’GCTCCTTTGGCTCCCTGGCTGACCAC. This protein was expressed and purified in the same manner as the wild type. DNA-PKcs was purified from HeLa cells as described in [115].

DNA substrates: Oligonucleotide based substrates (Integrated DNA technologies) were generated by annealing the oligonucleotides 5’CAGCTGGGAATTCCATATGAGTACTGCAGATGCACTTGCTCGATAGATCTAACATGAGCC and 5’Cy3-GTAGGGCTCATGTTAGATCTATCGAGCAAGTGCATCTGCAGTACTCATATGGAATTCCCAGCTGAG. 295 bp and 305 bp ligation substrates were PCR amplified from
the mouse Jk1 sequence using the respective primer sets,

5’GCTTCGATGAAGAGTGTCGGTGGAGGCACCAAG and

5’GCTTCGATGAAGAGGCTACCCTGCTTCTTTGAGC and

5’GCTTCGATGAAGAGTGTGGTGGACGTTCGGTG and

5’GCTTCGATGAAGAGGCTACCCTGCTTCTTTGAGC. These primers attach a TaqI restriction site to either end of the desired DNA fragment. The products were then cloned in the TOPO-TA vector (Invitrogen) and the resulting plasmids re-amplified with the vector specific primers 5’biotin-TEG-AGTGTGCTGGAATTCGCCCTT and 5’biotin-TEG-GTGATGGATATCTGCAATTCGCCCTT, before restriction digestion with TaqI. Uncut substrate was depleted by incubation with streptavidin beads (Roche) for 20 minutes at room temperature. Digested substrates were treated with Calf Intestinal Alkaline Phosphatase (NEB) for 30 minutes at 37°C, cleaned with a Qiaquick PCR Purification Kit (Qiagen), and subsequently 5’ labeled using $^{32}$P-γ-ATP and T4 Polynucleotide Kinase (NEB). Unincorporated label was removed by size exclusion on a G-25 sephadex spin column (GE Healthcare). The 300bp TT overhang substrate was similarly TOPO-TA cloned with substrate specific primers (described in [125]) before amplification in the presence of Cy5-α-dCTP (GE Healthcare) using the vector specific primers described above. This substrate was then digested with BtsCI (NEB), depleted of uncut substrate with streptavidin beads, and cleaned with a Qiaquick PCR Purification Kit.

*Electrophoretic mobility shift assay (EMSA):* NHEJ complexes were assembled by incubation of 10 nM Cy3 labeled oligoduplex with purified Ku, XRCC4-LigaseIV, and XLF in a standard reaction buffer (25 mM TRIS pH 7.5, 100 mM NaCl, 50 mM KCl, 0.5
mg/mL bovine serum albumin (BSA), 1 mM DTT, 0.1 mM EDTA) at room temperature for 30 minutes. 50 nM of hairpin DNA competitor was added with glycerol to 10% before separation of the resulting complexes by electrophoresis on a 1/3x TBE, 3.5 % polyacrylamide gel as described in [107].

**Ligation assay:** In vitro NHEJ reactions with complementary end substrates were conducted by incubating Ku, XRCC4-LigaseIV, DNA-PKcs, and XLF with the 305 bp DNA (at concentrations indicated in the figure legends and maintaining respective protein to DNA end ratios of 2:1:2:4) in a standard reaction buffer (25 mM Tris pH 7.5, 100 mM NaCl, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, and 200 ng supercoiled DNA competitor) supplemented with 6% polyethylene glycol at room temperature for 0 (under no pre-incubation conditions) or 30 minutes. Ligation was initiated by addition of MgCl$_2$ to 5 mM and ATP to 500 μM. Reactions were incubated at 37°C for 5 minutes and subsequently stopped with 0.1% SDS, 10 mM EDTA, and proteinase K. Reactions were phenol:chloroform extracted and separated on a 1X TBE, 5% polyacrylamide gel. Reactions utilizing the TT mismatch substrate incubated with Ku, XRCC4-LigaseIV, XLF, and pol μ contained slightly different protein to DNA end ratios (see Figure 4.4 legend), were supplemented with 10% polyethylene glycol, and pre-incubated for 10 minutes. These reactions were initiated by addition of MgCl$_2$ to 5 mM and dNTPs to 100 μM, and then processed as described above. Ligation efficiencies were calculated as in [126]. Sequence analysis of mismatch ligation junctions was conducted as in [126].

