DNA LIGASE IV STRUCTURALLY SUPPORTS END JOINING REPAIR OF DNA DOUBLE STRAND BREAKS.

By

Noah J. Goff

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Cell and Molecular Biology – Doctor of Philosophy

2024

ABSTRACT

DNA Double strand breaks (DSBs) are highly genotoxic lesions induced by external agents (ionizing radiation or chemotherapeutic drugs), cellular processes (DNA replication defects, reactive oxygen species, recombination intermediates in B/T cell development), or through increasingly prevalent gene editing techniques (CRISPR/Cas9). Mammals have evolved two major pathways for repairing DSBs: Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ). While NHEJ is frequently referred to as the "error prone" pathway, recent biochemical and *in vitro* single-molecule imaging studies have revealed mechanisms to protect ends from excessive end-processing through a two-stage synaptic model. First, DNA ends are synapsed by NHEJ factors in a long-range complex (LRC) with ends held ~115 Å apart. The LRC recruits downstream factors and only permits highly regulated end processing of bulky adducts. Second—following the recruitment of the NHEJ-associated DNA Ligase IV (L4)—ends are brought into direct proximity for final end-processing and ligation. In this dissertation, I report that catalytically inactive L4 promotes significant amounts of end joining in cell models, supporting the recent two-stage synaptic model for NHEJ. Furthermore, I characterize repair products from cells expressing catalytically inactive L4 showing that repair is significantly more mutagenic than in cells expressing active L4. Finally, I identify multiple interfaces in L4's DNA binding domain critical for maintaining promoting repair, providing insight into how L4 structurally supports synapsis of ends.

This dissertation is dedicated to my parents. It's impossible to overstate the significance of your support over the past 5 years.

ACKNOWLEDGMENTS

The past five years of research and time spent writing for my PhD would not have been possible without the generous and thoughtful support from dozens of people. It absolutely takes a village and I'm extremely grateful to the amazing community of people who have helped me over the past ~5 years. Special thanks to my committee for your valuable advice and well-thought questions throughout my time here at MSU; Dr. Daniel Vocelle from the MSU Flow Cytometry core for your dedication to teaching and technical know-how; the friends I've made through the BMS and CMB programs for their willingness to discuss science at any time of day (preferably over a beer); Dr. Jens Schmidt's lab for being so accommodating with microscope time (and room in their biosafety cabinets); the CMB program—specifically Dr. Margaret Petroff for being a fantastic graduate director and Alaina Burghardt for being the best program coordinator at the university; Emily Durocher for her unerring ability to make me smile after a long day; and finally my family for being there since day one, I'm so happy with how this journey has turned out and It's impossible to quantify how much having you nearby has meant along the way (not to mention all of the free food).

TABLE OF CONTENTS

Chapter 1: Evolving models of non-homologous end joining.

By

Noah J. Goff

Introduction and literature review

INTRODUCTION

DNA double strand breaks (DSBs) are extremely genotoxic lesions generated by a host of cellular processes and exogenous agents. Repair of DSBs is essential, and consequences of unrepaired breaks can impair core nuclear functions (transcription, DNA replication), stall cell cycle progression by triggering DNA damage checkpoints (DSB forming drugs form the backbone of many chemotherapeutic treatments), compromise genomic stability through aberrant genomic translocations, and ultimately an unrepaired break can lead to programmed cell death. Furthermore, failures during repair processes can further compromise genomic integrity, including short insertions/deletions, copy number variations, large intrachromosomal deletions, and inappropriate pairing of ends can exacerbate large chromosomal deletions. My work has focused on DNA end joining pathways and the structural role that DNA Ligase IV (L4) plays in supporting end joining through the canonical **n**on-**h**omologous **e**nd **j**oining pathway (NHEJ) and how loss of L4 catalytic activity promotes mutagenic **alt**ernative **e**nd **j**oining (alt-EJ).

DSB Repair pathways aim to accurately resolve breaks. DSBs are generally repaired through two mechanisms: NHEJ and **h**omologous **r**ecombination (HR). Briefly (DSB repair pathways are discussed in greater detail later in this chapter), HR utilizes a nearby homologous DNA sequence to perform high fidelity templated repair. NHEJ serves to resolve DSB repair through direct ligation of two free DNA ends (hence “end joining”). First, a DNA end joining pathway must sense free DNA ends, then generate a ligation substrate by modifying chemistry at the ends, and eventually catalytically seal the break by ligating both broken phosphodiester backbones. NHEJ solves a core challenge for cells: how to repair breaks that occur without a nearby homologous template (which is only present immediately following DNA replication). The flexibility of NHEJ (it can act on DNA ends through nearly the entire cell cycle) presents several challenges; namely, how to conserve sequence identity (preventing indels) and to conserve larger chromosomal structures (preventing large genomic translocations). Recent work has highlighted the remarkable fidelity of NHEJ and has sought to identify mechanisms to minimize errors generated by repair.

Conserved NHEJ structural factors promote repair. A structural heterodimer formed by Ku70 and Ku80 serves both as the key initial sensor of DNA damage as well as an anchor point for downstream factor recruitment. Early studies into Ku revealed that each factor exists

independently of each other but rapidly bind to DSBs and promote recruitment of other factors (citations). The two “C” shaped proteins clamp onto a double strand break, forming a donut shape that mediates protein-protein interactions with DNA-PKcs (forming the DNA-PK holoenzyme) and other repair factors (frequently though interactions between Ku binding motifs (KBM) and BRCT domains on other repair factors). In human cell strains, loss of either Ku gene is lethal, resulting in telomere fusions and cell death (1, 2).

Interestingly, the heterodimeric ring formed by Ku70/Ku80 is extremely stable and allowed to slide upstream of the break site—like beads on a string. As such, several mechanisms exist to prevent overloading of Ku onto a DNA an end and there’s significant discussion and ongoing research into how these tightly-bound DNA repair donuts are removed from repaired ends (3, 4). Several structural features of Ku have been well characterized, including the Ku80 C-Terminal Domain (in particular the final 14 residues) as being critical for functional interactions with DNA-PKcs to promote end joining (5–7).

X ray cross complimenting protein 4 (XRCC4) serves both as a key structural factor in supporting NHEJ synapsis and as the obligate binding partner of Ligase IV (L4). Two molecules of XRCC4 and one molecule of L4 form the ligation complex LX4, together a core NHEJ factor that is necessary for robust end joining. Each XRCC4 homodimer is shaped with an N-terminal globular head and long C-terminal alpha-helix stalk that mediates its interaction with L4 (8, 9). When homodimerized, the two XRCC4 head domains pair up with the C-terminal alpha-helices wound around each other 1-2 times (Figures 1, 2, 4). Loss of XRCC4 or L4 results in extreme hypersensitivity to DSB inducing agents (both as a result of radiation and drug induced damage) (10–16). Additionally, XRCC4 possesses a disordered c-terminal “tail” region that has been shown to transiently interact with L4’s DBD and directly with DNA (17). As part of the LX4 complex, XRCC4 facilitates interactions with XLF and Ku, providing support for end synapsis observed in recent Cryo-EM structures (18, 19).

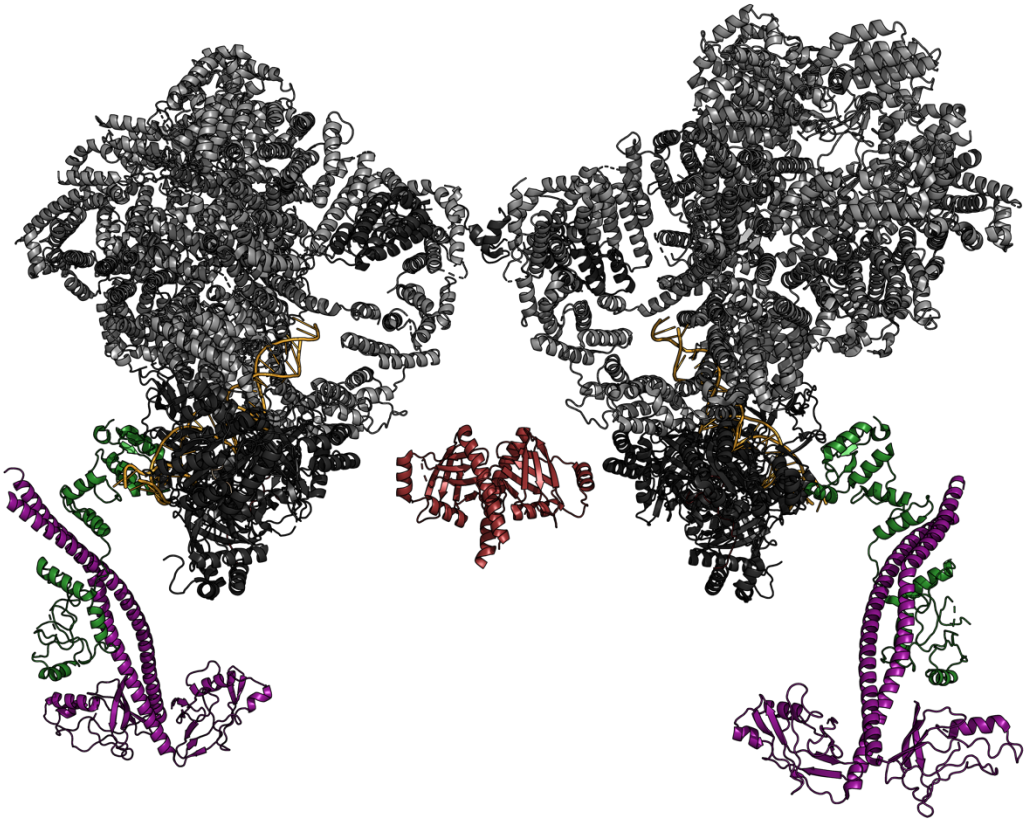


Figure 1: LX4 Has been observed in the Ku80-Mediated Long-Range Complex when incubated with PAXX. PDB 8BHY: dimeric structure of DNA-PKcs (grey), Ku70/Ku80 (grey/black), L4 (green), XRCC4 (purple), XLF (orange), and PAXX (red) work together to synapse DNA (orange) in a long-range synapse. Here, the LX4 complex are positioned away from the break, with interactions between DNA-PK and PAXX, along with *trans* interactions between Ku80 and DNA-PKcs supporting synapsis.

Additional structural factors XRCC4-like factor (XLF) and paralog of XRCC4 and XLF (PAXX) also work to synapse ends. XLF is a central mediator of synapsis in recent Cryo-EM structures, with the "XLF-mediated dimer" consisting of a group of structures where an XLF homodimer serves as a major contact in bridging the end-bound DNA-PK holoenzyme and LX4 complex (18, 19). The XLF-mediated dimer appears to fill the role of "end protection complex" originally proposed as the primary role of the long-range complex (LRC) proposed by Loparo and colleagues (20, 21). While complete ablation of XLF has variable impacts in different organisms, loss of XLF is remarkably synthetically lethal with a host of other DNA factors, and many studies implicate its role in promoting high-fidelity repair (12, 22–28). While typically not considered a "core" NHEJ factor (PAXX is functionally redundant with XLF), PAXX has recently been observed supporting LRC

via Cryo-EM, including a structure very similar to the Ku80-mediated dimer that includes PAXX and LX4 (providing evidence that the NHEJ dimers may all be capable of recruiting core factors required for transition into the SRC). Despite its apparent redundancy with XLF, PAXX has been identified in a wide range of species including plants (29).

The two core enzymatic NHEJ factors, DNA-PKcs and L4, each play unique structural roles in supporting NHEJ. DNA-PKcs is a large 465 kDa serine/threonine protein kinase that acts as the primary structural regulator of DNA end access. Furthermore, DNA-PKcs autophosphorylation plays a key role in driving conformation changes in NHEJ complexes, with well-studied phosphorylation motifs directly regulating access to DNA ends (11, 30–32). While not present in prokaryotes (many prokaryotes have end-joining with distant orthologs of the Ku and L4 molecules) (33, 34), DNA-PKcs is found in a wide range of eukaryotes with many of its key structural motifs (and autophosphorylation sites) conserved (35).

Mechanism and domains of mammalian DNA ligases: There are three mammalian DNA ligase genes: Ligase I (L1), Ligase III (L3) and Ligase IV (L4). All three genes share a conserved catalytic core consisting of a reactive Nucleotidyl Transferase Domain (NTD) flanked by a DNA-Binding Domain (DBD) and Oligonucleotide-Binding-Fold (OBF). While all three ligases serve DNA repair functions in the nucleus, L3 also provides mitochondrial DNA repair function through an alternate translational start site. This replaces the extreme N-terminus of the protein, substituting the nuclear localization signal for a mitochondrial localization signal (36). Each mammalian ligase is ATP dependent, capable of sealing a single nick in a deoxyribose-phosphate backbone (37–39). The general reaction progresses in three steps: the ligase 1) hydrolyzes ATP to adenylate a conserved lysine residue in NTD active-site, 2) binds to a nick and transfers the AMP moiety from its lysine to the 5’ phosphate, then 3) supports a nucleophilic attack from the remaining 3’ hydroxyl end to the 5’ PO_4-AMP bond resulting in a free AMP, uncharged ligase, and sealed single-strand break (37–41). Notably, the high degree of domain-conservation and architecture across all three DNA ligases implies that these genes share a common evolutionary ancestor (37, 39–41). Each Ligase NTD acts primarily through a conserved catalytic lysine that self-adenylated in the presence of ATP, providing energy to catalyze the break (37, 39).

Ligase IV is the NHEJ-exclusive ligase and well adapted for DSB repair. While Ligases I and III are frequently referred to as redundant, L4 serves exclusively in NHEJ-mediated DSB repair. While each of the mammalian ligases share a conserved mechanism and degree of structural homology in their catalytic cores, L4 has evolved several unique structural features to promote ligating ends from a DSB:

First, L4 has evolved a pair of tandem BRCT (BRCA1 c-terminus) domains that facilitate its interactions with NHEJ proteins. The tandem L4 BRCT domains are connected by a short, flexible linker termed the XRCC4 Interacting Region (XIR) for its role intercalating into the central alpha-helix stalk of an XRCC4-homodimer (37–41). The XIR is both sufficient and necessary to mediate L4's interaction with XRCC4 (of note, L3 also has evolved n-terminal BRCT domains that mediate its interaction with XRCC1). Second, unlike L1/L3, L4 rapidly adenylates following translation, and nearly all the L4 molecules present within the nucleus are prepared to catalyze break repair (37). This readiness could serve as a mechanism to minimize the amount of time required to perform the initial ligation step—converting the DSB into a much less harmful single-stranded break. Interestingly, while we were studying a catalytically inactive L4 (designed by mutating the conserved catalytic lysine), we observed residual ligation activity in *in vitro* joining assays. This residual ligase activity is lost following mutation of additional lysines positioned close to the DNA end, indicating that there is some flexibility within the catalytic core (10). To my knowledge, there have been no reports of alternate lysine adenylation in L1 or L3.

Third, L4 has a unique 8-amino acid loop structure in its DBD that promotes flexibility in DNA-substrate binding, including tolerance of mismatches, aberrant bases, and gaps in one end of the DNA substrate (42). This unique lesion tolerance is another representation of the priority given to double strand break repair—with failure to catalyze repair resulting in significant cell-growth defects (and loss of L4 is generally lethal) (14, 43). Notably in cell line models, loss of L1 or nuclear L3 generally does not result in cell fitness defect or even notable drug-sensitivity phenotypes (15, 44). The two single-stranded ligases are functionally redundant and loss of both is synthetically lethal; in contrast, L4 does not appear to complement DNA repair outside of NHEJ.

DNA-PKcs: A Synopsis Beyond Synapsis. The following sections have been reproduced from a review that I co-wrote with another graduate student, Maria Mikhova, alongside our advisors

Katheryn Meek and Jens Schmidt (45). The reproduced sections were primarily drafted by me and edited as a group. The figure referenced in this section was originally produced by Dr. Meek, originally produced as as “Figure 2” within the review. I’ve also removed chapter numbers, updated citation formats, citation numbers, and abbreviations to be consistent with the styles used in my dissertation.

A two-step model of DNA end synapsis in long and short-range NHEJ complexes. An emerging consensus in the field is a two-stage synaptic model for NHEJ driven by the presence and catalytic activity of DNA-PK. This model was initially proposed by Loparo and colleagues who have extensively employed *X. Laevis* egg extracts to measure NHEJ-mediated synapsis of two DNA ends using single-molecule fluorescence resonance energy transfer (smFRET, simultaneously measuring synapsis and proximity). Graham et al. provided the first evidence and model for distinct long-range and short-range synaptic complexes [LRC, SRC] (46). They proposed a LRC which notably holds DNA ends far apart (>100 Å); these LRCs are short-lived (on the order of several seconds), and only a subset progress to SRCs where ends are brought in direct proximity (as assessed by FRET). Formation of LRCs requires only Ku and DNA-PKcs. Critically, in this study, acceptor fluorescence was visualized independently of donor-mediated FRET allowing observation of synaptic events at distances beyond the range of FRET detection (>100 Å). This study established that progression to the SRC required the presence of Ku, L4, XRCC4, XLF, and the catalytic activity of DNA-PKcs. Interestingly, the catalytic activity of LX4, essential for the final ligation step in NHEJ, was entirely dispensable for SRC formation, supporting an important structural role for LX4 which had been proposed in previous studies (12, 47). Our recent cellular studies confirmed that LX4’s catalytic activity is dispensable for progression from long-range to short-range complexes (10). Moreover, these findings imply that nuclear DNA ligase III (L3) can access DNA ends maintained in NHEJ complexes by the presence of catalytically inactive LX4, a heretofore unappreciated level of cooperativity between mammalian DNA ligases both in living cells (10) and in animal models (14).

LX4 binding to DNA ends drives progression from long-range to short-range complexes. Work from Loparo and colleagues has recently addressed how NHEJ prioritizes ligation (of compatible ends) over end processing, promoting error-free repair. A 3-color imaging system was utilized to

independently assess LX4 binding at each end, the presence or absence of short-range synapsis, and LX4 stoichiometry, independent of its end-binding (48). This study demonstrates that LX4 dynamically binds DNA ends through residues in its DNA Binding Domain (DBD) prior to short-range synapsis, potentially as a sensor for large end adducts that could impede ligation or final end-processing. SRC assembly (as measured by FRET between DNA ends) occurs exactly at the time when a single LX4 complex binds to the two DNA ends. Surprisingly, SRC formation results in eviction of one of the two LX4 complexes that are present prior to short-range synapsis. Immediately at the time of short-range synapsis, loss of one LX4 complex was observed with the second LX4 residing for much longer consistent with the previously reported persistence of short-range complexes. A caveat of this study is that a blunt-ended DNA substrate was utilized that should be readily ligated, although LX4 end-binding progression to a short-range complex was similarly efficient with or without 5' phosphates. An important question remains as to whether the change in LX4 stoichiometry occurs prior to strand ligation. A similarly important issue is determining how LX4's dwell time is impacted if the DNA ends require fill-in synthesis or nucleolytic processing on one or both ends.