**Synapsis assay:** As with the ligation assay using complementary substrates, NHEJ factors were pre-incubated for 30 minutes with either a 295 bp DNA substrate or 305 bp DNA substrate in standard reaction buffer supplemented with 6% polyethylene glycol.
Complexes containing the 295 bp substrate were mixed with the corresponding complexes formed with the 305 bp substrate and incubated for 3, 5, 10, or 30 minutes. Ligation was then initiated with MgCl₂ to 5 mM and ATP to 500 μM and allowed to proceed for 5 minutes at 37°C. Reactions were then cleaned through a MinElute Enzyme Cleanup Kit (Qiagen) and digested with 1 unit HinfI (NEB) at 37°C for 15 minutes. Digestion was stopped by addition of 3 reaction volumes of formamide loading dye. Samples were heated to 95°C for 5 minutes and separated on an 8% formamide-urea gel. Fraction complex commitment, a measure of synaptic maintenance, was calculated as:

\[
\text{fraction commitment} = \frac{\%\text{hetero(pre)} - \%\text{hetero}}{\%\text{hetero(pre)}}
\]

RESULTS

**XLF clamps XRCC4-LigaseIV to DNA ends**

To initially address the role of XLF within nonhomologous end joining, we asked whether XLF could form defined complexes with other core NHEJ factors by electrophoretic mobility shift assay (EMSA). This assay utilizes a Cy3 labeled 66bp oligoduplex which migrates at the bottom of a polyacrylamide gel (Figure 4.1b, lane 1). Addition of a limiting amount of Ku produces a single protein-DNA complex migrating with reduced mobility relative to the free DNA substrate (Figure 4.1b, lane 3). Further incubation of the DNA substrate with Ku and XRCC4-LigaseIV generates a heterogeneous smear migrating slower than the defined Ku-DNA complex (Figure 4.1b, lane 4). This indicates a transient interaction of XRCC4-LigaseIV with Ku-bound DNA ends as the complex dissociates during the electrophoresis process. Interestingly, inclusion of even small concentrations of XLF stabilizes the Ku-XRCC4-LigaseIV-DNA complex.
complex generating a single, defined complex of unique mobility (Figure 4.1b, lane 5). This complex requires the presence of all three NHEJ factors indicating it formed through specific protein-protein interactions. Importantly, the R57G mutant of XLF known to cause human RS-SCID due to XLF deficiency [26] forms the ternary complex 10 fold less efficiently than wild type (Figure 4.1, lane 6). Amino acid R57 resides in an alpha helix suggested to facilitate polymerization of an XLF-XRCC4-LigaseIV filament [123] (Figure 4.1a). Thus, XLF’s biological role during NHEJ may involve clamping XRCC4-LigaseIV to DNA ends by filament formation.

**XLF stimulates complementary end ligation under stringent conditions**

Upon linking the phenotype of XLF deficiency to an inability to hold XRCC4-LigaseIV to DNA ends, we began investigating how this activity impacts end ligation during NHEJ. To address this question, we generated a 305 bp DNA substrate containing 2 bp complementary overhangs known to be readily ligated in vitro by a complex of Ku and XRCC4-LigaseIV [6]. We then asked whether dilution of these NHEJ reactions to make factor recruitment and synapsis more difficult reveals a requirement for XLF. We found that as previously described, a complex of Ku and XRCC4-LigaseIV efficiently ligates the DNA substrate when NHEJ reactions were conducted at an end concentration of 50 nM (Figure 4.2a, lanes 3). However, 5 fold dilution of these reactions (maintaining the protein to DNA ratios) abolishes this ligation and addition of either DNA-PKcs (a factor known to stimulate end alignment [3]) or XLF is required for rescue (Figure 4.2a, compare lanes 3, 6, and 7). The ability of XLF to mimic DNA-PKcs’s stimulation of ligation suggests that XLF also functions during synapsis. However, unlike DNA-PKcs, exclusion of the pre-incubation step of these ligation reactions reduces XLF’s ability to
rescue XRCC4-LigaseIV activity (Figure 4.2a, compare lanes 7 and 11). Furthermore, under these conditions inclusion of both DNA-PKcs and XLF has a multiplicative effect on in vitro NHEJ (Figure 4.2a lanes 10-12) suggesting that the mechanisms by which DNA-PKcs and XLF stimulate ligation are non-overlapping and may function in concert. Supporting this, when we attempt to form synaptic complexes under extremely dilute conditions (end concentrations of 0.5 nM), XLF-mediated ligation is reduced 5 fold where DNA-PKcs-mediated ligation is relatively unaffected (Figure 4.2b, compare lanes 2 to 4 and 5 to 7). XLF does stimulate ligation however when synaptic complexes are formed prior to dilution. This indicates XLF may stabilize synaptic complexes once they are formed (Figure 4.2b compare lanes 5 to 6). We therefore propose that while DNA-PKcs functions to actively align DNA ends, XLF functions to hold them together.

**XLF stabilizes the synaptic complex**

To directly test whether XLF increases the stability of a synaptic complex, we designed an experiment utilizing two DNA substrates: the 305 bp DNA described earlier and a second similar 295 bp DNA. Each of these DNAs was individually incubated with Ku, XRCC4-LigaseIV, and either XLF, DNA-PKcs, or both to pre-form synaptic complexes. After 30 minutes, the pre-formed complexes where mixed, incubated again for increasing time, and then ligation was initiated by addition of ATP and MgCl2. Ligation products were digested with a restriction enzyme that cuts internally in the two DNA substrates (HinfI) to resolve concatemers into head to head, head to tail, and tail to tail ligation products. Three potential head to head products exist: two are direct ligations of the pre-formed synaptic complexes (called homo-products) and the third arises from dissociation of the pre-formed complexes followed by re-association between DNA ends.
of the 295 bp and 305 bp DNA substrates (termed hetero-product) (Figure 4.3a). Formation of this hetero-product is thus a measure of synaptic complex stability and allows calculation of a synaptic complex half life.

Initially, all reactions display commitment to the original synaptic complex as indicated by the under-representation of the hetero-product seen when the two DNA substrates were mixed prior to complex formation (Figure 4.3b compare lanes 2 to 3, 7 to 8, and 12 to 13). Reactions containing Ku, XRCC4-LigaseIV, and DNA-PKcs however begin to accumulate hetero-product within 15 minutes of complex mixing (Figure 4.3b lane 4) even though the overall ligation of these reactions is similar to the corresponding XLF containing reactions. Ultimately, a synaptic complex consisting of Ku, XRCC4-LigaseIV, and DNA-PKcs maintains a half life of approximately 60 minutes (Figure 4.3c). XLF containing complexes are longer lived producing half lifes of approximately 3 hours (Figure 4.3c) indicating XLF functions in synaptic complex maintenance.

**XLF stimulates mismatch ligation**

Increased stability of synaptic complexes has several potential benefits. Depending upon the source of the damage, the end structures of chromosome double strand breaks can vary greatly producing a range of ligation difficulty [127]. Aligned overhangs containing mismatches, gaps, and nucleotide damages often utilize the activity of processing factors to make ends amenable to ligation [12, 15, 16, 128]. The use of end processing however requires XRCC4-LigaseIV to release the DNA ends in the synaptic complex. This re-organization increases the chances of synaptic complex dissociation and can potentially lead to DNA translocations or deletion [64]. Thus, XLF’s clamping of XRCC4-LigaseIV to DNA ends may provide a means to stabilize the synaptic
complex while processing occurs and thereby limit the risk of genomic instability. Further supporting this idea, XRCC4 is a known scaffold for the NHEJ processing factors, pol λ, pol μ, human Polynucleotide Kinase, and Aprataxin [13, 62, 129-131]. Thus, increased residence of XRCC4 at DNA ends would seem to increase the effectiveness of end processing. We tested the ability of XLF to stimulate ligation of non-complementary DNA ends by either mismatched joining or via pol μ dependent end processing. As previously reported, we find XLF significantly increases the ability of XRCC4-LigaseIV to join mismatched overhangs; however, the efficiency of this reaction is 5 to 10 fold less than pol μ mediated ligation (Figure 4.4 compare lanes 3 and 4 to lanes 5 and 6). XLF also surprisingly fails to stimulate pol μ mediated ligation. The ligation products of reactions containing both XLF and pol μ resemble a superposition of reactions containing XLF and pol μ individually suggesting that XLF functions independently of end processing (Figure 4.4, lanes 7-10). This independence was confirmed as preliminary sequence analysis of ligation products resulting from reactions containing both XLF and pol μ indicate that XLF never stimulates processing mediated ligation. Furthermore, high XLF concentrations appear to exclude processing (Table 4.1).