An attractive hypothesis supported by robust previous studies (49, 50) would be that the immediate capacity to ligate paired ends of one of the DNA strands would function even if the second strand requires substantial end-processing. Simply put, there is no better way to maintain synapsis of DNA ends than by sealing one strand. The unique characteristic (among vertebrate DNA ligases) of ligase 4 as a single-turnover enzyme (51) is curious, and obviously infers that the ligase 4 molecule that resolves the first strand break does not (necessarily) resolve the second strand break (which could be recognized as a nick or single-stranded DNA gap). This may reflect LX4's important structural role both in recruiting other enzymes, and in promoting stable synapsis. As noted above, emerging studies demonstrate not only that Lig3 can associate with NHEJ factors, but also that when LX4 is inactivated by mutation, Lig3 clearly functions to repair DSBs, in an inactive LX4-dependent manner, in living cells and animals (10, 14).

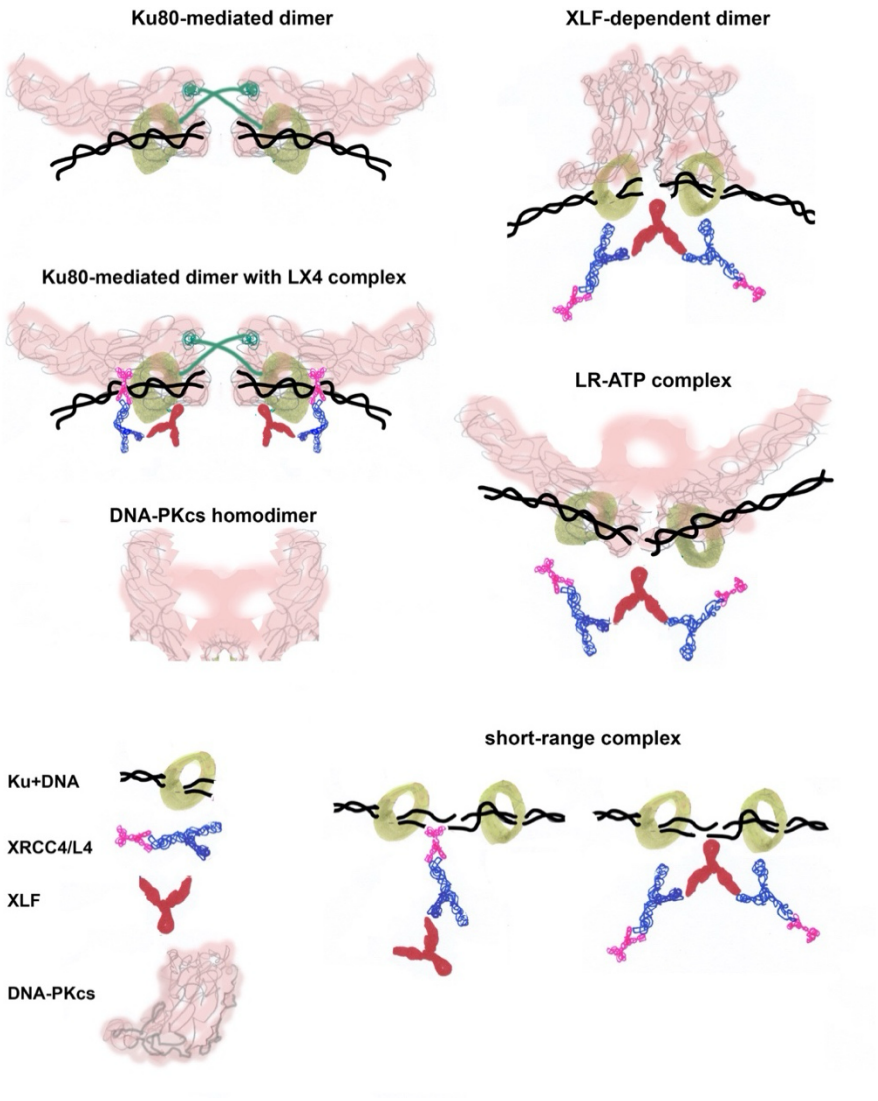


Figure 2: Numerous NHEJ complexes have been visualized by cryo-EM. Cartoons depicting NHEJ long-range complexes (Ku80-medieated and XLF-dependent dimer), long-range ATP (LR-ATP) complex, DNA-PKcs homodimer, and SRCs. DNA-PKcs is colored pink, Ku-green, XLF-red, XRCC4-blue, Lig4 (fuscia) (derived from PDB 6ZHE, PDB 7NFC, PDF 7LT3, PDB 8EZB, PDB 8BYH, PDB 7LSY, PDB 8EZ9).

Cryo-EM structures of NHEJ synaptic complexes. This two-stage synaptic model of NHEJ has been strengthened by recent structural work describing Cryo-EM structures of multimeric NHEJ complexes positioned around breaks. The published structures include descriptions of two major forms of NHEJ complexes that synapse DNA ends ~115Å apart (consistent with long-range synaptic complexes). These dimers are organized around two distinct interfaces: one mediated by a trans interaction between DNA-PKcs and Ku80's extreme C-terminus, and the other dependent

on a centrally placed XLF homodimer interacting with Ku and the LX4 complex, which also involves a large interface between the two DNA-PKcs molecules (Figure 2). These dimers have been termed the Ku80-mediated dimer (or domain-swap dimer), and the XLF-dependent dimer. Initial observations of the Ku80-mediated dimer only included the components of the DNA-PK complex, which would be consistent with the LRC requiring only Ku and DNA-PKcs. However, further studies show that the Ku80-mediated dimer can also contain XLF, PAXX, and LX4, or XLF and LX4, or PAXX and LX4 (22). The XLF-dependent dimer was first observed with two DNA-PK protomers, two LX4 complexes and one XLF homodimer. Recent studies have shown that PAXX can also assemble into this complex (22). Figure 2 presents cartoon depictions of the dimers (for simplicity, the complexes that include PAXX are not represented); structural models of these complexes are available in recent excellent reviews (52, 53)

He and colleagues have shown that in the presence of ATP, the XLF-dependent dimer can progress to a new form, termed long-range ATP (LR-ATP) (54). In this complex, ABCDE phosphorylation has occurred (likely in trans) resulting in a major conformational change: each DNA-PK protomer has a remarkably similar structure to the ABCDE phosphorylated monomeric complex with the DEB dissolved, and the phosphorylated ABCDE sites anchored by K/R residues in M-HEAT. These authors also observed a DNA-PKcs homodimer that shares some structural similarity with homodimers of other PIKKs, ATM and ATR. They hypothesized that this dimer formed after ATP-induced dissociation of DNA-PKcs from Ku/DNA. Another idea proposed was that this homodimer might serve as a DNA-PKcs reservoir. Notably, although these DNA-PKcs dimers were observed without the other components of the XLF-dependent dimer, no complexes consistent with SRC were observed in the cryo-EM experiments with ATP. Cryo-EM structures of NHEJ complexes consistent with SRCs have only been obtained in experiments utilizing Ku, LX4, and XLF with a DNA substrate that facilitates end synapsis (18). The organization of this short-range synaptic complex is remarkably like the organization observed in the XLF-dependent dimer, with a single XLF monomer mediating the synapsis between two Ku-bound DNA ends, and LX4/XLF/LX4 bridge promoting synapsis (Figure 2). Of note, in the short-range structure, although the catalytic domain of only one LX4 is structured and positioned over the perfectly juxtaposed DNA ends, the second LX4 complex is clearly present, which is not consistent with the

stoichiometry discussed above. Stinson et al. suggest that one LX4 complex is released at the time of long-range to short-range transition but noted that another LX4 may associate (48) (thus, two potential SRCs are depicted in Figure 2). Finally, a DNA-PK trimer was observed (not shown), consisting of a central DNA-PK protomer, interacting with another to form an XLF-dependent dimer, and concurrently interacting with a third DNA-PK protomer via the Ku80-mediated domain swap interaction. It is not entirely clear how this trimeric structure would function, but at least two distinct DSBs would be required (55).

Alternative models of NHEJ synapsis. The observation that DNA-PKcs catalytic activity is required for SRC formation in the *X. Laevis* extract model has generated substantial discussion over the role of DNA-PKcs when presenting complete models for NHEJ synapsis. Although the idea that DNA-PK facilitates synapsis was proposed decades ago, originally by Chu and colleagues (56) and supported by others (57), in other elegant smFRET studies using purified human NHEJ components, stable synapsis of DNA ends is not dependent on DNA-PKcs6 (58–60) Still, other experimental systems using purified human proteins, demonstrate a clear role for DNA-PKcs in synapsis (61). In the studies from Lieber and colleagues (58–60), adding DNA-PKcs decreased the total number of synaptic foci. These investigators also examined the role of DNA end chemistry on LX4 mediated synapsis and ligation; their studies showed a strong dependence for 5'phosphates for synapsis in the absence of DNA-PKcs (59). A caveat of this and other smFRET approaches is the inability to differentiate between synapsis-mediated or covalent ligation of the two oligonucleotides. Further *in vitro* single-molecule work (58) provided evidence for a long-range synaptic complex formed from purified human Ku70/80 and LX4. Addition of purified XLF to these reactions induced a transition into a short-range synaptic complex, a consistent observation with their earlier findings of DNA-PKcs-independent short-range complex formation. Thus, additional work will be required to elucidate how distinct NHEJ complexes facilitate repair.

Still, it seems likely that there are complex, cell-wide effects of losing DNA-PKcs beyond its role in synapsing and regulating access to DNA ends in NHEJ. There is ample evidence for a multifaceted role for DNA-PKcs in both directly synapsing ends and participating in the cell-wide DNA damage response; moreover, DNA-PK has been implicated to function in numerous other cellular functions [reviewed in (62)]. One argument suggesting that DNA-PKcs can be dispensable

for NHEJ is that cells lacking DNA-PKcs are notably less sensitive to DSB inducing drugs than isogenic cells lacking other core NHEJ factors (63, 64). However, this conclusion is complicated by the fact that cells lacking DNA-PKcs also lose significant expression of ATM, leading to an impaired apoptotic response (65, 66). In fact, loss of even one allele of ATM substantially rescues the severe phenotype of Lig4 deficient mice (67).

There are also obvious examples of DNA-PKcs-independent NHEJ. Although recent phylogenetic work from Lees-Miller et al. have established that DNA-PKcs is highly represented in many species far removed from vertebrates (35), there are still many species with functional NHEJ pathways that lack DNA-PKcs homologues (for example, some insects, yeast, and bacteria) (68). Additionally, the recently published cryo-EM structure of the putative short-range complex was observed in cryo-EM experiments that lacked DNA-PKcs; thus, short-range complex assembly can occur in the absence of DNA-PKcs (18). Moreover, because of the iterative nature of NHEJ, we must consider that there may be multiple paths that lead to the same outcome: there is more than one way to tie a knot/ligate an end. In sum, while recent studies and technological advancements have provided new insights, substantial gaps in knowledge remain that must be resolved before a full understanding of how NHEJ functions in living cells to rapidly repair DSB while minimizing genome alteration.

Many, but not all NHEJ end-processing steps occur in short-range complexes. Another transformative study from Loparo and colleagues (21) provided compelling biochemical evidence that factors required for the formation of SRCs [Ku, XLF, LX4, and the catalytic activity of DNA-PK] are also essential to enable fill-in synthesis mediated by X-family polymerases [polƛ and polμ, 3’ adduct cleavage mediated by Tdp1, phosphorylation of 5’ hydroxyls by PNKP, and the activity of an unidentified 5’>3’ exonuclease on 5’ flaps (21). Moreover, DNA ends were observed to be protected from end-processing activities in the LRC.

At first, these results might seem at odds with our conclusion that the two long-range complexes promote different aspects of end-processing; this is not the case. The mutational studies suggest that the SRC derives from the XLF-mediated dimer; the enzymatic activities absolutely ascribed to the SRC by Stinson et al., (except for the unidentified 5'>3" nuclease) should all promote fill-in end processing. So, the XLF-dependent dimer promotes fill-in end

processing by promoting progression to the SRC. Part of the argument from Stinson et al that nuclease activity is limited to the SRC was that DNA-PK inhibition blocked nuclease activity, presumably because this blocks progression to the SRC. However, DNA-PK inhibition also blocks the ABCDE phosphorylated form of DNA-PK activation that is requisite to Artemis activation and potentially required for the activation of other nucleases that may function in NHEJ (Mre11, Apollo)(1, 69). Thus, what is less clear is what nucleases (besides Artemis) contribute to NHEJ and in what NHEJ complexes do these nucleases function. This concludes the work adapted from DNA-PK: A synopsis beyond synapsis.

L4's role in dimeric NHEJ complexes: In recent NHEJ cryo-EM structures, the XRCC4-L4 complex has been observed in variations of both classes of NHEJ dimers, with the catalytic core only observed in the SRC and one unpublished version of the Ku80-mediated LRC (4–6). The LX4 complex is seen to interact with Ku70 and Ku80 through XRCC4's head domain and L4's BRCT repeats (Figure 1, 3) (18, 19, 22, 71). In the XLF-mediated dimer, LX4 instead facilitates an interaction with an XLF homodimer spanning the gap between ends, bridging each Ku "donut" (there is also a conformation change in DNA-PKcs to more tightly bind ends with a *trans* PKcs-PKcs interface supporting synapsis in addition to LX4-XLF-LX4, Figure 3). The Ku80-mediated dimer was initially observed in the absence of LX4, however recent studies from Chaplin and colleagues incorporating PAXX and LX4 have since resolved a version of the Ku80-mediated dimer with LX4 bound to DNA-PK, but not supporting the synapse (in contrast to their role in supporting the XLF-mediated dimer). The short-range complex structure contains two LX4 complexes, with one interacting only though its LX4 structural domain and the other directly bound to the break (Figure 4)(18)—notably contradicting recent smFRET work from Loparo and colleagues (48) that indicates one molecule of L4 leaves the break prior to SR synapsis.

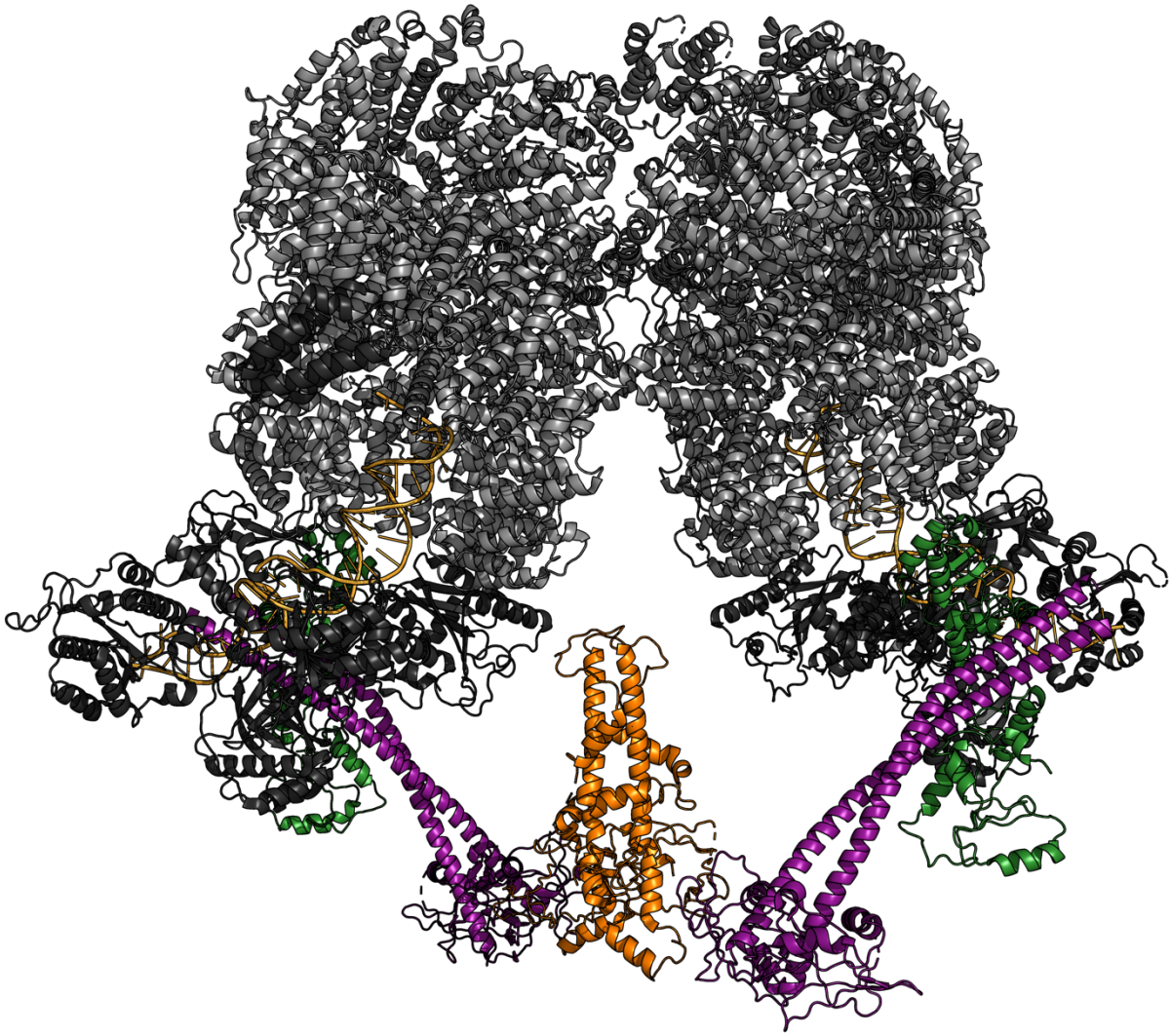


Figure 3: XLF-Mediated Long-Range Complex. PDB 7LSY: a second dimeric structure of DNA-PKcs (grey), Ku70/Ku80 (grey/black), L4 (green), XRCC4 (purple), and XLF (orange), work together to synapse DNA (orange) in a long-range synapse. Here, LX4 synapses ends through interactions with Ku80 and XLF, along with a DNA-PKcs *trans* interaction in the head domain.

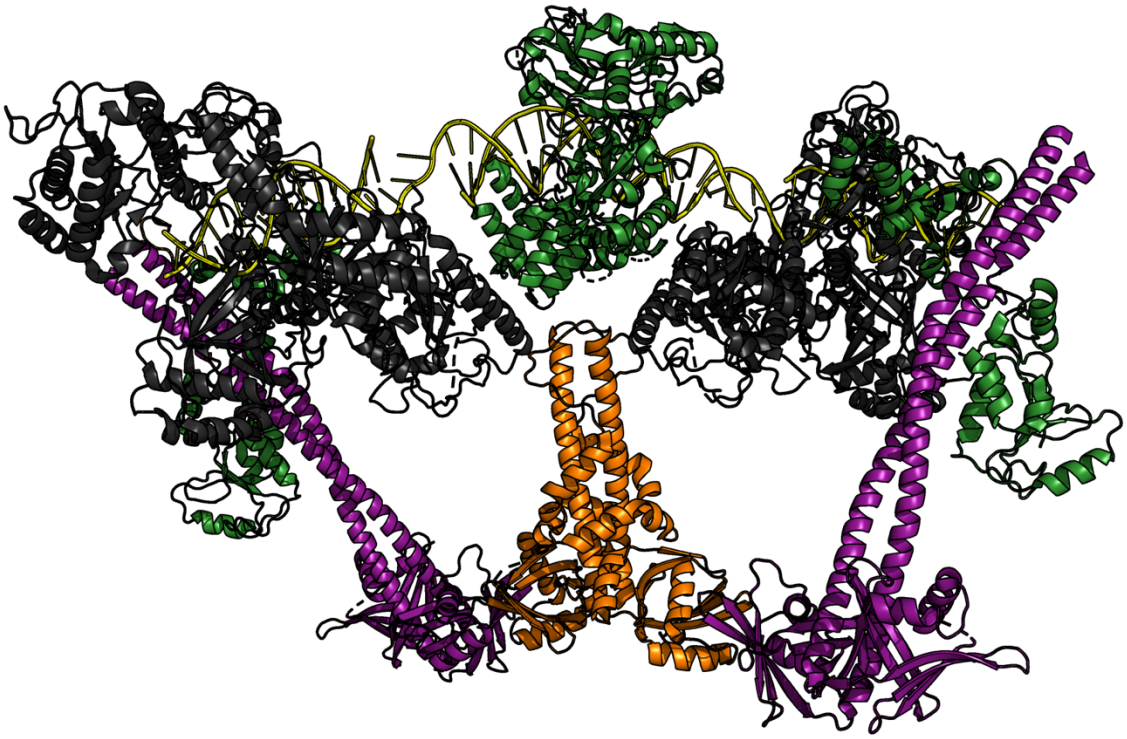


Figure 4: DNA Ligase IV plays a pivotal role in the observed CryoEM Complex. PDB 7LSY: Ku70/Ku80 (Black), L4 (Green), XRCC4 (Purple), XLF (Orange) work together to synapse DNA (yellow) in a ligation-competent state. This structure features two LX4 complexes, with the rightmost L4 catalytic core bound to the DNA end. Only the structural domains (L4 BRCT, XRCC4 homodimer) are observed for the second LX4 complex (left).