DISCUSSION

Based upon its stimulation of low fidelity mismatch ligation and lack of involvement during processing mediated ligation, we propose XLF’s clamping of XRCC4-LigaseIV at DNA ends establishes a subpathway within NHEJ designed to ensure ligation of difficult DNA ends (Figure 4.5). As shown here, and previously, XLF
has relatively little effect on in vitro ligation of complementary overhangs, but greatly stimulates ligation of mismatched and blunt ends [65, 66]. We suggest that because complementary ends are easily aligned and ligated due to base pairing, these ends require only transient interaction with XRCC4-LigaseIV and consequently XLF is dispensable.

Contrastingly, increased residence of XRCC4-LigaseIV at complex DNA ends (possibly established through XLF-mediated filament formation) potentially allows LigaseIV to sample multiple alignments of the 3’OH and 5’PO4 of opposing ends until it finds a geometry conducive to ligation. Thus with the aid of XLF, damaged ends could be directly ligated incorporating the damaged nucleotides but eliminating the risk of DNA translocation and deletion associated with the double strand break (Figure 4.5). Opposing this benefit, low fidelity NHEJ has a significant danger. Our preliminary data suggests specific types of end structures, particularly gaps in the aligned DNA ends or the presence of 5’dRP residues, may prohibit direct ligation by low fidelity NHEJ. Attempts to ligate such ends without the aid of end processing appear to uncouple DNA-LigaseIV’s ligation mechanism. This process results in the adenylation of the DNA substrate, but failure to direct the nucleophilic attack needed to re-establish the phosphodiester backbone. Adenylated DNA ends are toxic to ligation [131]. Therefore, inappropriate attempts to utilize XLF-mediated ligation may further complicate double stand break repair. These situations may require a means to reset the NHEJ synaptic complex potentially through removal of adenylates by Aprataxin [132] and phosphorylation of XRCC4 by DNA-PKcs [133]. After any adenylates are removed and the XLF stabilized synaptic complex is dissociated, a new complex can form allowing processing based repair. Ultimately, the ability to switch between these two
complementary pathways provides would provide an effective means to quickly avoid DNA translocation while maximizing sequence retention.