Diverse DNA substrates and the cell cycle dictate repair pathway choice: While it's somewhat unconventional to think of DNA repair enzymes in terms of classical enzyme kinetics—since the concentration of a repair factor typically dwarfs the concentration of DSBs—the model works well for considering DNA end chemistry as a key element in pathway choice. The diverse mechanisms that generate DSBs results in a very wide range of substrates—ranging from unique end adducts formed by many chemotherapeutic agents, secondary structures biased near the break, single-ended DSBs formed at failed replication forks, to ends compromised by previous repair steps. To this end, eukaryotic organisms have evolved several mechanisms to sense breaks and engage the correct repair mechanism depending on substrate chemistry and cell cycle status.

Homologous recombination: The other major DSB repair pathway, homologous recombination (HR), serves as a slow, energy intensive, high-fidelity repair pathway that can repair nearly every variation of DSB chemistry (with the notable exception of programmed DSBs generated in

VDJ/CSR recombination). The primary drawback to HR is timing: namely, it's only available in the S/G2 phases of the cell cycle after DNA replication generates a nearby homologous template for repair. This limits repair to actively dividing cells that encounter breaks behind replication forks. Furthermore, instead of directly repairing ends, HR first resects large stretches of DNA to generate long single-stranded ends. This resection (mediated by MRN and Exo1) serves to clear lesions and facilitate homology search prior to synthesis of new DNA—thereby creating a "high fidelity" repaired region (assuming the homologous template matches the original sequence). The resection and homology search are a time-intensive processes, requiring multiple rounds of processing (including resolution of the DNA tetramer holiday junctions) prior to finalizing repair.

The intense processing of ends directed to HR results in intermediates incompatible with NHEJ; as such, the pathways act competitively and there are multiple mechanisms to restrain either pathway depending on context. These mechanisms act 1) to restrain NHEJ and prevent premature ligation (possibly by preventing transition into the SRC)(21) and 2) actively remove NHEJ factors from acting on ends that are targeted towards HR (3, 72). Interestingly, a recent live-cell SMI study indicates the same abortive mechanism engages in cells lacking L4 (73)—implying this mechanism may more broadly engage if NHEJ is incapable of rapidly transitioning ends into the SRC.

Alternative end joining: (alt-EJ) serves as less well-defined collection of factors that's available to join breaks not resolved by either of the "canonical" NHEJ or HR pathways. While the pathway is loosely defined, recent studies (including experiments reported in Chapter 3) suggest that DNA Polymerase Theta (Polθ) is a critical factor essential for many alt-EJ repair events. Polθ is a large and uniquely contains both a highly error-prone polymerase domain and a helicase domain capable of bridging ends utilizing 1-5 bp of terminal microhomology (74). The dual helicase-polymerase function is generally thought to facilitate synthesis through extremely difficult to replicate stretches of DNA (g quadruplexes, other complicated secondary structures) in addition to its role as a backup end-joining factor (74–76).

Of note, Polθ prefers free DNA end substrates with long sections of 3' ssDNA (77, 78) and its end-joining functions likely support joining of failed HR intermediates (where resection can produce long single-stranded ends that are poor substrates for canonical NHEJ). While further

work needs to be done, current literature supports MMEJ—and its dependence on Polθ helicase/polymerase—as a key mediator of joining in the absence of NHEJ.

Alt-EJ serves as an interesting foil to NHEJ, being both lower fidelity and significantly slower. These qualities reflect alt-EJ's typical substrates: rare complex breaks that "fail" out of the two major DSB repair pathways. As such, there is little evolutionary pressure for alt-EJ to have developed mechanisms to respond to many breaks, quickly. This is in stark contrast to NHEJ, where factors have evolved to be both fast, highly regulated, and relatively accurate. It may be more accurate to characterize alt-EJ as a collection of backup mechanisms that may really be secondary functions of many other DNA repair pathways working together to resolve DSBs. Many factors hypothesized to resolve breaks in the absence of NHEJ have notable roles in other pathways. For example, there is still some residual end joining function in the absence of L4, but neither Ligase I nor Ligase III becomes essential following loss of L4 (nuclear loss of both L1 and L3 is synthetically lethal regardless of L4 status). Furthermore, the slower kinetics of MMEJ indicate that it is not likely in direct competition to NHEJ or HR (which do act in direct competition), and instead works as a "catch-all" mechanism that only acts if higher-fidelity pathways are unable to resolve a particular break in a timely manner.

CONCLUSIONS

I would like to leave this chapter highlighting the remarkable advancements in our understanding of NHEJ over the course of my PhD. When I began, there was significant disagreement over DNA synapsis models—with arguments over the significance of DNA-PKcs vs XRCC4/XLF filaments as a primary synaptic mechanism. The application of clever SMI experiments to address NHEJ factor kinetics and dedication of Cryo-EM groups to generate 3D-maps of NHEJ complexes has truly deepened the field's appreciation for how flexible the pathway can be. I look at this project through the lens of those advancements. From March 1st, 2020, I've sought to understand L4's structural role in supporting NHEJ—before we understood anything about the distinct LR complexes or true organization of the SR complex. I've since focused my PhD work—including much work that is not included in this dissertation—through the lens of structure-function relationships between NHEJ's structural factors (XRCC4, XLF, Ku70/80) and L4, advancing our understanding the structural role of L4 and how it promotes end joining beyond catalysis.

REFERENCES

1. Sonmez,C., Toia,B., Eickhoff,P., Matei,A.M., El Beyrouthy,M., Wallner,B., Douglas,M.E., de Lange,T. and Lottersberger,F. (2024) DNA-PK controls Apollo's access to leading-end telomeres. *Nucleic Acids Research*, 10.1093/nar/gkae105.

2. Wang,Y., Ghosh,G. and Hendrickson,E.A. (2009) Ku86 represses lethal telomere deletion events in human somatic cells. *Proceedings of the National Academy of Sciences*, **106**, 12430–12435.

3. Bossaert,M., Moreno,A.T., Peixoto,A., Pillaire,M.-J., Chanut,P., Frit,P., Calsou,P., Loparo,J.J. and Britton,S. (2024) Identification of the main barriers to Ku accumulation in chromatin. *Cell Reports*, **43**.

4. Kragelund,B.B., Weterings,E., Hartmann-Petersen,R. and Keijzers,G. (2016) The Ku70/80 ring in Non-Homologous End-Joining: easy to slip on, hard to remove. *Front Biosci (Landmark Ed)*, **21**, 514–527.

5. Weterings,E., Verkaik,N.S., Keijzers,G., Florea,B.I., Wang,S.-Y., Ortega,L.G., Uematsu,N., Chen,D.J. and van Gent,D.C. (2009) The Ku80 Carboxy Terminus Stimulates Joining and Artemis-Mediated Processing of DNA Ends. *MCB*, **29**, 1134–1142.

6. Inagawa,T., Wennink,T., Lebbink,J., Keijzers,G., Florea,B., Verkaik,N. and van Gent,D. (2020) C-Terminal Extensions of Ku70 and Ku80 Differentially Influence DNA End Binding Properties. *INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES*, **21**.

7. Singleton,B.K., Torres-Arzayus,M.I., Rottinghaus,S.T., Taccioli,G.E. and Jeggo,P.A. (1999) The C Terminus of Ku80 Activates the DNA-Dependent Protein Kinase Catalytic Subunit. *Mol. Cell. Biol.*, **19**, 3267–3277.

8. Sibanda,B.L., Critchlow,S.E., Begun,J., Pei,X.Y., Jackson,S.P., Blundell,T.L. and Pellegrini,L. (2001) Crystal structure of an Xrcc4–DNA ligase IV complex. *Nat Struct Mol Biol*, **8**, 1015–1019.

9. Grawunder,U., Zimmer,D. and Lieber,M.R. (1998) DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. *Current Biology*, **8**, 873–879.

10. Goff,N.J., Brenière,M., Buehl,C.J., de Melo,A.J., Huskova,H., Ochi,T., Blundell,T.L., Mao,W., Yu,K., Modesti,M., *et al.* (2022) Catalytically inactive DNA ligase IV promotes DNA repair in living cells. *Nucleic Acids Research*, 10.1093/nar/gkac913.

11. Meek,K. (2020) Activation of DNA-PK by hairpinned DNA ends reveals a stepwise mechanism of kinase activation. *Nucleic Acids Research*, **48**, 9098–9108.

12. Roy,S., de Melo,A., Xu,Y., Tadi,S., Negrel,A., Hendrickson,E., Modesti,M. and Meek,K. (2015) XRCC4/XLF Interaction Is Variably Required for DNA Repair and Is Not Required for Ligase IV Stimulation. *MOLECULAR AND CELLULAR BIOLOGY*, **35**, 3017–3028.

13. Yu,Y., Wang,W., Ding,Q., Ye,R., Chen,D., Merkle,D., Schriemer,D., Meek,K. and Lees-Miller,S. (2003) DNA-PK phosphorylation sites in XRCC4 are not required for survival after radiation or for V(D)J recombination. *DNA REPAIR*, **2**, 1239–1252.

14. Medina-Suarez,D., Han,L., O'Reily,S., Liu,J., Wei,C., Brenière,M., Goff,N.J., Chen,C., Modesti,M., Meek,K., *et al.* (2024) DNA ligases 3 and 4 cooperate in vivo to facilitate DNA repair and organism viability. *In Review.*

15. Masani,S., Han,L., Meek,K. and Yu,K. (2016) Redundant function of DNA ligase 1 and 3 in alternative end-joining during immunoglobulin class switch recombination. *PNAS*, **113**, 1261–1266.

16. Oh,S., Harvey,A., Zimbric,J., Wang,Y., Nguyen,T., Jackson,P.J. and Hendrickson,E.A. (2014) DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells. *DNA Repair*, **21**, 97–110.

17. Vu,D.-D., Bonucci,A., Brenière,M., Cisneros-Aguirre,M., Pelupessy,P., Wang,Z., Carlier,L., Bouvignies,G., Cortes,P., Aggarwal,A.K., *et al.* (2024) Multivalent interactions of the disordered regions of XLF and XRCC4 foster robust cellular NHEJ and drive the formation of ligation-boosting condensates in vitro. *Nat Struct Mol Biol*, 10.1038/s41594-024-01339-x.

18. Chen,S., Lee,L., Naila,T., Fishbain,S., Wang,A., Tomkinson,A.E., Lees-Miller,S.P. and He,Y. (2021) Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature*, 10.1038/s41586-021-03458-7.

19. Chaplin,A.K., Hardwick,S.W., Stavridi,A.K., Buehl,C.J., Goff,N.J., Ropars,V., Liang,S., Oliveira,T.M.D., Chirgadze,D.Y., Meek,K., *et al.* (2021) Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular Cell*, **0**.

20. Buehl,C.J., Goff,N.J., Hardwick,S.W., Gellert,M., Blundell,T.L., Yang,W., Chaplin,A.K. and Meek,K. (2023) Two distinct long-range synaptic complexes promote different aspects of end processing prior to repair of DNA breaks by non-homologous end joining. *Molecular Cell*, 10.1016/j.molcel.2023.01.012.

21. Stinson,B.M., Moreno,A.T., Walter,J.C. and Loparo,J.J. (2020) A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell*, **77**, 1080-1091.e8.

22. Seif-El-Dahan,M., Kefala-Stavridi,A., Frit,P., Hardwick,S.W., Chirgadze,D.Y., Maia De Oliviera,T., Britton,S., Barboule,N., Bossaert,M., Pandurangan,A.P., *et al.* (2023) PAXX binding to the NHEJ machinery explains functional redundancy with XLF. *Science Advances*, **9**, eadg2834.

23. Tadi,S., Tellier-Lebegue,C., Nemoz,C., Drevet,P., Audebert,S., Roy,S., Meek,K., Charbonnier,J. and Modesti,M. (2016) PAXX Is an Accessory c-NHEJ Factor that Associates with Ku70 and Has Overlapping Functions with XLF. *CELL REPORTS*, **17**, 541–555.

24. Roy,S., Andres,S., Vergnes,A., Neal,J., Xu,Y., Yu,Y., Lees-Miller,S., Junop,M., Modesti,M. and Meek,K. (2012) XRCC4's interaction with XLF is required for coding (but not signal) end joining. *NUCLEIC ACIDS RESEARCH*, **40**, 1684–1694.

25. Balmus,G., Barros,A.C., Wijnhoven,P.W.G., Lescale,C., Hasse,H.L., Boroviak,K., Sage,C. le, Doe,B., Speak,A.O., Galli,A., *et al.* (2016) Synthetic lethality between PAXX and XLF in mammalian development. *Genes Dev.*, **30**, 2152–2157.

26. Cisneros-Aguirre,M., Lopezcolorado,F.W., Tsai,L.J., Bhargava,R. and Stark,J.M. (2022) The importance of DNAPKcs for blunt DNA end joining is magnified when XLF is weakened. *Nat Commun*, **13**, 3662.

27. Graham,T.G.W., Carney,S.M., Walter,J.C. and Loparo,J.J. (2018) A single XLF dimer bridges DNA ends during nonhomologous end joining. *Nat Struct Mol Biol*, **25**, 877–884.

28. Bhargava,R., Sandhu,M., Muk,S., Lee,G., Vaidehi,N. and Stark,J.M. (2018) C-NHEJ without indels is robust and requires synergistic function of distinct XLF domains. *Nat Commun*, **9**, 2484.

29. Ochi,T. and Khan,H. (2023) Plant PAXX has an XLF-like function and stimulates DNA end-joining by the Ku-DNA ligase IV-XRCC4 complex. 10.1101/2023.05.26.542399.

30. Cui,X., Yu,Y., Gupta,S., Cho,Y.-M., Lees-Miller,S.P. and Meek,K. (2005) Autophosphorylation of DNA-Dependent Protein Kinase Regulates DNA End Processing and May Also Alter Double-Strand Break Repair Pathway Choice. *MCB*, **25**, 10842–10852.

31. Neal,J.A., Sugiman-Marangos,S., VanderVere-Carozza,P., Wagner,M., Turchi,J., Lees-Miller,S.P., Junop,M.S. and Meek,K. (2014) Unraveling the Complexities of DNA-Dependent Protein Kinase Autophosphorylation. *Molecular and Cellular Biology*, **34**, 2162–2175.

32. Liu,L., Chen,X., Li,J., Wang,H., Buehl,C.J., Goff,N.J., Meek,K., Yang,W. and Gellert,M. (2021) Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Molecular Cell*, **0**.

33. Öz,R., Wang,J.L., Guerois,R., Goyal,G., KK,S., Ropars,V., Sharma,R., Koca,F., Charbonnier,J.-B., Modesti,M., *et al.* (2021) Dynamics of Ku and bacterial non-homologous end-joining characterized using single DNA molecule analysis. *Nucleic Acids Research*, **49**, 2629–2641.

34. Manoj,F. and Kuhlman,T.E. (2023) Phylogenetic Distribution of Prokaryotic Non-Homologous End Joining DNA Repair Systems in Bacteria and Archaea. 10.1101/2023.09.30.560322.

35. Lees-Miller,J.P., Cobban,A., Katsonis,P., Bacolla,A., Tsutakawa,S.E., Hammel,M., Meek,K., Anderson,D.W., Lichtarge,O., Tainer,J.A., *et al.* (2021) Uncovering DNA-PKcs ancient

phylogeny, unique sequence motifs and insights for human disease. *Progress in Biophysics and Molecular Biology*, **163**, 87–108.

36. Lakshmipathy,U. and Campbell,C. (1999) The Human DNA Ligase III Gene Encodes Nuclear and Mitochondrial Proteins. *Molecular and Cellular Biology*, **19**, 3869.

37. Ellenberger,T. and Tomkinson,A.E. (2008) Eukaryotic DNA Ligases: Structural and Functional Insights. *Annual Review of Biochemistry*, **77**, 313–338.

38. Sallmyr,A., Khattri Bhandari,S., Naila,T. and Tomkinson,A.E. (2023) Mammalian DNA ligases; roles in maintaining genome integrity. *Journal of Molecular Biology*, 10.1016/j.jmb.2023.168276.

39. Sallmyr,A., Rashid,I., Bhandari,S.K., Naila,T. and Tomkinson,A.E. (2020) Human DNA ligases in replication and repair. *DNA Repair*, **93**, 102908.

40. Doherty,A.J. and Suh,S.W. (2000) Structural and mechanistic conservation in DNA ligases. *Nucleic Acids Research*, **28**, 4051–4058.

41. Shuman,S. (2009) DNA Ligases: Progress and Prospects*. *Journal of Biological Chemistry*, **284**, 17365–17369.

42. Conlin,M.P., Reid,D.A., Small,G.W., Chang,H.H., Watanabe,G., Lieber,M.R., Ramsden,D.A. and Rothenberg,E. (2017) DNA Ligase IV Guides End-Processing Choice during Nonhomologous End Joining. *Cell Reports*, **20**, 2810–2819.

43. Chistiakov,D.A., Voronova,N.V. and Chistiakov,A.P. (2009) Ligase IV syndrome. *European Journal of Medical Genetics*, **52**, 373–378.

44. Han,L., Masani,S., Hsieh,C. and Yu,K. (2014) DNA Ligase I Is Not Essential for Mammalian Cell Viability. *Cell Reports*, **7**, 316–320.

45. Goff,N.J., Mikhova,M., Schmidt,J.C. and Meek,K. (2024) DNA-PK: A synopsis beyond synapsis. *DNA Repair*, **141**, 103716.

46. Graham,T.G.W., Walter,J.C. and Loparo,J.J. (2016) Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular Cell*, **61**, 850–858.

47. Meek,K., Lees-Miller,S. and Modesti,M. (2012) N-terminal constraint activates the catalytic subunit of the DNA-dependent protein kinase in the absence of DNA or Ku. *NUCLEIC ACIDS RESEARCH*, **40**, 2964–2973.

48. Stinson,B.M., Carney,S.M., Walter,J.C. and Loparo,J.J. (2024) Structural role for DNA Ligase IV in promoting the fidelity of non-homologous end joining. *Nat Commun*, **15**, 1250.

49. Waters,C.A., Strande,N.T., Pryor,J.M., Strom,C.N., Mieczkowski,P., Burkhalter,M.D., Oh,S., Qaqish,B.F., Moore,D.T., Hendrickson,E.A., *et al.* (2014) The fidelity of the ligation

step determines how ends are resolved during nonhomologous end joining. *Nat Commun*, **5**, 4286.

50. Gu,J., Lu,H., Tippin,B., Shimazaki,N., Goodman,M.F. and Lieber,M.R. (2007) XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *The EMBO Journal*, **26**, 1010–1023.

51. Chen,X., Ballin,J.D., Della-Maria,J., Tsai,M.-S., White,E.J., Tomkinson,A.E. and Wilson,G.M. (2009) Distinct kinetics of human DNA ligases I, IIIα, IIIβ, and IV reveal direct DNA sensing ability and differential physiological functions in DNA repair. *DNA Repair*, **8**, 961–968.

52. Amin,H., Zahid,S., Hall,C. and Chaplin,A.K. (2024) Cold snapshots of DNA repair: Cryo-EM structures of DNA-PKcs and NHEJ machinery. *Progress in Biophysics and Molecular Biology*, **186**, 1–13.

53. Vogt,A., He,Y. and Lees-Miller,S.P. (2023) How to fix DNA breaks: new insights into the mechanism of non-homologous end joining. *Biochemical Society Transactions*, 10.1042/BST20220741.

54. Chen,S., Vogt,A., Lee,L., Naila,T., McKeown,R., Tomkinson,A.E., Lees-Miller,S.P. and He,Y. (2023) Cryo-EM visualization of DNA-PKcs structural intermediates in NHEJ. *Science Advances*, **9**, eadg2838.

55. Hardwick,S.W., Stavridi,A.K., Chirgadze,D.Y., De Oliveira,T.M., Charbonnier,J.-B., Ropars,V., Meek,K., Blundell,T.L. and Chaplin,A.K. (2023) Cryo-EM structure of a DNA-PK trimer: higher order oligomerisation in NHEJ. *Structure*, **31**, 895-902.e3.

56. DeFazio,L.G., Stansel,R.M., Griffith,J.D. and Chu,G. (2002) Synapsis of DNA ends by DNA-dependent protein kinase. *The EMBO journal*, **21**, 3192–3200.

57. Hammel,M., Yu,Y., Mahaney,B.L., Cai,B., Ye,R., Phipps,B.M., Rambo,R.P., Hura,G.L., Pelikan,M., So,S., *et al.* (2010) Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *The Journal of biological chemistry*, **285**, 1414–1423.

58. Zhao,B., Watanabe,G., Morten,M.J., Reid,D.A., Rothenberg,E. and Lieber,M.R. (2019) The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis. *Nat Commun*, **10**, 3588.

59. Reid,D.A., Conlin,M.P., Yin,Y., Chang,H.H., Watanabe,G., Lieber,M.R., Ramsden,D.A. and Rothenberg,E. (2017) Bridging of double-stranded breaks by the nonhomologous end-joining ligation complex is modulated by DNA end chemistry. *Nucleic Acids Research*, **45**, 1872–1878.

60. Reid,D.A., Keegan,S., Leo-Macias,A., Watanabe,G., Strande,N.T., Chang,H.H., Oksuz,B.A., Fenyo,D., Lieber,M.R., Ramsden,D.A., *et al.* (2015) Organization and dynamics of the

nonhomologous end-joining machinery during DNA double-strand break repair. *Proceedings of the National Academy of Sciences*, **112**, E2575–E2584.

61. Wang,J.L., Duboc,C., Wu,Q., Ochi,T., Liang,S., Tsutakawa,S.E., Lees-Miller,S.P., Nadal,M., Tainer,J.A., Blundell,T.L., *et al.* (2018) Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nature structural & molecular biology*, **25**, 482–487.

62. Dylgjeri,E., McNair,C., Goodwin,J.F., Raymon,H.K., McCue,P.A., Shafi,A.A., Leiby,B.E., Leeuw,R., Kothari,V., McCann,J.J., *et al.* (2019) Pleiotropic impact of DNA-PK in cancer and implications for therapeutic strategies. *Clinical cancer research : an official journal of the American Association for Cancer Research*, **25**, 5623–5637.

63. Beucher,A., Birraux,J., Tchouandong,L., Barton,O., Shibata,A., Conrad,S., Goodarzi,A.A., Krempler,A., Jeggo,P.A. and Lobrich,M. (2009) ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *The EMBO journal*, **28**, 3413–3427.

64. Rooney,S., Alt,F.W., Lombard,D., Whitlow,S., Eckersdorff,M., Fleming,J., Fugmann,S., Ferguson,D.O., Schatz,D.G. and Sekiguchi,J. (2003) Defective DNA repair and increased genomic instability in Artemis-deficient murine cells. *The Journal of experimental medicine*, **197**, 553–565.

65. Peng,Y., Woods,R.G., Beamish,H., Ye,R., Lees-Miller,S.P., Lavin,M.F. and Bedford,J.S. (2005) Deficiency in the catalytic subunit of DNA-dependent protein kinase causes down-regulation of ATM. *Cancer Res*, **65**, 1670–1677.

66. Neal,J. and Meek,K. (2019) Deciphering phenotypic variance in different models of DNA-PKcs deficiency. *DNA REPAIR*, **73**, 7–16.

67. Sekiguchi,J., Ferguson,D.O., Chen,H.T., Yang,E.M., Earle,J., Frank,K., Whitlow,S., Gu,Y., Xu,Y., Nussenzweig,A., *et al.* (2001) Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. In *Proceedings of the national academy of sciences of the united states of america*.Vol. 98, pp. 3243–3248.

68. Bertrand,C., Thibessard,A., Bruand,C., Lecointe,F. and Leblond,P. (2019) Bacterial NHEJ: a never ending story. *Mol Microbiol*, **111**, 1139–1151.

69. Deshpande,R.A., Myler,L.R., Soniat,M.M., Makharashvili,N., Lee,L., Lees-Miller,S.P., Finkelstein,I.J. and Paull,T.T. (2020) DNA-dependent protein kinase promotes DNA end processing by MRN and CtIP. *Sci Adv*, **6**, 0922.

70. Chaplin,A.K., Hardwick,S.W., Liang,S., Kefala Stavridi,A., Hnizda,A., Cooper,L.R., De Oliveira,T.M., Chirgadze,D.Y. and Blundell,T.L. (2021) Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nat Struct Mol Biol*, **28**, 13–19.

71. Chen,X., Xu,X., Chen,Y., Cheung,J.C., Wang,H., Jiang,J., de Val,N., Fox,T., Gellert,M. and Yang,W. (2021) Structure of an activated DNA-PK and its implications for NHEJ. *Molecular Cell*, **81**, 801-810.e3.

72. Balmus,G., Pilger,D., Coates,J., Demir,M., Sczaniecka-Clift,M., Barros,A.C., Woods,M., Fu,B., Yang,F., Chen,E., *et al.* (2019) ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. *Nat Commun*, **10**, 87.

73. Mikhova,M., Goff,N.J., Janovič,T., Heyza,J.R., Meek,K. and Schmidt,J.C. (2024) Single-molecule imaging reveals the kinetics of non-homologous end-joining in living cells. 10.1101/2023.06.22.546088.

74. Schaub,J.M., Soniat,M.M. and Finkelstein,I.J. (2022) Polymerase theta-helicase promotes end joining by stripping single-stranded DNA-binding proteins and bridging DNA ends. *Nucleic Acids Research*, 10.1093/nar/gkac119.

75. Beagan,K. and McVey,M. (2016) Linking DNA polymerase theta structure and function in health and disease. *Cell. Mol. Life Sci.*, **73**, 603–615.

76. Li,C., Zhu,H., Jin,S., Maksoud,L.M., Jain,N., Sun,J. and Gao,Y. (2023) Structural basis of DNA polymerase θ mediated DNA end joining. *Nucleic Acids Research*, **51**, 463–474.

77. Wyatt,D.W., Feng,W., Conlin,M.P., Yousefzadeh,M.J., Roberts,S.A., Mieczkowski,P., Wood,R.D., Gupta,G.P. and Ramsden,D.A. (2016) Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell*, **63**, 662–673.

78. He,P. and Yang,W. (2018) Template and primer requirements for DNA Pol θ-mediated end joining. *Proceedings of the National Academy of Sciences*, **115**, 7747–7752.

Chapter 2: Catalytically inactive DNA ligase IV promotes DNA repair in living cells.

By:

Noah J. Goff[1,3,4], Manon Brenière[5], Christopher J. Buehl[1,3,4], Abinadabe J. de Melo[5], Hana Huskova[5], Takashi Ochi[6], Tom Blundell[7], Weifeng Mao[2,3,8], Kefei Yu[2,3], Mauro Modesti[5,], and Katheryn Meek[1,3,4*].

Data chapter

[1]College of Veterinary Medicine, [2]College of Human Medicine, [3]Department of Microbiology & Molecular Genetics, [4]Department of Pathobiology & Diagnostic Investigation, Michigan State University, East Lansing, MI 48824, USA; and [5]Centre de Recherche en Cancérologie de Marseille, CNRS UMR7258, INSERM U1068, Institut Paoli-Calmettes, Aix-Marseille Université, Marseille, France; [6]The Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9TJ, UK; [7]Department of Biochemistry , University of Cambridge, Cambridge CB2 1GA, UK; [8]current address, Department of Biotechnology, Dalian Medical University, Dalian 116044, China.*Correspondence: kmeek@msu.edu

ABSTRACT

DNA double strand breaks (DSBs) are induced by external genotoxic agents (ionizing radiation or genotoxins) or by internal processes (recombination intermediates in lymphocytes or by replication errors). The DNA ends induced by these genotoxic processes are often not ligatable, requiring potentially mutagenic end-processing to render ends compatible for ligation by non-homologous end-joining (NHEJ). Using single molecule approaches, Loparo and colleagues propose that NHEJ fidelity can be maintained by restricting end-processing to a ligation competent short-range NHEJ complex that "maximizes the fidelity of DNA repair". These *in vitro* studies show that although this short-range NHEJ complex requires DNA ligase IV (Lig4), its catalytic activity is dispensable. Here using cellular models, we show that inactive Lig4 robustly promotes DNA repair in living cells. Compared to repair products from wild-type cells, those isolated from cells with inactive Lig4 show a somewhat increased fraction that utilize micro-homology (MH) at the joining site consistent with alternative end-joining (a-EJ). But unlike a-EJ in the absence of NHEJ, a large percentage of joints isolated from cells with inactive Lig4 occur with no MH -- thus, clearly distinct from a-EJ. Finally, biochemical assays demonstrate that the inactive Lig4 complex promotes the activity of DNA ligase III (Lig3).

INTRODUCTION

A growing consensus is emerging that DNA double-strand break repair by the non-homologous end-joining [NHEJ] pathway proceeds through distinct steps (3-10). Loparo and colleagues have shown that a long-range synaptic complex that positions the DNA ends ~115 Angstroms (Å) apart is dependent on the catalytic subunit of the DNA dependent protein kinase [DNA-PKcs, DNA-PK] and its regulatory subunit the DNA end-binding factor Ku. Transition of these long-range complexes to short-range synaptic complexes requires the catalytic activity of DNA-PK and the NHEJ ligase complex including XRCC4, XLF, and DNA ligase IV [Lig4]. These *in vitro* studies establish that whereas the long-range complex effectively blocks DNA end-processing, the short-range complex facilitates both end-processing and ligation (6,11). These studies are in good agreement from cellular studies from Ramsden and colleagues who propose that the ligase complex helps to limit end-processing to promote more error-free repair (12). Of note, in these *in vitro* studies, the catalytic activity of Lig4 is not required to either promote formation of the short-range complex or to facilitate end-processing that occurs in the complex. This result is consistent with previous studies proposing various structural roles for the Lig4 complex including promotion of end-processing (13), promotion of end-synapsis (2,14,15), and notably a report from Chiruvella et al who show that in yeast, a catalytically inactive Lig4 complex increases chromosomal end-joining (16). Support for this emerging two-step model of NHEJ has been bolstered by recent cryo-EM studies that precisely delineate potential long-range and short-range NHEJ complexes (7-9).

In these previous studies, although end-processing mediated by the X family polymerases poll and polm, as well as TDP1, and PNKP was clearly dependent on the short-range complex, end-processing by the Artemis nuclease was not (11). Artemis functions exclusively with DNA-PKcs; its role in facilitating opening of the hairpin DNA termini associated with VDJ recombination [the process that provides for the generation of a diverse repertoire of antibodies and T cell receptors] has been well-defined (17-19). However, Artemis also functions to repair a subset of DSBs with non-ligatable DNA ends (20). Work from our laboratory also strongly suggests that Artemis hairpin opening is not restricted to the short-range complex (21). Briefly, using episomal end-joining assays, we found that cells which lack the NHEJ ligase complex robustly open hairpin

termini. These data are consistent with experiments in developing mouse lymphocytes from NHEJ deficient mice. In these studies, opened hairpin coding joints are observed in mice with defects in the ligase complex (22), but not in mice with defects in either DNA-PK or Artemis (23,24); but, coding end-joining is dramatically impaired in mice with any of these defects in NHEJ (23). Finally, these results are completely consistent with recent structural studies that demonstrate that DNA-PK and Artemis function in a monomeric complex that does not require the Lig4 complex that is apparently requisite for other end-processing activities (25).

It has been known for decades that NHEJ defective cell lines robustly repair DNA DSBs on episomes using the alternative non-homologous end-joining pathway (a-EJ)(26,27). However, our recent study demonstrates that cells lacking the XRCC4/Lig4 complex are similarly defective in joining the opened hairpins as are cells lacking either DNA-PKcs or Artemis where the hairpins remain sealed (21). This suggests that the DNA-PK/Artemis complex that facilitates hairpin opening also, via some undefined mechanism, shields the opened hairpin ends from a-EJ; this "undefined mechanism" would be consistent with the function of the long-range complex (blocking end-processing and ligation) proposed by Loparo and colleagues (11).

Here we explore the potential basis for Lig4's non-catalytic role in NHEJ. To test whether the presence of the ligase complex (but not its ligase activity) is sufficient to promote end-processing in living cells, we generated (via Crispr strategies) both Lig4 deficient and enzymatically inactive Lig4 mutant cell strains. We find that whereas Lig4 deficient cells are similarly impaired in joining DSBs with hairpin termini as are other NHEJ defective cells, cells expressing inactive Lig4 are remarkably proficient in rejoining these DSBs. We conclude that the DNA-PK complex and perhaps the long-range synaptic complex protects DSBs from other DNA ligases and end-processing factors (except for Artemis) in living cells. Moreover, our cellular experiments demonstrate that a catalytically inactive Lig4 complex can efficiently promote end-joining in living cells, consistent with the model that NHEJ-mediated end-processing is limited to a short-range synaptic complex. It follows that either Lig3 or Lig1 must be capable of facilitating ligation of DNA ends processed by NHEJ's short-range complexes. Thus, we used *in vitro* assays and establish that the catalytically inactive Lig4 complex robustly promotes the activity of Lig3 but not Lig1. Finally, repaired DSBs recovered from cells expressing catalytically inactive Lig4 have

increased utilization of short sequence micro-homologies (MH) at the joining site, a well-established characteristic of a-EJ (28).

RESULTS

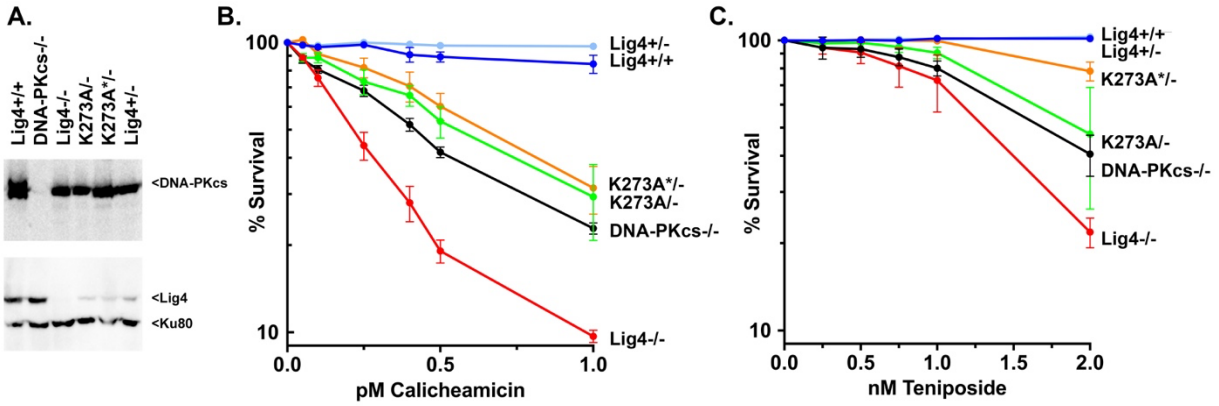


Figure 5: Cells expressing catalytically inactive DNA ligase IV or completely deficient in DNA ligase IV are markedly sensitive to the DSB-inducing agents calicheamicin and teniposide. (A) Western blot of Ligase IV in wild-type 293T cells (Lig4+/+), Lig4 haplo-insufficient cells (Lig4+/), or 293T cells deficient in Lig4 (Lig4-/-), deficient in DNA-PKcs (DNA-PKcs-/-) or with mutations in Lig4 [K273A/-, and K273A*/-]. (These are two independent clones possessing one copy of catalytically inactive Lig4 and a frameshift mutation on the second Lig4 allele; the K273A clone containing an additional mutation, I270V is labelled K273A*/-, for brevity.) **(B)** Sensitivity of the same panel of cell strains to calicheamicin **(B)** or teniposide **(C)** was assessed. Briefly, cells were plated in 24 well plates into complete medium with increasing doses of calicheamicin or teniposide. After 7 days, MTT was added, and cell survival assessed by colorimetry. Survivakl assays were performed four times in duplicate (B) or three times in duplicate (C).

Cells expressing catalytically inactive DNA ligase IV or completely deficient in DNA ligase IV are markedly sensitive to DSB-inducing agents. To begin to address whether Lig4 has a non-catalytic function in living cells, a CRISPR strategy was devised using a single gRNA that targets the codon AAG, encoding K273 [the well-conserved catalytic site (16,29,30) in Lig4 and a single stranded oligonucleotide to direct an HDR mediated K273A mutation at this site (as well as the introduction of a novel HhaI site). From these transfections in 293T cells, single clones were isolated, and DNA was isolated for PCR/HhaI digestion. Amplicon sequencing was performed on clones that included addition of the HhaI site in the initial PCR screen. Four clones were chosen for further analyses (Supplemental Table 1). Sanger sequencing from one clone (clone 17, that lacks a wild-

type allele by initial PCR screen) revealed two alleles with frameshift mutations, and we conclude that this clone is completely deficient in Lig4. Amplicon sequencing of two clones (that have a novel HhaI site ascertained by the initial PCR screen) reveal that one (clone 144) has the targeted K273A mutation on one allele, and a frameshift mutation on the second allele. The second clone (clone 195) contains the K273A mutation as well as an additional mutation, I270V on one allele, and the second allele includes a frameshift mutation (Supplemental Table 1). A heterozygous clone (clone 104) was isolated that has an inactivating frameshift mutation on one allele, but the second allele is wild-type.