ACKNOWLEDGEMENTS

Steven Roberts is responsible for all the data and writing within this chapter.
Table 4.1: Sequence analysis of pol μ and XLF mediated junctions. Percent Activity is relative to reactions containing only pol μ. Junction sequences are represented in three columns. Column 3 contains the right side DNA end, Column 5 contains the left side DNA end, and Column 4 contains inserted nucleotides. Number of sequences recovered is indicated in parentheses in column 4. Column 6 indicates whether synthesis-mediated or direct ligation occurred. 4 bp deletions fit neither pathway and likely represent inappropriately digested substrate.
Figure 4.1: Interaction between XLF and other core NHEJ factors. (a) Position of R57G (R57 in blue) mutation in XLF (orange) (PDB 2R9A [122]). R57 resides in α-helix 2 of XLF, a structural element predicted to support polymerization of an XLF, XRCC4 (blue)-LigaseIV (yellow) filament (image adapted from [122]). (b) 10 nM Cy3 labeled oligoduplex was incubated with 2 nM Ku, 10 nM XRCC4-LigaseIV (XL), and either 10 nM wild type or R57G mutant XLF as indicated. NHEJ complexes were formed for 30 minutes and subsequently separated by electrophoresis.
Figure 4.2: In vitro NHEJ of complementary substrates. (a) 25 nM radiolabeled DNA substrate (50 nM ends) was incubated with 100 nM Ku and 50 nM XRCC4-LigaseIV (XL) (lanes 1-3) or 5 nM substrate (10 nM ends) was incubated with 20 nM Ku, 10 nM XL, 20 nM DNA-PKcs (Cs), and 40 nM XLF (lanes 4-12) as indicated. Room temperature pre-incubation was omitted in lanes 9-12. Substrate and ligation products are indicated with S and P, respectively. (b) Ligation reactions containing 5 nM substrate and 20 nM Ku, 10 nM XL, 20 nM Cs, and 40 nM XLF (as indicated) were diluted 10 fold before (B) or after (A) a 30 minute room temperature pre-incubation. Ligation in (a) and (b) was then allowed for 5 minutes. Ligation efficiencies are indicated below each image.
Figure 4.3: Synaptic Maintenance. (a) Synaptic complexes containing 20 nM Ku, 10 nM XLF, and 20 nM Cs, 40 nM XLF, or both were formed on either 5 nM 295 bp or 305 bp complementary DNA substrate. The corresponding 295 bp and 305 bp complexes were mixed for increasing time before initiation of ligation. Synaptic maintenance is measured by dissociation of pre-formed synaptic complexes (homo-products) and subsequent formation of new synaptic complexes between a 295 bp substrate and a 305 bp substrate (hetero-product). (b) The synaptic complexes described in (a) were mixed for increasing time (3, 5, 10, 30 minutes) before ligation for 5 minutes. Ligation products were digested before homo-product and hetero-product head to head junctions were separated by formamide-Urea electrophoresis. Pre indicates mixing of 295 and 305 bp substrates prior to synaptic complex formation. Sub indicates a loading control. (c) Quantification of hetero-product formation in (b). Time mixing is the total time of complex mixing and ligation.
Figure 4.4: Mismatch Ligation by NHEJ. 5 nM TT mismatch substrate (10 nM ends) was incubated with 25 nM Ku, 50 nM XL, 5 or 25 nM pol μ, and 10 or 50 nM XLF as indicated. DNA-protein complexes were formed for 10 minutes before a 5 minute ligation. Ligation efficiency is presented below the image.
Figure 4.5: Model of XLF-mediated Low Fidelity Ligation. DNA damage results in difficult DNA ends (containing oxidized bases or mismatches) which are aligned by Ku and DNA-PKcs. Recruitment of XRCC4-LigaseIV and associated processing factors (right path) results in end processing to produce complementary ends before high fidelity ligation. Alternatively, when recruited with XLF, XRCC4-LigaseIV (left path) attempts direct ligation. If successful, mismatches or damaged nucleotides are incorporated in a low fidelity ligation product. Failed ligation results in adenylation of the DNA end. Aprataxin is needed to remove the adenylate before ligation can be attempted again. Low fidelity ligation may also be aborted by phosphorylation of XRCC4 resulting in dissociation of the synaptic complex, allowing a new processing based complex to form.
To avoid cell death, all chromosome double strand breaks must be rejoined. Nonhomologous end joining however faces a variety of obstacles during repair of these breaks. The location and causative agent of the damage can produce DNA ends that are obscured by DNA binding proteins, have flanking nucleotide damage, or contain mismatches when re-aligned. Obscured, damaged, or non-complementary ends each inhibit end ligation by XRCC4-LigaseIV [89, 127, 134] and thus additional NHEJ activities are required to overcome these obstacles. I have demonstrated that specific NHEJ factors have the potential to participate in each of these tasks. Ku can to serve as a wedge to remove protein obstructions from DNA ends, thereby providing a means to increase XRCC4-LigaseIV’s end access [107]. Furthermore, Ku is an efficient AP lyase capable of excising end-associated abasic sites that block ligation. While end processing provides the potential to conservatively repair non-complementary ends, XLF/Cernunnos increases XRCC4-LigaseIV’s ability to ligate mismatched ends. NHEJ employs these activities and a number of additional factors to confront a broad spectrum of end contexts and thereby ensure that even extremely complex breaks are repaired. Recruitment of these activities greatly expands the scope of the repair machinery, transforming this “simple” DSB repair pathway into a very flexible and sophisticated process.
Chromatin Contexts

The hierarchal packaging of DNA into chromatin presents the initial obstruction to end ligation as it has the potential to disrupt end recognition. The restrictive nature of chromatin has lead several groups to hypothesize that double strand break repair in general is inhibited by chromatin structure and that repair pathways must utilize specific remodeling activities to relieve this inhibition (reviewed in [135]). While the respective roles of chromatin remodelers Rad54 and Ino80 in homology searching [136, 137] and histone eviction preceding end resection [105, 138] during homologous recombination have been well documented, chromatin remodeling during NHEJ is less well-understood.

Like homologous recombination, the need for chromatin remodeling in NHEJ is clear as even small repair complexes containing only Ku and XRCC4-LigaseIV require at least 50 bp of DNA for ligation [97]. Furthermore, XRCC4-LigaseIV activity is readily inhibited by the presence of proteins obstructing DNA ends [89]. Yet, the mechanism that provides XRCC4-LigaseIV the end access it needs has remained elusive. The yeast ATP-dependent chromatin remodelers Ino80 and RSC have each been independently implicated in NHEJ as their impairment results in cell sensitivity to double strand breaks that cannot be repaired by homologous recombination. The kinetics of their recruitment however raises questions as to their general involvement in end recognition. Specifically, recruitment of RSC to double strand breaks is dependent on the presence of the NHEJ factor Ku and precedes that of Ino80 [135, 139-141]. Thus, Ku has already completed the end recognition step of NHEJ before either of these ATP-dependent chromatin remodeling factors arrives at the break. My results are consistent with these findings as in vitro Ku readily binds a variety of protein-occluded ends without the need for prior
chromatin remodeling. Interestingly, upon binding Ku also evicts weakly bound DNA binding proteins (i.e. histone H1) and peels DNA from strongly positioned nucleosomes. These results suggest Ku itself may establish the end access needed for ligation instead of utilizing more disruptive ATP-dependent remodeling complexes (Figure 5.1). Precedent exists for cellular NHEJ occurring without large changes in nucleosome positioning. NHEJ is often responsible for the ligation of uncapped telomeres, but ligation of these DNA ends occurs without any detectible changes in chromatin structure [142]. In this case, break recognition and limited, localized remodeling by Ku may be sufficient to allow ligation to occur.

My work also indicates that Ku may not always be sufficient to remove protein obstructions from DNA ends. Ku is unable to access the DNA ends bound by the RAG proteins in a signal end complex. Therefore, an additional activity (potentially protein modification or proteolysis) must remove this obstruction before these ends are ligated into a signal junction. By analogy, Ku alone may not always be able to provide sufficient end access in the presence of unusually difficult nucleosome obstructions. This situation would potentially require the additional activity of the ATP-dependent chromatin remodelers and explain their requirement for NHEJ repair of the breaks described earlier. In this situation, Ku’s ability to load on protein-associated DNA ends may allow it to specifically recruit these remodelers (as shown for RSC) to difficult breaks to generate additional end access (Figure 5.1). While my in vitro data suggests a role for Ku in establishing end access and recruiting additional remodeling activities to chromosome double strand breaks, the identification of a clear separation of function mutant is needed to test these activities in a cellular system.
End Structure

After a chromosome double strand break is recognized and any end obscuring proteins are removed, XRCC4-LigaseIV is confronted with a second level of obstacles: the structure of the broken DNA ends. Chromosome double strand breaks are often associated with a broad spectrum of flanking nucleotide damages and mismatched overhangs that each inhibit XRCC4-LigaseIV activity [127, 134]. The causative agent of the damage determines the break’s specific end structure. Breaks generated by aborted base excision repair and during class switch recombination are associated with 5’dRP residues [42], whereas breaks produced during V(D)J recombination often contain mismatched overhangs when aligned [121]. Ionizing radiation however generates breaks with a variety of associated nucleotide damages incompatible with repair, including, 8-oxoguanine, abasic sites, 3’phosphoglycolates, and 5’hydroxyls [29, 30]. Consequently, NHEJ utilizes a similarly broad repertoire of processing factors to generate compatible ends suitable for ligation. For each of the incompatibilities described above, at least one and often several NHEJ factors have been implicated in the incompatibility’s removal or repair.