Cells with defects in the NHEJ pathway are markedly sensitive to numerous different DNA damaging agents that can induce DNA DSBs by various mechanisms. We chose two drugs, calicheamicin [generating DSBs with 3' overhangs with 3' phosphoglycolate (3'PG)] (31) that generates DSBs throughout the cell cycle (31) and teniposide a topoisomerase II (topII) inhibitor that generates two-ended DSBs primarily in S phase [with 5' overhangs with 5' hydroxyls] (32). Cellular calicheamicin and teniposide sensitivity was assessed for wild-type 293T as well as Lig4+/-, K273A/-, K273A+I270V/- (labelled K273A*/-, for brevity), and Lig4-/-, clones as well as a previously described DNA-PKcs deficient clone (33) using MTT staining as a measure of cellular viability (Figure 5B + C). As can be seen, cells completely deficient in Lig4 are remarkably sensitive to calicheamicin and teniposide; the two clones expressing only catalytically inactive Lig4 are also sensitive to both drugs, but substantially more resistant than cells that are completely deficient in Lig4. In contrast, at these doses, the Lig4+/- heterozygous clone is similarly resistant to both drugs as wild-type 293T cells. Finally, the previously described DNA-PKcs-/- 293T cells are also hypersensitive to calicheamicin and teniposide, but more resistant than Lig4-/- cells consistent with previous studies (23). These data suggest that catalytically inactive Lig4 promotes survival (albeit inefficient) to calicheamicin and teniposide.

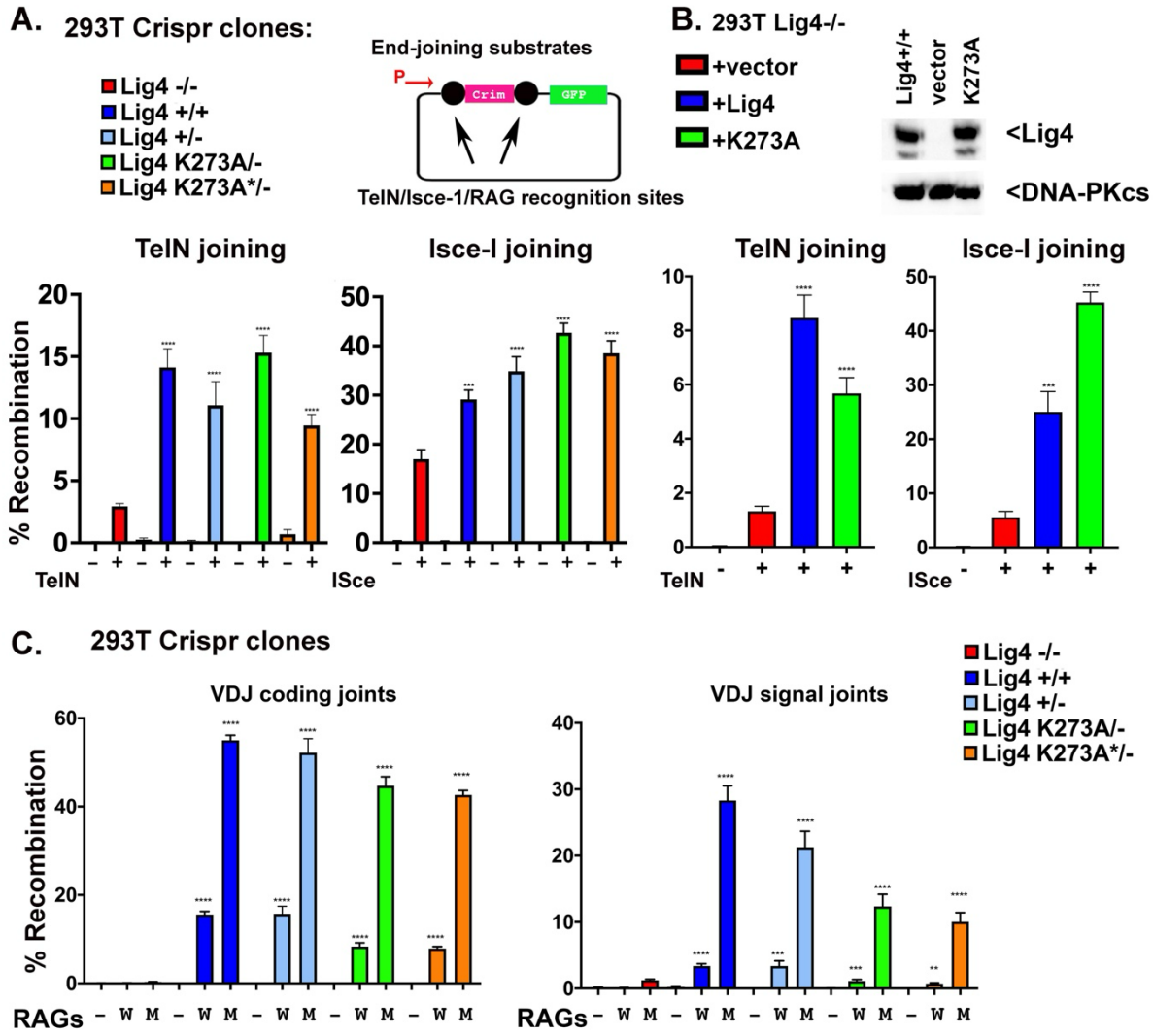


Figure 6: Catalytically inactive DNA ligase IV promotes joining of DSBs in episomal substrates in 293T cells. (A) Fluorescent substrates (depicted in top panels indicate promoter with arrow) were utilized to detect TelN, ISce-1, or V(D)J coding and signal joints in 293T cells of the indicated genotypes. 293T clones with the indicated genotypes were co-transfected with TelN or I-SceI plasmid substrate, with and without enzyme (-/+) and analyzed for red/green fluorescence via flow cytometry (lower panels). (**B)** Lig4-/- 293T cells were tested for TelN and I-SceI joining by co-transfecting the TelN/I-SceI substrates and expression plasmids with or without co-transfection of expression plasmids encoding either WT or catalytically inactive human ligase IV as indicated. (**C)** 293T clones were co-transfected with wild-type (W) or hypermutant (M) RAG expression plasmids with either plasmid substrates to detect coding or signal end-joining as indicated by analysis of red/blue fluorescence via flow cytometry. Episomal V(D)J assays testing joining of coding (hairpin) and signal (blunt) ends. Cells were transfected with substrate and either: no Rag2 (-), WT rags (W), or mutant rag2 (M) and analyzed via flow cytometry. In A, B, and C, student's T test comparing joining rates between Lig4-/- and either Lig4+/+, Lig4+/-, or K273A were performed; ****P<0.0001; ***P<0.001; ns=not significant in two-tailed unpaired t test.

Catalytically inactive DNA ligase IV promotes joining of DSBs on episomal substrates in 293T cells. We have previously developed plasmid substrates with recognition sequences for DNA endonucleases that can produce a range of DNA end structures. The cutting sites flank the coding sequence for a red fluorescent protein (Crimson, RFP) driven by a CMV promoter and by co-transfecting substrate with appropriate enzymes, end-joining efficiency can be assessed by measuring the fraction of cells that express GFP or CFP after excision of the RFP cassette. The TelN protelomerase and I-SceI homing endonuclease produce hairpin DNA ends or 4 nucleotide overhangs respectively. To test whether 293T cells that either lack Lig4 entirely or retain a single copy of a catalytically inactive Lig4 gene can rejoin TelN or I-SceI-induced DSBs, a series of episomal end-joining assays were performed (Figure 6). Our previous work has shown that cells with defects in either the DNA-PK complex, Artemis, or the Lig4 complex are all significantly and similarly impaired in joining TelN-induced hairpin ends, even though the hairpinned ends are opened in cells deficient in the Lig4 complex (that retain DNA-PK and Artemis), but are obviously sealed in cells lacking DNA-PK or Artemis (21). Consistent with this previous study, cells deficient in Lig4 have a marked decrease in TelN joining (Figure 6A). In contrast, rejoining I-SceI DSBs is only ~50% reduced in cells deficient in Lig4, because restriction enzyme-induced DSBs are available to the a-EJ pathway [consistent with previous studies, (26,27,34)]. These data suggest that engagement of the Artemis/DNA-PK complex to resolve the closed hairpins must in some undefined way restrict the opened hairpins to the NHEJ pathway. Strikingly, in contrast to Lig4 deficient cells, the two independent clones expressing only catalytically inactive Lig4 rejoin significant levels of both I-SceI and TelN-induced DSBs (Figure 6A). These data suggest that the Lig4 complex (whether active or not) promotes end-joining by a-EJ presumably by relaxing the NHEJ restriction imposed by hairpin opening by DNA-PK/Artemis.

To confirm that the observed differences in end-joining are the result of the targeted Lig4 mutations, we utilized the Lig4 deficient 293T cells and expression vectors encoding wild-type or catalytically inactive human Lig4 (Figure 6B) in a complementation experiment. As can be seen, both wild-type and catalytically inactive Lig4 (but not empty vector) substantially reverse both TelN and ISce-1 joining in 293T cells completely deficient in Lig4.

In cells that lack NHEJ, DSBs (including those induced by restriction enzymes, those introduced during class switch recombination, or by CRISPR or other gene editing strategies) can all be efficiently re-joined by the a-EJ pathway. In contrast, NHEJ-deficient cells do not join VDJ-associated DSBs because VDJ recombination intermediates are tightly restricted to the NHEJ pathway by mechanism(s) that are still incompletely understood. To corroborate the TelN joining experiments, we next assessed V(D)J coding and signal end-joining in the same panel of 293T cells. In these experiments we used both wild-type RAG expression vectors as well as a well-studied hyper-RAG2 mutant that substantially increases VDJ joining by de-stabilizing the RAG post-cleavage complex(s) that function to limit end-joining to the NHEJ pathway (35). As can be seen, 293T cells lacking Lig4 are effectively incapable of joining RAG-induced DSBs, either blunt-ended signal ends or hairpinned coding ends. In contrast, cells expressing catalytically inactive Lig4 join both coding and signal ends at levels only modestly reduced as compared to levels observed in wild-type 293T or Lig4 haplo-insufficient 293T cells. The hyper RAG2 mutant substantially increases the level of both coding and signal joining in wild-type cells, but joining is minimal in cells completely deficient in Lig4. In contrast, both signal and coding end-joining is robust, and only modestly reduced in cells expressing catalytically inactive Lig4 in experiments utilizing the hyper-RAG mutant (Figure 6C). It is interesting to note that cells expressing catalytically inactive Lig4 have a more severe deficit in signal end-joining than in coding end-joining. To our knowledge, this is unlike VDJ deficits in any other NHEJ mutant studied to date, and perhaps suggest that catalytically inactive Lig4 can promote joining of over-hanged DNA ends better than blunt DNA ends. In sum, we conclude that catalytically inactive Lig4 efficiently promotes rejoining of RAG-induced DSBs. These data reveal that once V(D)J intermediates are released to a complete Lig4 complex (potentially, the short-range NHEJ complex), the strict requirement for NHEJ-only dependent joining of RAG-induced DSBs has been fulfilled, and the RAG-induced DSBs can be joined by the a-EJ ligases just as well as non-RAG induced DSBs.

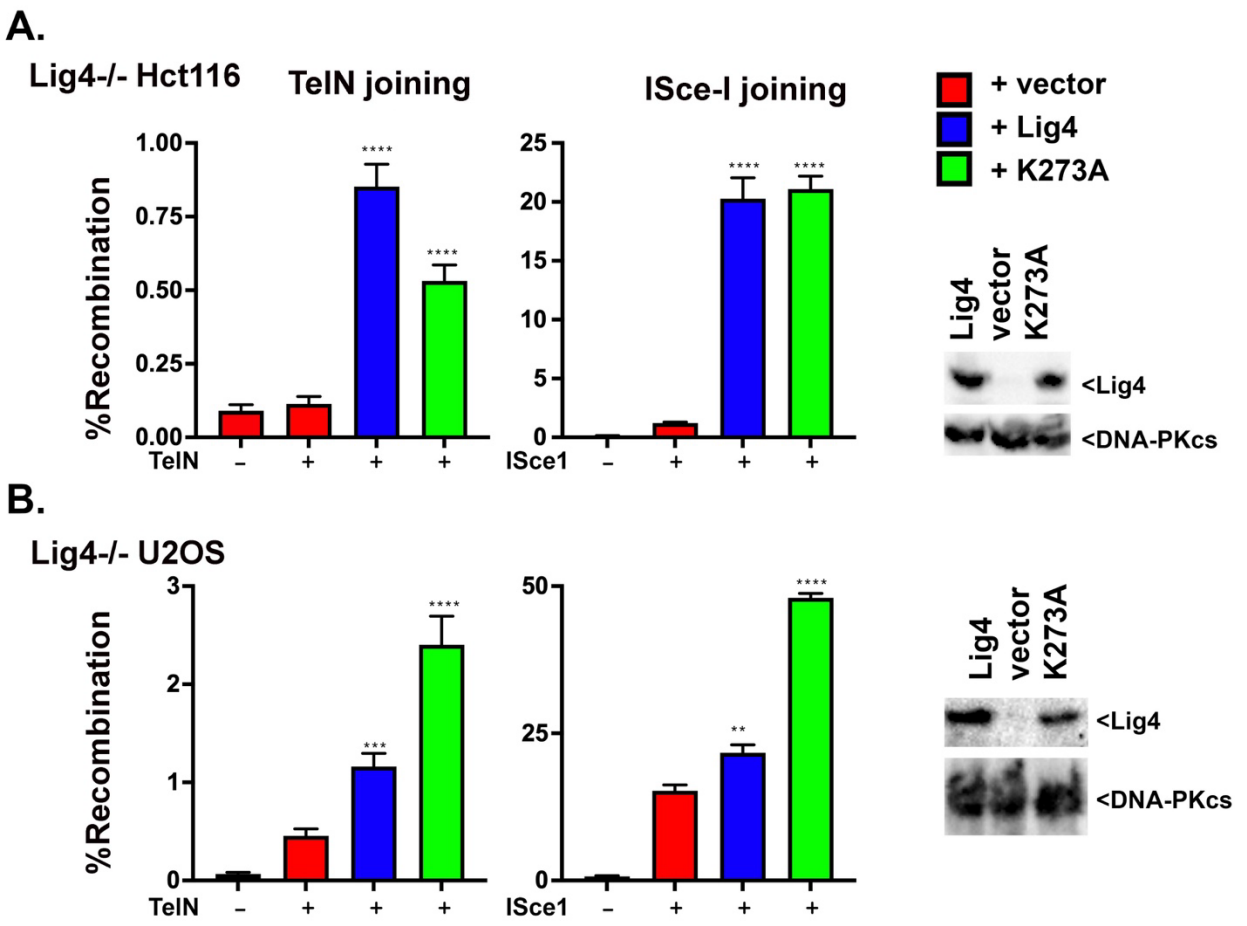


Figure 7: Catalytically inactive DNA ligase IV promotes joining of DSBs in episomal substrates in a variety of cell types. (**A)** Lig4-/- HCT116 cells were tested for TelN and I-SceI joining by co-transfecting the TelN/I-SceI substrates and expression plasmids with or without co-transfection of expression plasmids encoding either WT or catalytically inactive human ligase IV as indicated. (**B)** Lig4-/- U2OS cells were tested for TelN and I-SceI joining by co-transfecting the TelN/I-SceI substrates and expression plasmids with or without co-transfection of expression plasmids encoding either WT or catalytically inactive mouse ligase IV as indicated. In A and B, student's T test comparing joining rates between vector and Lig4 or K273A were performed; ****P<0.0001; ***P<0.001; **P<0.01; ns=not significant in two-tailed unpaired t test.

Catalytically inactive DNA ligase IV promotes joining of DSBs on episomal substrates in diverse cell types. To extend these studies, a CRISPR/Cas9 strategy was utilized to ablate Lig4 from the U2OS cell strain, and a Lig4 deficient HCT116 cell strain was obtained from Dr. Eric Hendrickson (36). As can be seen, in these cell types, both wild-type and catalytically inactive Lig4 (but not empty vector) substantially reverse both TelN and ISce-I end-joining (Figure 7), completely

analogous to the results observed in Lig4-/- 293T cells. These data substantiate our conclusion that the Lig4 complex facilitates end-joining of a variety of DSBs on episomal substrates.

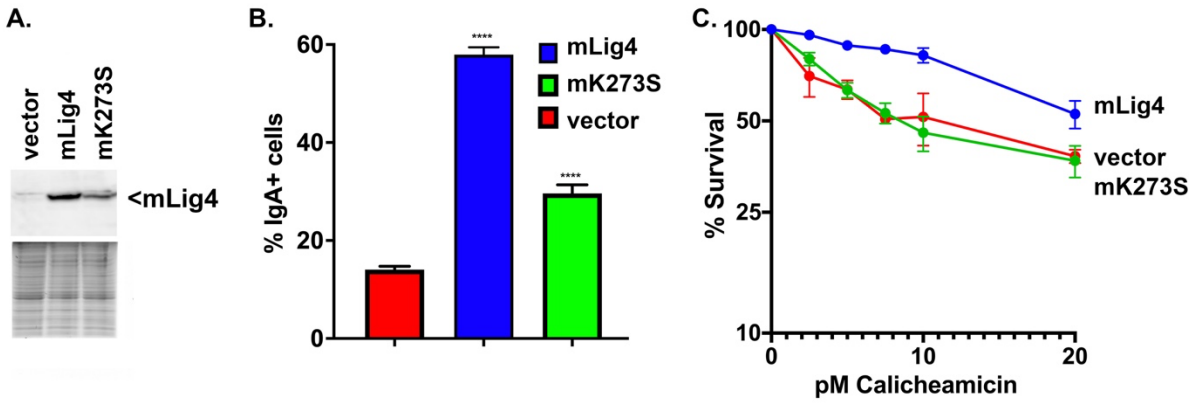


Figure 8: Catalytically inactive Ligase IV is only moderately deficient in CSR. (A) Western blot of murine Ligase IV expressed in Lig4-/- CH12 cells. **(B)** CSR efficiency (percentage of IgA-positive cells, assessed by flow cytometry) in mouse CH12 cells expressing WT ligase IV, catalytically inactive Ligase IV (dLig4), or no ligase IV. Error bars indicate SE of three independent experiments. **(C)** Sensitivity of the same panel of cell strains to calicheamicin was assessed. Briefly, cells were plated in 24 well plates into complete medium with increasing doses of calicheamicin. After 7 days, MTT was added, and cell survival assessed by colorimetry. In B, student's T test comparing joining rates between vector and Lig4, or vector K273S were performed; ****P<0.0001; in two-tailed unpaired t test. Survival assay was performed three times in duplicate.

Catalytically inactive DNA ligase IV supports chromosomal rearrangement during class switch recombination. Previously we studied the role of Lig4 in the process of immunoglobulin class switch recombination (CSR) using a well-studied mouse B cell model, CH12 cells (37). Retroviral vectors encoding either wild-type or K273S murine Lig4 were prepared and used to transduce a previously described Lig4-/- Ch12 cell strain. When CH12 cells are stimulated with an anti-CD40 antibody, interleukin 4, and TGFβ1 (+CIT), CSR (from IgM to IgA) is robustly induced. This produces clear populations of IgM+/IgA-, IgM+/IgA+, and IgM-/IgA+ cells that can be analyzed by flow cytometry. Like VDJ recombination, CSR is a lymphocyte specific DNA recombination event that proceeds through a double-strand break intermediate; but unlike VDJ recombination, a-EJ can facilitate CSR (albeit at reduced levels) in the absence of most NHEJ factors. Our previous study demonstrated that either Lig1 or Lig3 could facilitate a-EJ in a Lig4 deficient murine B cell line (CH12 cells) that were induced to undergo class switch recombination.

We next assessed CSR in the CH12 Lig4-/- cells expressing wild-type Lig4, K273S, or vector only from lentivirus expression vectors. Consistent with our previous study, only ~10% of Lig4 deficient CH12 cells expressing vector alone undergo CSR compared to ~60% of cells expressing wild-type murine Lig4. Cells expressing K273S Lig4 display an intermediate level of CSR (Figure 8). We conclude that catalytically inactive Lig4 promotes chromosomal end-joining of cell-programmed DSBs. We also assessed cellular resistance to calicheamicin in this panel of CH12 cells. As can be seen, whereas cells expressing wild-type Lig4 are substantially more resistant to calicheamicin than cells expressing no Lig4, cells expressing catalytically inactive Lig4 are similarly sensitive to calicheamicin as are cells lacking Lig4. We conclude that catalytically inactive Lig4 cannot restore cellular resistance to DNA damaging agents in all cell types.

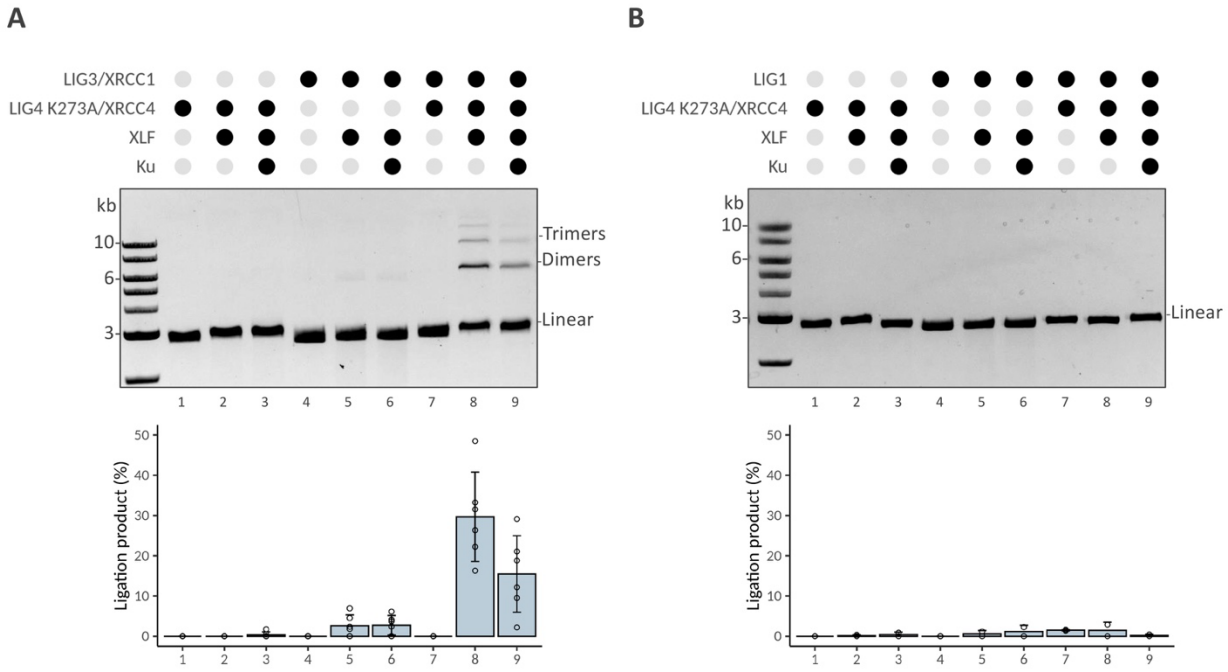


Figure 9: Catalytically inactive Lig4/XRCC4/XLF promotes Lig3-mediated joining of both cohesive and blunt DNA ends *in vitro.* The capacity for catalytically inactive Lig4/XRCC4 to promote ligation by Lig3 **(A)**, or Lig1 **(B)** was assessed by incubating a blunt DNA ligation substrate (linearized pUC19) with purified recombinant proteins as indicated at the following concentrations: Ku, 250nM; XLF, 200nM; LIG4/XRCC4, 500nM; LIG3/XRCC1, 40nM; and LIG1,180nM. 10ul reactions were incubated for 30 minutes at room temperature and analyzed by gel electrophoresis and quantified with Image J. Bar graph represents three independent experiments and error bars represent SD.

Catalytically inactive Lig4/XRCC4/XLF promotes Lig3-mediated joining of both cohesive and blunt DNA ends *in vitro.* The robust joining facilitated by catalytically inactive Lig4 in cellular assays suggests that the NHEJ short-range complex might be capable of facilitating ligation of DNA ends by either Lig3 or Lig1. To directly test this possibility, Lig1 and Lig3/XRCC1 were expressed and purified from *E. coli* (Sup. Figure 1A) and tested for activity (Sup. Figure 1B). Then purified Lig1 or Lig3/XRCC1 were tested in ligase assays that included the components of the NHEJ short range complex (Ku, X4/Lig4, and XLF) using catalytically inactive Lig4, on substrates with either blunt or cohesive DNA ends. In these assays, higher levels Lig4/XRCC4 were used as compared to either Lig1 (2.8 molar excess) or Lig3/XRCC1 (12.5 molar excess). As can be seen (Figure 9A), using a blunt ligation substrate, no ligation products are observed with just the NHEJ components, and no activity is observed with only Lig3/XRCC1. In contrast, when Lig3/XRCC1 is added to reactions containing catalytically inactive Lig4 (with or without Ku), robust ligation is observed. This activity is completely dependent on both Lig4/XRCC4 and XLF, but Ku is dispensable. In parallel assays, Lig1 activity is not significantly enhanced by catalytically inactive Lig4 complex (Figure 9B) under any of the conditions tested.

This assay was repeated using DNA with cohesive ends (Sup. Figure2); as can be seen, ligation of compatible ends is robust in the presence of catalytically inactive Lig4 complex and Lig3/XRCC1, but not Lig1, suggesting that catalytically inactive Lig4 complex can stimulate Lig3/XRCC1-mediated joining of cohesive ends. However, (to our surprise) we consistently observed very low levels of ligation products in reactions with XLF and XRCC4/Lig4 (catalytically inactive, K273A) without Lig3/XRCC1 or Lig1. The activity represents <5% the activity observed in assays with wild-type Lig4 complexes (compare Sup. Figure 1C and 2A). This residual activity could not be attributed to a contaminating ligase since the proteins are prepared in *E. coli* that has only NAD dependent ligase. Moreover, adenylation assays demonstrate minimal adenylation of the K273A mutant Lig4 (Sup. Figure 3). From the recent cryo-EM study of the short-range complex (7), we observed that 4 additional lysine residues are very close to the 5' phosphate of the DNA end; whereas K273 is ~7Å away, lysines 449, 451, 352, and 345 are between 8Å and 9.5Å away from the 5' phosphate. We considered that one of these four lysines might serve as a back-up adenylation site explaining the minimal ligase activity of the K273A mutant. These four residues

were substituted with arginine in the K273A mutant construct. As can be seen, this 5X lysine mutant has no residual activity in ligase assays but retains the capacity to stimulate Lig3 activity towards both blunt and cohesive ends *in vitro* (Supplemental Figure 4A) to a similar extent as the K273A mutant. Moreover, in cellular assays, the 5X lysine mutant joining activity that is indistinguishable from that of the K273A mutant (Supplemental Figure 4B). From these data, we conclude that the catalytically inactive Lig4/XRCC4/XLF complex promotes end-joining of Lig3 *in vitro,* potentially explaining the robust joining activity observed in living cells expressing the catalytically inactive complex.

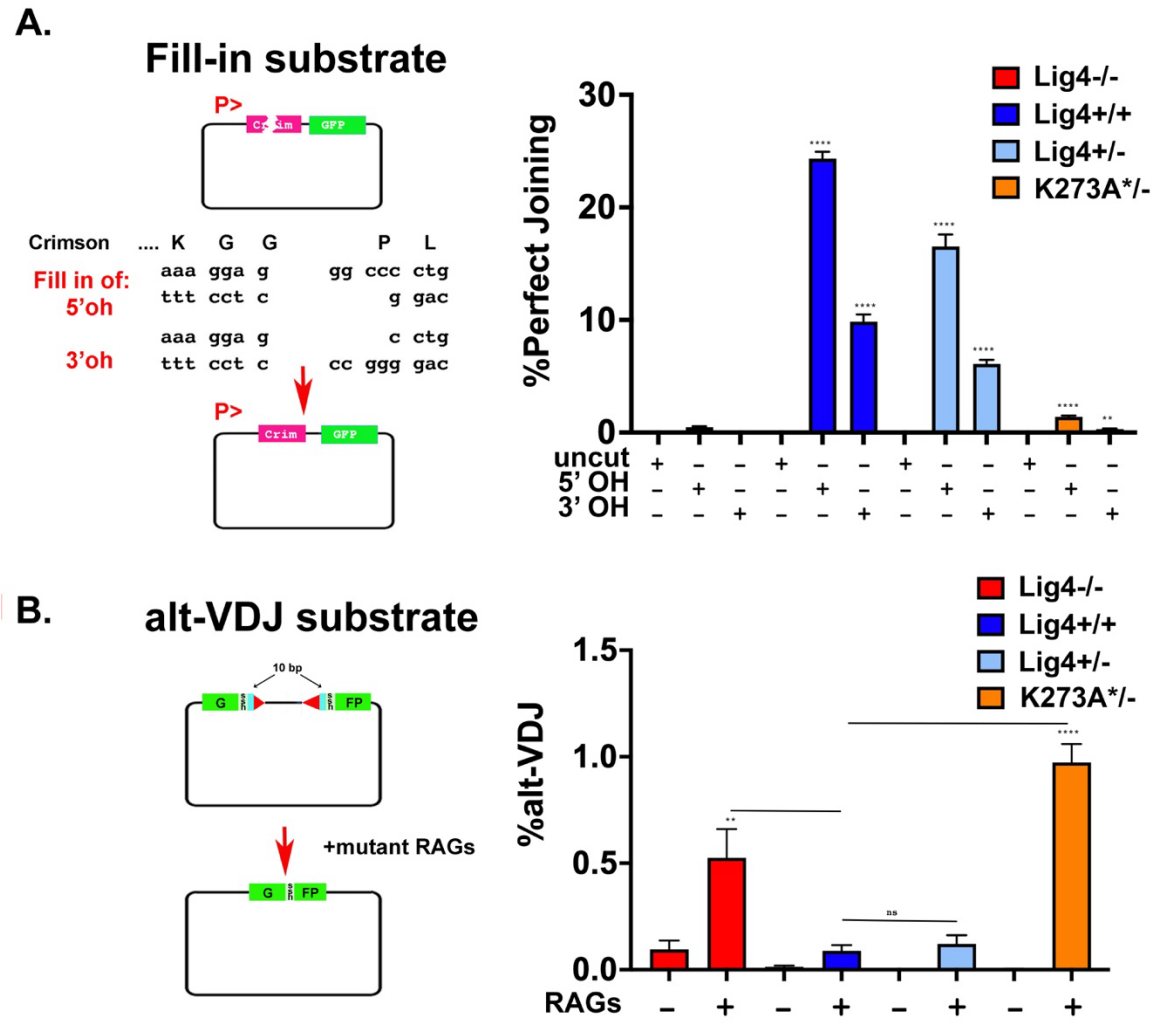


Figure 10: DSBs repaired in cells expressing K273A Lig4 have characteristics of a-EJ. **(A)** 293T clones of the indicated genotypes were transfected with Fill-in joining substrate either uncleaved, or cleaved with Eco53KI and PspOMI (to generate blunt and 5' overhangs), or Eco53K1 and Apa1(to generate blunt and 3' overhangs). Cells were assessed for RFP and GFP expression 72-

Figure 10 (cont'd)
hours later. Student's T test comparing joining rates between Lig4-/- and either Lig4+/+, Lig4+/-, or K273A*/- were performed; ****P<0.0001; in two-tailed unpaired t test. **(B)** 293T clones of the indicated genotypes were co-transfected with the alt-VDJ substrate, a ds-RED expression plasmid, and hypermutant RAG expression plasmids. Restoration of the GFP reading frame requires joining via a 9bp region of MH that occurs 10bp from the termini of both coding ends. Cells were assessed for RFP and GFP expression 72 hours after transfection. Student's T test comparing joining rates between Lig4-/- and either Lig4+/+, Lig4+/-, or K273A*/- as well as betwee Lig4+/+ and either Lig4+/- or K273A*/- were performed; ****P<0.0001; ***P<0.001; **P<0.01; in two-tailed unpaired t test.

Cells expressing K273A Lig4 cannot perform fill-in end-processing, but robustly promote a-EJ in episomal assays. Clearly catalytically inactive Lig4 enhances end-joining of episomal substrates and in some cell types, cellular survival after exposure to agents that induce DSBs. To gain knowledge as to how end-joining is enhanced by the catalytically inactive Lig4 complex, we used assays that clarify structural characteristics of repaired DSBs. More specifically, we focused on determining whether catalytically inactive Lig4 promotes end-joining that is similar to authentic NHEJ [i.e. generating joints ranging from perfect end-joining with no or minimal base pair loss, or joining at sites of MH (1-3bp)], or alternatively, if catalytically inactive Lig4 promotes joining that is more similar to a-EJ [characterized by increased loss of terminal nucleotides, and a strong dependence on the presence of MH (often longer than 3bp)].

Although NHEJ is generally characterized as an error-prone repair mechanism, numerous studies have documented the fidelity of NHEJ when joining most DSBs (11,12). For example, perfect joining of compatible ends are highly favored and incompatible ends are rejoined to minimize nucleotide loss (11,12). To assess the capacity of catalytically inactive Lig4 to promote joining that is similar to that of NHEJ, a joining substrate that measures fidelity of joining was generated (Figure 10A). Briefly, the substrate plasmid is restricted with two enzymes to generate a blunt end on one side and an over-hanged end (either 5' or 3') on the other side as illustrated. To restore the crimson open reading frame, the ends must be aligned so that the missing bases across from the overhangs are filled in. In both cellular and *in vitro* models of NHEJ, perfect fill-in of DNA ends like these is efficiently mediated by NHEJ (11,12). If uncut plasmid is transfected, crimson is not expressed because of the disruption in the open reading frame, but GFP is expressed by use of its own ATG. When restricted plasmids are transfected, if any plasmid

rejoining occurs, GFP is expressed. With this assay, "Perfect rejoining" represents the percent GFP positive cells that also express crimson, and in these assays, GFP expression is robust in all cell types. As would be expected, cells completely lacking Lig4-/- are completely unable to perfectly rejoin the transfected substrate promoting crimson expression, whereas in wild-type 293T cells more than 30% of the cells that rejoin the plasmid, rejoin the plasmid perfectly and crimson is robustly expressed (Figure 10A). In contrast to other joining assays (Figure 6), cells expressing only catalytically inactive Lig4 are similarly deficient in perfect rejoining as are cells completely lacking Lig4. We conclude that catalytically inactive Lig4 cannot promote perfect rejoining of the fill-in episomal substrate. It follows that the catalytically inactive complex either does not efficiently support fill-in end-processing, or that the cellular ligase facilitating end-joining (Lig3 or Lig1) does not efficiently join the filled-in ends.

To assess whether a-EJ-like events facilitate repair in cells expressing catalytically inactive Lig4, we utilized an assay developed by Roth and colleagues (38). It is well-appreciated that short sequence homologies at opened coding-end termini can facilitate coding end-joining (39); these short sequence homologies are generally small (1-3 base pair) and occur close to the DNA terminus. We utilized their a-EJ assay (Figure 10B, termed alt-VDJ, using mutant RAG expression constructs that destabilize the RAG post-cleavage complex(s) to promote more alt-VDJ) to determine whether cells expressing catalytically inactive Lig4 depend on MH for joining. This assay requires nucleotide deletions of 10bp from each coding end, and then use of 9 base pair of short sequence homology to restore the GFP open-reading frame. Whereas minimal alt-VDJ is observed in wild-type 293T cells or Lig4 haplo-insufficient cells where end-processing generally precludes loss of the required 10bp from each coding-end, alt-VDJ is readily detected in Lig4-/- 293T cells (Figure 10B). Strikingly, alt-VDJ joining is robust in cells expressing catalytically inactive Lig4; these data suggest that K273A mutant Lig4 promotes end-joining that is similar to a-EJ.