My results extend our understanding of the scope of NHEJ processing by explaining how this repair pathway copes with a specific form of end associated nucleotide damage: an abasic site. As previously mentioned, abasic sites commonly flank double strands breaks generated by ionizing radiation, aborted base excision repair, and during class switch recombination [29, 42]. In vitro and in cells, NHEJ utilizes an AP lyase to precisely excise abasic sites from DNA ends to allow efficient ligation. My work implicates the NHEJ core factor, Ku, as the source of this lyase activity. Ku
efficiently cleaves 3’ of end associated abasic sites and is the predominant protein in both human and hamster cell extracts that cross-links with NaBH₄ to these substrates. Ku therefore likely functions to remove abasic sites at DNA ends. Identification of the lyase active site lysine in Ku will be important to confirm that Ku is the lyase active at double strand breaks in cells.

With ongoing research, the manner in which NHEJ handles many end associated nucleotide damages is becoming increasingly clear. This double strand break repair pathway employs specific processing factors to remove abasic sites and 3’ phosphoglycolates [60, 61], convert 5’hydroxyls to 5’ phosphates [62, 143], bypass 8-oxoguanine [63], and fill gaps in aligned overhangs [15, 113] (Figure 5.2). While these activities may facilitate end ligation, their use specifically requires XRCC4-LigaseIV to release the broken DNA ends [64, 127, 144]. This re-organization of the synaptic complex threatens it integrity and thus increases the chances of DNA translocations and deletions. My research and that of others indicate that NHEJ may complement processing based joining with a low fidelity, XLF/Cernunnos dependent pathway to limit the risk of genomic instability. XLF has been shown to increase ligation of mismatched overhangs [65, 66] suggesting the protein may facilitate the ligation of difficult double strand breaks. My work indicates that XLF stimulates mismatch ligation by clamping XRCC4-LigaseIV to DNA ends and stabilizing the synaptic complex. Furthermore, inclusion of XLF in in vitro ligation reactions containing substrates traditionally viewed as polymerase μ dependent can exclude the processing based ligation in favor of mismatch ligation. Thus, XLF appears to support a low fidelity NHEJ pathway. Genetic analysis of XLF mutants supports this idea as XLF appears to function within a subset of
total double strand breaks. While XLF deficient human patients display a RS-SCID phenotype characteristic of core NHEJ factor deficiencies [7, 26, 27], closer investigation of the root cause of this phenotype indicates that the immunodeficiency seen in these individuals is not due to a defect in V(D)J recombination (as is the case for other core NHEJ factors) but instead an inability to resolve class switch recombination breaks [124]. Since radiation induced breaks and the breaks generated during class switch recombination are each associated with flanking nucleotide damage, XLF may be involved in resolution of these complex DNA ends.

Based upon my studies of Ku’s lyase activity and XLF’s stimulation of mismatch repair as well as previously published work on processing during NHEJ and XLF modification of XRCC4-LigaseIV activity, I propose a model of NHEJ consisting of two complementary paths to accommodate complex double strand breaks. The first pathway involves processing of flanking nucleotide damages and mismatched overhangs to produce complementary overhangs and permit ligation that conserves genetic information. Alternatively, XLF can clamp XRCC4-LigaseIV to the DNA ends and exclude end processing. The clamping of XRCC4-LigaseIV to DNA ends stabilizes the synaptic complex and allows LigaseIV to sample many end geometries until one suitable for ligation is found. Ultimately, the combined use of these two pathways allows NHEJ to maintain the maximum amount of genetic information while ensuring that all breaks are repaired regardless of their complexity.
Coordination of NHEJ factors

The recruitment of such a large collection of factors (Figure 5.2) to a break so that generation of end access, end processing, and ligation can take place presents NHEJ with a potential problem. Each of these factors is potentially recruited to the same double strand break and consequently, their roles within NHEJ must be tightly coordinated to allow each factor appropriate access to the DNA end. Further complicating this situation, many of the processing factors (e.g. Aprataxin, PNK, and APLF) are recruited to breaks by binding overlapping regions of XRCC4 [130, 145, 146]. Thus, the way NHEJ regulates the recruitment of these factors, their access to the DNA ends, and whether to forgo processing for mismatch joining is of current interest.