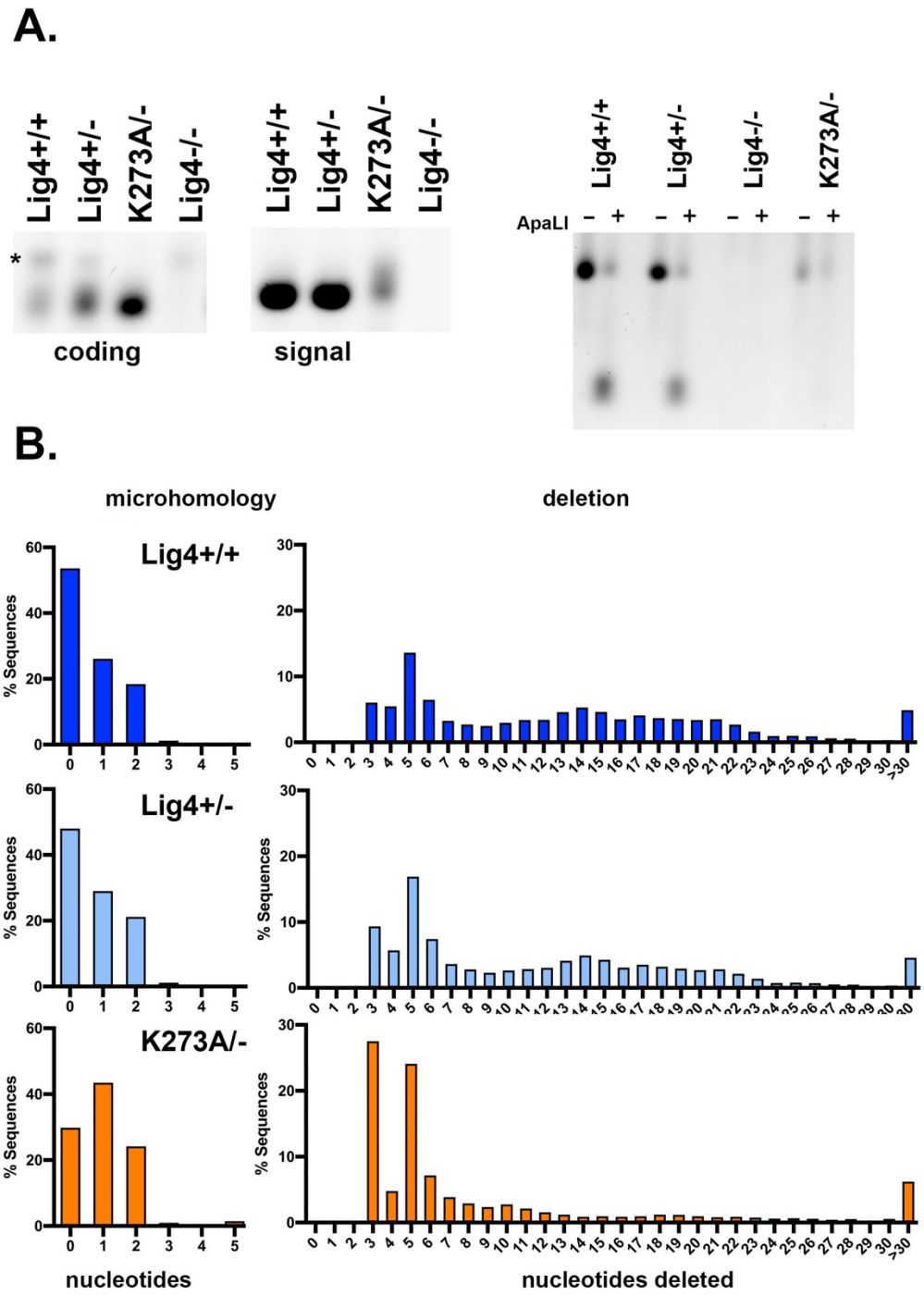


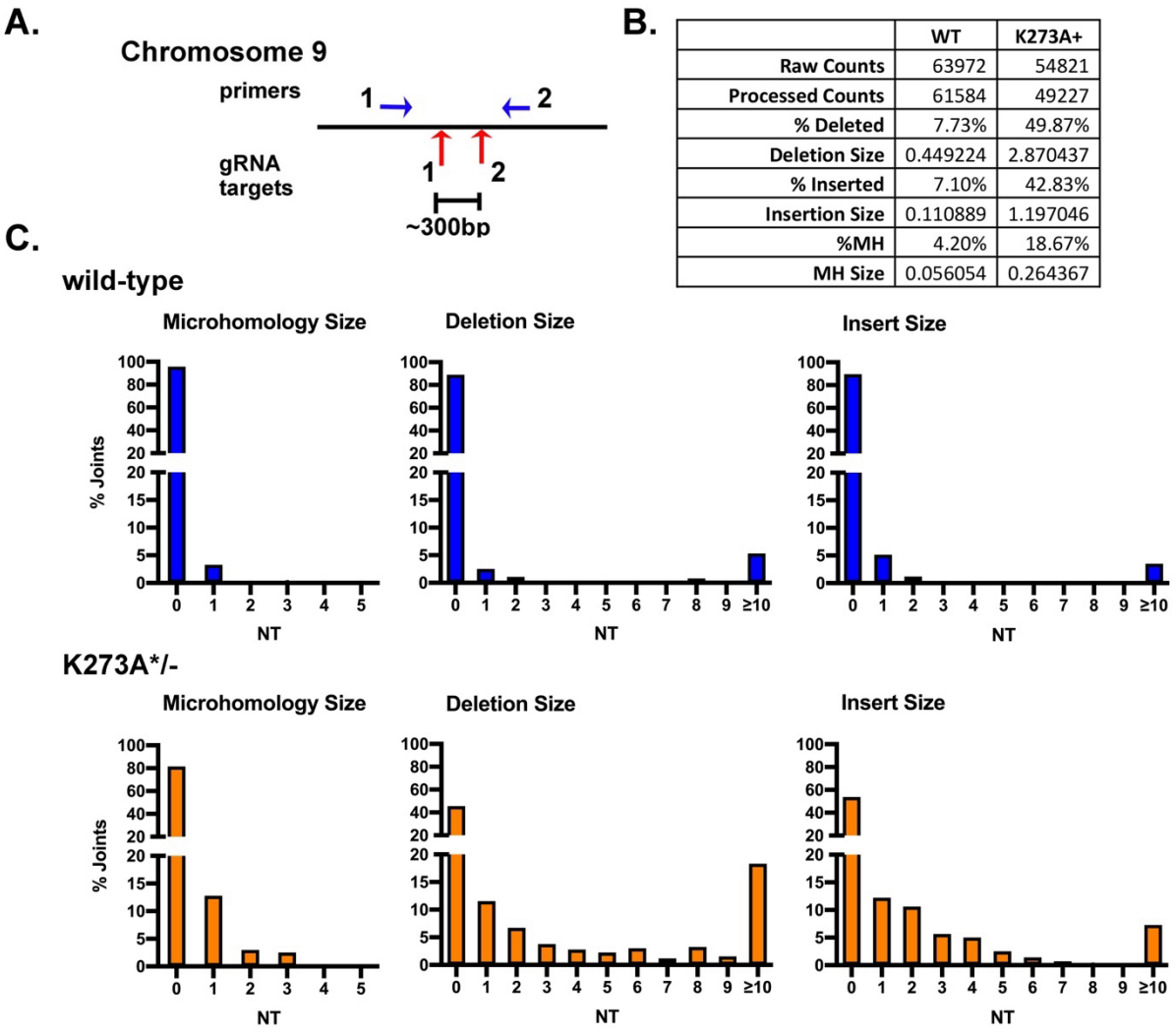
Figure 11: Rejoined episomal VDJ coding and signal joints in cells expressing K273A Lig4 have characteristics of a-EJ mediated joining. **(A)** Coding (left) and signal joints (middle) were PCR amplified from substrates isolated from cells of the indicated genotypes 72 hours after transfection. (*) indicates a PCR artifact. Half of the PCR reaction for signal joint amplification was digested with ApaLI prior to electrophoresis (right). **(B)** Isolated coding joints from the indicated cell strains were subjected to amplicon sequencing. Histograms depicting numbers of nucleotides deleted or extent of microhomology at site of joining.

Rejoined episomal DSBs in cells expressing K273A Lig4 have characteristics of a-EJ mediated joining. To further characterize DSB joining events in cells expressing catalytically inactive Lig4, we characterized VDJ joints isolated from cells expressing the inactive complex. In cells with intact NHEJ, VDJ coding end-joining generates a diverse array of rejoined opened hairpin coding ends. In contrast, rejoining of the blunt signal ends is usually a precise, perfect head-to-head joining of the two heptamers. This recapitulates VDJ joining characteristics in developing lymphocytes. It has been shown that the rare VDJ joints generated in NHEJ deficient cells or animals have characteristics of a-EJ, such that coding joints have increased levels of nucleotide loss and the joints often occur at regions of MH; whereas signal ends are not perfectly rejoined. As shown above (Figure 6C), in cells completely deficient in Lig4, the joining rate of both coding and signal ends is severely reduced. Coding and signal joints from episomal assays were PCR amplified and analyzed by gel electrophoresis. As can be seen (Figure 11A, left), coding joints amplified from wild-type or Lig4 haplo-insufficient 293T cells, generate a diverse "smear" of coding joints, whereas coding joints amplified from 293T cells that express only catalytically inactive Lig4 are homogeneous, and slightly smaller than those recovered from wild-type cells.

As expected, signal joints from wild-type or Lig4 haplo-insufficient 293T cells are uniform, but signal ends isolated from cells that express catalytically inactive Lig4 are diverse, generating a smear of joints (Figure 11A middle). Rejoining of signal ends (without nucleotide loss or gain) generates a novel ApaLI site; so fidelity of signal end-joining can be ascertained by restricting signal joints with this enzyme. Signal joints amplified from assays in wild-type cells are uniform and largely susceptible to ApaLI cleavage consistent with perfect rejoining of signal joints in NHEJ-proficient cells (40). In contrast, signal joints amplified from cells that lack Lig4 or express only catalytically inactive Lig4 are completely resistant to ApaLI cleavage indicating nucleotide loss or addition from the signal ends prior to rejoining, suggesting catalytic inactive Lig4 complex is not proficient at promoting perfect blunt end-joining (Figure 11A, right).

PCR amplified coding joints were submitted for amplicon sequencing; as can be seen (Figure 11B), joints isolated from wild-type or Lig4 haplo-insufficient cells are virtually indistinguishable, and ~50% do not have sequence MH at the site of joining. In these assays, there is a bias for joining at particular sites of microhomology resulting over-representation of

sequences with 3bp and 5bp deletions; these over-represented joints can be appreciated in all three samples. In wild-type cells, the predominate 3bp and 5bp deletion products account for 5.99% and 12.12% of all sequences, in L4 haplo-insufficient cells, 9.28% and 14.66%, but in K273A mutant cells these two particular joints account for 27.4% and 18.37% of all joints. These data support the conclusion that joints facilitated by catalytically inactive Lig4 have characteristics of a-EJ.



	WT	K273A+
Raw Counts	63972	54821
Processed Counts	61584	49227
% Deleted	7.73%	49.87%
Deletion Size	0.449224	2.870437
% Inserted	7.10%	42.83%
Insertion Size	0.110889	1.197046
%MH	4.20%	18.67%
MH Size	0.056054	0.264367

Figure 12: Chromosomal DSBs repaired in cells expressing K273A Lig4 have characteristics of a-EJ. (A) Diagram of region on chromosome nine targeted by two different gRNAs and position of primers utilized to detect chromosomal deletions induced by Cas9 and indicated gRNAs. **(B)** Summary of amplicon analyses of PCR amplified deletional joints from indicated cell strains. **(C)** Histograms depicting numbers of nucleotides deleted, inserted, or extent of microhomology at site of joining.

Chromosomal end-joining in cells expressing K273A Lig4 has characteristics of a-EJ. To extend these findings to chromosomal end-joining, we exploited CRISPR/Cas9 targeting of the FANCG gene on chromosome 9, which we found to be remarkably efficient in 293T cells (41). Although this strategy does not allow quantification of joining, the quality of joining can be assessed by sequencing. Briefly, cells of the indicated genotypes were transfected with two gRNA/cas9/puro plasmids that target sequences ~300bp apart; after 48 hours, cells were subjected to puromycin selection. After 72 hours, genomic DNA was isolated and PCR was utilized to assess deletional rejoining of the two DSBs. Isolated PCR fragments were submitted for amplicon sequencing. In wild-type cells, ~92% of recovered joints are perfect (of the blunt-ended Cas9 cleavage sites), whereas in K273A mutant cells, only ~80% of the joints are perfect. In wild-type cells, the average nucleotide loss/joint is 0.45bp, and ~4% of joints occur at sites of MH. In contrast, in K273A mutant cells, the average loss/joint was 2.9bp, and ~19% occur at sites with MH. The range of nucleotide loss, nucleotide insertion, and utilization of short sequence homologies for joints from wild-type or K273A mutant cells is shown in Figure 12C. We conclude that joints from K273A mutant cells are consistent with joining via a-EJ.

DISCUSSION

In other mutational studies that block enzymatic function of NHEJ factors (ie Artemis or DNA-PKcs), inactivation results in a cellular phenotype that is similar to (42-44) or in certain cell types worse than complete loss (33,45). Thus, it appears that the Lig4 complex is unique in that in its inactive form, Lig4 can promote the function of other repair factors.

There have been several previous studies proposing various structural roles for the Lig4 complex. In 2007, Chu and colleagues reported that the NHEJ ligase complex promotes end-processing activities *in vitro* (13). This is consistent with more recent reports, where in an *in vitro* model of NHEJ, Stinson et al. demonstrated that Lig4 (either active or inactive) was necessary for formation of a short-range non-homologous end-joining complex that closely juxtaposed two DNA ends, promoting both end-processing and ligation (11).

Besides the studies from Loparo and colleagues suggesting a two-stage model of NHEJ (6,11), several other reports support a role in DNA end synapsis for the Lig4 complex (2,14,15). Reid et al demonstrated that the Lig4 complex can promote end-synapsis *in vitro* (using purified

proteins in a FRET assay), and the ability to synapse the ends was strongly impacted by DNA end structure, especially by the presence of the 5' phosphate (2). Of note, a K273A mutant Lig4 complex was defective in synapsing DNA ends in their assays using purified proteins (2), whereas similar FRET assays from Graham et al showed that K273A using Xenopus extracts, fully support end synapsis (6). Conlin et al, extended the studies of end synapsis to show that a unique pocket in Lig4 allows synapsis of DNA ends with mis-matched termini (15). There are also cellular studies that support a role for the Lig4 complex in DNA end synapsis. Cottarel et al. demonstrated that cells deficient in Lig4 are defective in damage-induced autophosphorylation of DNA-PKcs at serine 2056 (S2056) (14). We first showed that S2056 autophosphorylation could occur in trans, and that S2056 phosphorylation occurs almost exclusively by autophosphorylation (46). Recent structural studies strongly support the conclusion that S2056 phosphorylation occurs in *trans* (25). Cottarel et al. demonstrated that catalytically inactive Lig4 could restore the deficit in S2056 phosphorylation in Lig4 deficient cells as well as wild-type Lig4. From these studies they proposed that the catalytically inactive Lig4 complex facilitates synapsis, promoting trans-autophosphorylation of DNA-PKcs at S2056 (14). We extended this finding by demonstrating that XLF also participates in the capacity of the Lig4 complex to promote S2056 phosphorylation in living cells, and that only XLF that was proficient in interacting with XRCC4 (to generate DNA-bridging filaments) could promote S2056 phosphorylation (47). Finally, Chiruvella et al reported that a catalytically inactive Lig4 complex increases not only synapsis, but chromosomal end-joining in yeast; this impact on end-joining was absolutely dependent on XLF (16). This result is entirely consistent with cellular assays presented here, and with the *in vitro* assays demonstrating that catalytically inactive Lig4/XRCC4 only stimulates Lig3/XRCC1 joining in the presence of XLF (Figure 9).

Our working model is that Lig4 itself is important in promoting synapsis and that formation of a short-range synaptic complex (even one that lacks Lig4 activity): 1) facilitates joining of many DSBs, 2) promotes cellular resistance to drugs that induce DSBs in some cell types, and 3) fulfills the RAG-induced restriction step of VDJ recombination so that even RAG-induced DSBs (both coding and signal ends) can be joined by non-NHEJ DNA ligases.

In sum, data presented here extend these previous studies demonstrating a non-catalytic role for Lig4 that robustly stimulates both end-joining and cellular resistance to DNA damage. Joining that is facilitated by the inactive Lig4 complex shares characteristics with a-EJ (increased nucleotide loss, and increased use of microhomologies). This observation presents an unanticipated possibility, that the inactive Lig4 complex (likely consistent with short-range complexes proposed recently) may facilitate a-EJ mediated by Lig3. These studies present an important unanswered question: Do complexes with catalytically active Lig4 utilize Lig3 to facilitate joining in normal cells? Work is ongoing to address this possibility.

MATERIALS AND METHODS

Cell culture, genome editing, and survival assays. 293T, U2OS and HCT116 cells were cultured in Dulbecco's Modified Eagle Medium (Life Technologies) supplemented with 10% fetal bovine serum (Atlanta Biologicals, GA), 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U/ml penicillin, 100 µg/ml streptomycin (Life Technologies) and 10 µg/ml ciprofloxacin. CH12F3 cells were cultured in RPMI 1640 medium supplemented with 10% (vol/vol) FBS and 50 µM β-mercaptoethanol.

Cas9-targeted gene disruption was performed using methods similar to those reported by Mali et al. (48). Briefly, duplex oligonucleotides (Integrated DNA Technologies) generating a gRNA specific for a PAM site targeting the K273 codon were cloned into pCas2A-puro. Cells were transfected with 2ug plasmid and a 120 nucleotide oligonucleotide encoding a K273A mutation as well as a silent mutation generating a novel restriction site (HhaI) (see Supplemental Table I). 48 hours after transfection, cells were replated at cloning densities in media containing puromycin (1ug/ml). Puromycin was removed after 72 hours. Isolated clones were selected, and DNA isolated with DNAzol (Sigma) according to the manufacturer's protocol. Restriction digestion of PCR products was utilized to detect the K273A mutation. Western blotting was used to confirm expression and genotypes were confirmed by Sanger sequencing or Amplicon sequencing (Genewiz) as depicted in Supplemental Table 1.

MTT staining was performed to assess cell viability for both 293T cells and CH12. 30,000 to 50,000 cells were plated in each well of a 24-well plate, containing medium with varying concentrations of zeocin. After 5 to 7 days of calicheamicin treatment, cells were treated with 1

mg/ml MTT (Sigma) solution for 1 hr. Medium containing MTT was then removed and formazan crystals thus produced were solubilized in acidic isopropanol. Absorbance was read at 570 nm to determine relative survival.

Episomal end-joining assays. The fluorescent VDJ coding, signal, and alt-VDJ substrates have been described (21,38,49). The Fill-in substrate was generated by synthesizing (Integrated DNA Technologies) a fragment encoding crimson with the addition of restriction endonuclease sites Eco53KI, PspOMI, and Apa1 that disrupt the Crimson open-reading frame. Cleavage with these restriction enzymes generate blunt and over-hanged ends as described in Figure 11D. The Crimson fragment included NheI and BamHI which were used to subclone this fragment between the promoter and GFP open reading frame in the substrate plasmids. Briefly, extrachromosomal fluorescent joining assays were performed on cells plated at 20-40% confluency into 24-well plates in complete medium. Cells were transfected with 0.125 μg substrate, and either 0.25 I-Sce1, TelN, or RAG 1+2 expression plasmids per well using polyethylenimine (PEI, 1 ug/mL, Polysciences) at 2 μL/1 μg DNA. Cells were harvested 72 hours after transfection and analyzed for GFP and RFP expression by flow cytometry. The percentage of recombination was calculated as the percentage of live cells expressing GFP divided by the percentage expressing RFP. Data presented represents at least three independent experiments, which each includes triplicate transfections.

In figure 11, the coding joint substrate was modified so that the coding flanks included only A or T sequences; in addition, the 23RSS was inverted to provide for inversional joining which facilitates PCR amplification of coding joints. Transfected plasmids were isolated by alkaline lysates 72 hours after transfection. Coding joints from each transfection were PCR amplified and analyzed by electrophoresis and amplicon sequencing by Genewiz.

Class switch recombination assays. CSR assays were performed as described previously (37). Briefly, cells were seeded in the presence of 1 μg/mL anti-CD40 antibody (16-0402-86; bioscience), 5 ng/mL of IL-4 (404-ML; R&D Systems), and 0.5 ng/mL TGF-β1 (R&D Systems 240-B) and grown for 72 h. Cells were stained with a FITC-conjugated anti-mouse IgA antibody (BD Bio-sciences 559354) and analyzed on a LSR II flow cytometer (BD Biosciences). CSR efficiency is determined as the percentage of IgA-positive cells.

Amplicon sequencing of Crispr/Cas9-induced chromosomal DSBs. We have used a Crispr/Cas9 strategy to sequence DSBs from within the FancG locus previously (41). To induce deletional DSBs, two gRNA/Cas9/puro plasmids were transfected into 293T cells with the indicated genotypes. After 48 hours, cells were place in puromycin selection media. Cells were harvested and DNA prepared 72 hr later and PCR performed to detect chromosomal deletions. PCR fragments were isolated and subjected to Amplicon sequencing provided by Genewiz.

Mammalian expression vectors and recombinant protein expression constructs and purification. Expression plasmids and purification procedures for XRCC4-WT, XLF-WT, Lig4-WT, and Lig4-K273A have been described (41).The LIG4/XRCC4 complex was produced as described (14). All proteins batches were dialyzed against 150 mM KCl, 20 HEPES pH8, 1 mM EDTA, 2 mM DTT and 10% (v/v) glycerol, flash frozen in liquid nitrogen and stored at -80˚C. T4 DNA ligase was obtained from New England Biolabs. LIG4 fragments 1-620 (WT or K273A), DBD, NTase, and OBD were prepared as described (50). Mammalian expression constructs for wild-type Lig4 and Lig4-K273A were generated by subcloning from plasmids described above. The 5XLys mutant was prepared by subcloning a geneblock (IDT) encoding K273A, K449R, K451R, K352R, and K345R via PflMI/BlpI subcloning.