A temporal organization likely explains the coordination of events occurring prior to synaptic complex formation. As Ku is the first NHEJ protein to the break and subsequently nucleates formation of the synaptic complex, this protein contends with few other activities requiring end access. Thus, Ku’s removal of end-associated abasic sites likely occurs before or during recruitment of the rest of the NHEJ machinery. The chromatin context of the break further favors early implementation of Ku mediated activities (end recognition, chromatin remodeling, and AP lyase activity) as other end-associated proteins exclude the rest of the NHEJ factors from accessing the break. Consequently, Ku or an ATP-dependent remodeler must generate end access at the break before XRCC4-LigaseIV and any of its associated factors can be recruited. DNA-PKcs provides a second level of organization between Ku-mediated processes and other NHEJ activities as the unphosphorylated form of this protein covers the broken DNA ends and excludes end access to additional processing factors [147]. This mode of binding allows
Ku based activities to continue but inhibits all other processing events. Autophosphorylation of DNA-PKcs is required to relieve this inhibition and allow additional NHEJ factors into the synaptic complex [147].

Once the synaptic complex is formed and all necessary processing factors are recruited, the coordination of end processing and ligation becomes less obvious. Phosphorylation of other NHEJ factors may fine tune the regulation of end access by controlling the specific recruitment of needed processing factors. Multiple NHEJ processing factors are each recruited to phosphorylation sites on XRCC4. Specifically, Aprataxin, PNK, and the putative NHEJ nuclease APLF are all recruited to phosphorylated XRCC4 via forkhead associated domains [130, 145, 146]. Differential recruitment of the individual processing factors could thus be controlled by the specific kinase responsible for the XRCC4 phosphorylation as well as subtle structural differences in the processing factors’ protein-protein interaction domains [148]. Additional DNA-PKcs phosphorylation sites in XRCC4-LigaseIV, Ku, and XLF may also have the potential to modulate end processing and ligation. Phosphorylation of XRCC4-LigaseIV is known to decrease the complex’s affinity for DNA and could thus alter its ability to serve as a scaffold as well as function within a stable synaptic complex [133]. The effects of DNA-PKcs phosphorylation of Ku, XRCC4-LigaseIV, and XLF are likely subtle, however, as to this date mutation of any of these sites has failed to produce a significant biological NHEJ defect [149-151].

CONCLUDING REMARKS

The work discussed within this thesis expands the current understanding of the complexity of nonhomologous end joining. While often described as a “simple” method
for repair of chromosome double strand breaks, my work and that of others indicates that NHEJ employs a surprising degree of flexibility in the preparation of DNA ends for ligation. Implication of Ku and specific ATP-dependent chromatin remodelers in the generation of end access has provided initial insights as to how this repair pathway overcomes the inhibitory nature of chromatin to locate double strand breaks and prepare DNA ends for subsequent repair. Similarly, identification of new and unique processing activities, such as Ku’s lyase activity, as well as the description XLF’s role in mismatch ligation have suggested new mechanisms by which NHEJ can circumvent the specialized obstacles inherent in complex DNA ends. As our knowledge of these processes and their coordination continues to expand, it is becoming increasingly clear that this simple system for repair chromosome double strand breaks is a quite sophisticated.
Figure 5.1: Removal of protein-obstructions at DSBs in cells. Ku, nucleosomes, linker histones and other linker-associated factors, additional NHEJ components (e.g. DNA-PKcs, XRCC4-Ligase IV), and ATP-dependent chromatin remodelers depicted as noted.
Figure 5.2: Complexity of NHEJ. The NHEJ core factors (grey) recruit many additional factors involved in overcoming the obstacles of obscured and incompatible DNA ends. Polymerases (blue) are recruited to LigaseIV for gap filling. Factors involved in removing or repairing end associated nucleotide damage (green) are often recruited to XRCC4. Artemis (yellow) is recruited to DNA-PKcs for hairpin opening, flap removal, as well as removal of damaged bases. XLF (red) binds mainly to XRCC4 and modulates LigaseIV's ability to ligate non-complementary ends. Ku is unique, displaying multiple activities involved in removing nucleotide damage (abasic sites) and removal of protein-obstructions (either by itself or recruited ATP-dependent remodelers (white, i.e. RSC)).
REFERENCES


88


150. Lee, K.J., et al., Identification of DNA-PKcs phosphorylation sites in XRCC4 and effects of mutations at these sites on DNA end joining in a cell-free system. DNA Repair (Amst), 2004. 3(3): p. 267-76.