Ligation Assay. Ligation assays were performed in a volume of 10 µL containing 50 ng of linearized pUC19 plasmid DNA (with XbaI for cohesive ends and SmaI for blunt ends), 1 mM ATP, 1 mM DTT, 20 mM Tris-HCl pH8.0, 2 mM $MgCl_2$, and 60 mM KCl, 1% (v/v) glycerol with addition of proteins at the indicated concentrations. Reactions were incubated at room temperature for 30 min before addition of Proteinase K at 1.4 mg/mL final concentration and 1X of a loading buffer containing 0.01% SDS (NEB #B7024S) and incubated for 10 min at 55°C. Samples were next resolved by 0.8% agarose gel electrophoresis in Tris-Borate-EDTA buffer. Gels were stained in Tris-Borate-EDTA buffer supplemented with 0.5 µg/mL ethidium bromide and destained in deionized water. Images where captured under UV light using a Bio-Rad chemidoc and quantitatively analyzed with Image J.

Oligonucleotides and antibodies. Oligonucleotides used in this study are as follows; only one strand of the oligonucleotides used for targeting PAM sites are presented, and are without BMHI overhangs.

For 293T Lig4 CRISPR experiments:

gRNA Lig4: CATACGTTCACCATCTAGCT

Lig4 K273A HDR:

CTATTGCAGATATTGAGCACATTGAGAAGGATATGAAACATCAGAGTTTCTACATAGAAACAGCGCTAG

ATGGTGAACGTATGCAAATGCACAAAGATGGAGATGTATATAAATACTTCTCTCGAAATGG

For U2OS CRISPR experiments:

gRNA Lig4: GTTCAGCACTTGAGCAAAAG.

The U20S clone used in the experiments had +1 frameshift mutations on both alleles.

For 293T polQ CRISPR experiments:

gRNA1 polQ:TGAAGCGGGTTTTGGAAATG

gRNA2 polQ:TCTGATCAATCGCCTCATAG

For 293T FANCG CRISPR experiments:

gRNA1 FancG:GGGCCAGGCCTGGGTTCAAC

gRNA2 FancG:GACTTAAGAGAAAGGGACTG

5’ FancG PCR: CCCAAGATGTCCCGGCTGTGGG

3’FancG PCR:CCATGGGCCTCTCTGTCCTTGCAC

Oligonucleotides for substrate PCR:

5’ coding joint PCR: CGGTGGGAGGTCTATATAAGCA

3’ coding joint PCR:CTACACCGTGGTGGAGCAGTA

5’ signal joint PCR:ACCTTGAAGCGCATGAAGGGC

3’ signal joint PCR:TCCATGCGGTACTTCATGGTC

Antibodies utilized in this study include: anti-DNA Lig 4 (Abcam, 26039), anti-Ku80 (Invitrogen, 111), anti-DNA-PKcs (generous gift Tim Carter).

Data availability: The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials; the amplicon sequencing files from this study are available from the corresponding author.

Funding: USDA National Institute of Food and Agriculture [1019208]; Public Health Service [AI048758, AI147634 to K.M.] and [AI 38345, AI 39039 to KY]; French National Research Agency

(ANR-17-CE12-0020-01, ANR-18-CE29-0003-04) and French National League against cancer (équipe labellisée) to M.M.

Acknowledgements: We are particularly indebted to both Dale Ramsden and Adam Luthman for their generous and invaluable assistance in analyzing amplicon sequencing projects for differences in junctional diversity and utilization of short sequence homologies.

Contribution by the author (CMB dissertation requirement): This work contains work from multiple authors. As first author, Noah Goff performed the bulk of the data collection, analysis, and contributed to the conceptualization/writing/editing of this manuscript—with the exception of Figure 9 and supplemental figures 1-3. Note the supplemental figures are available in the online version of this manuscript at Nucleic Acids Research.

REFERENCES

1. Reid, D.A., Keegan, S., Leo-Macias, A., Watanabe, G., Strande, N.T., Chang, H.H., Oksuz, B.A., Fenyo, D., Lieber, M.R., Ramsden, D.A. *et al.* (2015) Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E2575-2584.

2. Reid, D.A., Conlin, M.P., Yin, Y., Chang, H.H., Watanabe, G., Lieber, M.R., Ramsden, D.A. and Rothenberg, E. (2017) Bridging of double-stranded breaks by the nonhomologous end-joining ligation complex is modulated by DNA end chemistry. *Nucleic acids research*, **45**, 1872-1878.

3. Zhao, B., Watanabe, G., Morten, M.J., Reid, D.A., Rothenberg, E. and Lieber, M.R. (2019) The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis. *Nature communications*, **10**, 3588.

4. Jackson, S.P. (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis*, **23**, 687-696.

5. Yoo, S. and Dynan, W.S. (1999) Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic acids research*, **27**, 4679-4686.

6. Graham, T.G., Walter, J.C. and Loparo, J.J. (2016) Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular cell*, **61**, 850-858.

7. Chen, S., Lee, L., Naila, T., Fishbain, S., Wang, A., Tomkinson, A.E., Lees-Miller, S.P. and He, Y. (2021) Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature*, **593**, 294-298.

8. Chaplin, A.K., Hardwick, S.W., Stavridi, A.K., Buehl, C.J., Goff, N.J., Ropars, V., Liang, S., De Oliveira, T.M., Chirgadze, D.Y., Meek, K. *et al.* (2021) Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular cell*, **81**, 3400-3409 e3403.

9. Chaplin, A.K., Hardwick, S.W., Liang, S., Kefala Stavridi, A., Hnizda, A., Cooper, L.R., De Oliveira, T.M., Chirgadze, D.Y. and Blundell, T.L. (2021) Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nature structural & molecular biology*, **28**, 13-19.

10. Wang, J.L., Duboc, C., Wu, Q., Ochi, T., Liang, S., Tsutakawa, S.E., Lees-Miller, S.P., Nadal, M., Tainer, J.A., Blundell, T.L. *et al.* (2018) Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nat Struct Mol Biol*, **25**, 482-487.

11. Stinson, B.M., Moreno, A.T., Walter, J.C. and Loparo, J.J. (2020) A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular cell*, **77**, 1080-1091 e1088.

12. Waters, C.A., Strande, N.T., Pryor, J.M., Strom, C.N., Mieczkowski, P., Burkhalter, M.D., Oh, S., Qaqish, B.F., Moore, D.T., Hendrickson, E.A. *et al.* (2014) The fidelity of the ligation step

determines how ends are resolved during nonhomologous end joining. *Nature communications*, **5**, 4286.

13. Budman, J., Kim, S.A. and Chu, G. (2007) Processing of DNA for nonhomologous end-joining is controlled by kinase activity and XRCC4/ligase IV. *The Journal of biological chemistry*, **282**, 11950-11959.

14. Cottarel, J., Frit, P., Bombarde, O., Salles, B., Negrel, A., Bernard, S., Jeggo, P.A., Lieber, M.R., Modesti, M. and Calsou, P. (2013) A noncatalytic function of the ligation complex during nonhomologous end joining. *The Journal of cell biology*, **200**, 173-186.

15. Conlin, M.P., Reid, D.A., Small, G.W., Chang, H.H., Watanabe, G., Lieber, M.R., Ramsden, D.A. and Rothenberg, E. (2017) DNA Ligase IV Guides End-Processing Choice during Nonhomologous End Joining. *Cell reports*, **20**, 2810-2819.

16. Chiruvella, K.K., Liang, Z., Birkeland, S.R., Basrur, V. and Wilson, T.E. (2013) Saccharomyces cerevisiae DNA ligase IV supports imprecise end joining independently of its catalytic activity. *PLoS genetics*, **9**, e1003599.

17. Lu, H., Shimazaki, N., Raval, P., Gu, J., Watanabe, G., Schwarz, K., Swanson, P.C. and Lieber, M.R. (2008) A biochemically defined system for coding joint formation in V(D)J recombination. *Molecular cell*, **31**, 485-497.

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

19. Goodarzi, A.A., Yu, Y., Riballo, E., Douglas, P., Walker, S.A., Ye, R., Harer, C., Marchetti, C., Morrice, N., Jeggo, P.A. *et al.* (2006) DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *The EMBO journal*, **25**, 3880-3889.

20. Riballo, E., Kuhne, M., Rief, N., Doherty, A., Smith, G.C., Recio, M.J., Reis, C., Dahm, K., Fricke, A., Krempler, A. *et al.* (2004) A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Molecular cell*, **16**, 715-724.

21. Meek, K. (2020) Activation of DNA-PK by hairpinned DNA ends reveals a stepwise mechanism of kinase activation. *Nucleic acids research*, **48**, 9098-9108.

22. Helmink, B.A., Tubbs, A.T., Dorsett, Y., Bednarski, J.J., Walker, L.M., Feng, Z., Sharma, G.G., McKinnon, P.J., Zhang, J., Bassing, C.H. *et al.* (2011) H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature*, **469**, 245-249.

23. Rooney, S., Sekiguchi, J., Zhu, C., Cheng, H.L., Manis, J., Whitlow, S., DeVido, J., Foy, D., Chaudhuri, J., Lombard, D. *et al.* (2002) Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. *Mol Cell*, **10**, 1379-1390.

24. Roth, D.B., Menetski, J.P., Nakajima, P.B., Bosma, M.J. and Gellert, M. (1992) V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. *Cell*, **70**, 983-991.

25. Liu, L., Chen, X., Li, J., Wang, H., Buehl, C.J., Goff, N.J., Meek, K., Yang, W. and Gellert, M. (2021) Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Molecular cell*.

26. Kabotyanski, E.B., Gomelsky, L., Han, J.O., Stamato, T.D. and Roth, D.B. (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic acids research*, **26**, 5333-5342.

27. Cui, X. and Meek, K. (2007) Linking double-stranded DNA breaks to the recombination activating gene complex directs repair to the nonhomologous end-joining pathway. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 17046-17051.

28. Sallmyr, A. and Tomkinson, A.E. (2018) Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. *The Journal of biological chemistry*, **293**, 10536-10546.

29. Buck, D., Moshous, D., de Chasseval, R., Ma, Y., le Deist, F., Cavazzana-Calvo, M., Fischer, A., Casanova, J.L., Lieber, M.R. and de Villartay, J.P. (2006) Severe combined immunodeficiency and microcephaly in siblings with hypomorphic mutations in DNA ligase IV. *European journal of immunology*, **36**, 224-235.

30. Chen, S.H. and Yu, X. (2019) Human DNA ligase IV is able to use NAD+ as an alternative adenylation donor for DNA ends ligation. *Nucleic acids research*, **47**, 1321-1334.

31. Dedon, P.C. and Goldberg, I.H. (1992) Free-radical mechanisms involved in the formation of sequence-dependent bistranded DNA lesions by the antitumor antibiotics bleomycin, neocarzinostatin, and calicheamicin. *Chem Res Toxicol*, **5**, 311-332.

32. Wu, C.C., Li, T.K., Farh, L., Lin, L.Y., Lin, T.S., Yu, Y.J., Yen, T.J., Chiang, C.W. and Chan, N.L. (2011) Structural basis of type II topoisomerase inhibition by the anticancer drug etoposide. *Science*, **333**, 459-462.

33. Neal, J.A. and Meek, K. (2019) Deciphering phenotypic variance in different models of DNA-PKcs deficiency. *DNA repair*, **73**, 7-16.

34. Roth, D.B. and Wilson, J.H. (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Molecular and cellular biology*, **6**, 4295-4304.

35. Corneo, B., Wendland, R.L., Deriano, L., Cui, X., Klein, I.A., Wong, S.Y., Arnal, S., Holub, A.J., Weller, G.R., Pancake, B.A. *et al.* (2007) Rag mutations reveal robust alternative end joining. *Nature*, **449**, 483-486.

36. Oh, S., Harvey, A., Zimbric, J., Wang, Y., Nguyen, T., Jackson, P.J. and Hendrickson, E.A. (2014) DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells. *DNA repair*, **21**, 97-110.

37. Masani, S., Han, L., Meek, K. and Yu, K. (2016) Redundant function of DNA ligase 1 and 3 in alternative end-joining during immunoglobulin class switch recombination. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 1261-1266.

38. Deriano, L., Stracker, T.H., Baker, A., Petrini, J.H. and Roth, D.B. (2009) Roles for NBS1 in alternative nonhomologous end-joining of V(D)J recombination intermediates. *Molecular cell*, **34**, 13-25.

39. Meek, K. (1990) Analysis of Junctional Diversity during Lymphocyte-B Development. *Science*, **250**, 820-823.

40. Lewis, S. and Gellert, M. (1989) The mechanism of antigen receptor gene assembly. *Cell*, **59**, 585-588.

41. Normanno, D., Negrel, A., de Melo, A.J., Betzi, S., Meek, K. and Modesti, M. (2017) Mutational phospho-mimicry reveals a regulatory role for the XRCC4 and XLF C-terminal tails in modulating DNA bridging during classical non-homologous end joining. *Elife*, **6**.

42. Kienker, L.J., Shin, E.K. and Meek, K. (2000) Both V(D)J recombination and radioresistance require DNA-PK kinase activity, though minimal levels suffice for V(D)J recombination. *Nucleic acids research*, **28**, 2752-2761.

43. Kurimasa, A., Kumano, S., Boubnov, N.V., Story, M.D., Tung, C.S., Peterson, S.R. and Chen, D.J. (1999) Requirement for the kinase activity of human DNA-dependent protein kinase catalytic subunit in DNA strand break rejoining. *Molecular and cellular biology*, **19**, 3877-3884.

44. Lu, H., Saha, J., Beckmann, P.J., Hendrickson, E.A. and Davis, A.J. (2019) DNA-PKcs promotes chromatin decondensation to facilitate initiation of the DNA damage response. *Nucleic acids research*, **47**, 9467-9479.

45. Jiang, W., Crowe, J.L., Liu, X., Nakajima, S., Wang, Y., Li, C., Lee, B.J., Dubois, R.L., Liu, C., Yu, X. *et al.* (2015) Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. *Molecular cell*, **58**, 172-185.

46. Meek, K., Douglas, P., Cui, X., Ding, Q. and Lees-Miller, S.P. (2007) trans Autophosphorylation at DNA-dependent protein kinase's two major autophosphorylation site clusters facilitates end processing but not end joining. *Molecular and cellular biology*, **27**, 3881-3890.

47. Roy, S., de Melo, A.J., Xu, Y., Tadi, S.K., Negrel, A., Hendrickson, E., Modesti, M. and Meek, K. (2015) XRCC4/XLF Interaction Is Variably Required for DNA Repair and Is Not Required for Ligase IV Stimulation. *Molecular and cellular biology*, **35**, 3017-3028.

48. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.

49. Neal, J.A., Xu, Y., Abe, M., Hendrickson, E. and Meek, K. (2016) Restoration of ATM Expression in DNA-PKcs-Deficient Cells Inhibits Signal End Joining. *Journal of immunology*, **196**, 3032-3042.

50. Ochi, T., Gu, X. and Blundell, T.L. (2013) Structure of the catalytic region of DNA ligase IV in complex with an Artemis fragment sheds light on double-strand break repair. *Structure*, **21**, 672-679.

Chapter 3: New insight into how the DNA binding domain of DNA ligase IV facilitates end-joining, independent of its catalytic activity.

By:

Noah J. Goff, Saanjh Sundar, Addy Walia, Gargi Parkhi, Katheryn Meek

Data Chapter

ABSTRACT

DNA double strand breaks (DSBs) are highly genotoxic lesions generated throughout the cell cycle by cellular processes (DNA tension, replication errors, metabolic stress) or exogenous agents (ionizing radiation, chemotherapeutic drugs). Cells have several pathways to resolve DSBs and mitigate their insult to genomic integrity. End-joining pathways, most notably non-homologous end joining (NHEJ), are the most readily available DSB repair mechanisms—being available throughout the cell cycle and capable of repairing a wide range of damage adducted DNA-ends. Recent studies have revealed a two-stage synaptic mechanism for NHEJ, where DNA Ligase IV (L4) acts both catalytically to ligate the broken strands and as an essential structural factor that drives progression from long-range synaptic complexes (LRCs) into a ligation-capable short-range synaptic complex (SRC). Here we report a novel cryo-EM structure containing density for L4's DNA binding domain (DBD) in the LRC, something not previously seen in any reported structures. The L4 side of that interface is shared with a trans interaction between L4 and Ku70 in the SRC. We characterize how the L4 DBD promotes end-joining in the absence of L4 catalytic activity through interactions in the SRC with DNA and Ku70. Each of these interactions are essential to the capacity of catalytically inactive L4 to promote joining. Moreover, we have discovered that catalytically inactive L4 requires a specific LRC to promote end-joining. Finally, we have observed several instances of a single structural interface in one protein mediating interactions with different NHEJ components, and we discuss potential implications of this finding.

INTRODUCTION

Preservation of genetic material is paramount to long-term health and survival of all organisms. Threats to genomic stability either result from common environmental agents or are byproducts of essential cellular functions, including reactive byproducts of metabolism, torsional stress within the nucleus, autonomous nuclear elements, and some viruses among other sources. Failure of DNA repair pathways in eukaryotes often results in harmful lesions that can be converted into mutations, impair cellular function, contribute to development of cancers, or eventually drive cell death. DNA **d**ouble **s**trand **b**reaks (DSBs) are considered the most genotoxic lesions—with inappropriate repair resulting in a range of outcomes from small deletions to large chromosomal rearrangements (79–81, 52, 45).

Direct end-joining is the primary repair mechanism available throughout the cell cycle and is primarily mediated through the canonical **n**on-**h**omologous **e**nd-**j**oining pathway (NHEJ) (82). A general model of NHEJ begins with sensing damage, synapsis of two ends, processing of chemical adducts to produce compatible 3’ hydroxyl and 5’ phosphate groups and completion of the process by direct ligation of the DNA ends. The core NHEJ factors in mammals (and many other eukaryotes)(35) include the **DNA** dependent **p**rotein **k**inase [DNA-PK, consisting of the DNA end-binding heterodimer, (Ku70/Ku80), and the **c**atalytic **s**ubunit DNA-PKcs] that acts as a damage sensor, structural factors XRCC4, **X**RCC4-**L**ike **F**actor (XLF) and **pa**ralog of **X**RCC4 and **X**LF (PAXX), and the NHEJ-specific DNA **l**igase **IV** (L4) that performs the final ligation step. NHEJ’s availability throughout the cell cycle addresses a central challenge of cellular DNA repair: how to repair lesions when a homologous template is rarely available? As a result, NHEJ is both remarkably flexible and fast, but increases the risk of small indels or larger chromosomal instability. Recent work in the field has developed a more detailed two-stage synaptic model of NHEJ, where ends are synapsed and protected from aberrant end processing in a “long-range” synaptic complex, essential end processing is performed and required downstream factors are recruited before ultimately transitioning into a “short-range” complex where ends become accessible to either L4 or further end-processing by X-family DNA **pol**ymerases **l**ambda and **m**u (Poll and Polm), the phosphodiesterase TDP1, and the nucleases Artemis, MRN, and others (21, 69, 83, 84).

Work from Loparo and colleagues first proposed a two-stage model of DNA end synapsis by NHEJ factors (46). Their single molecule FRET studies revealed both **s**hort-**r**ange and long-**r**ange synapses where ends are held in either direct proximity (**s**hort-**r**ange **c**omplex) or ~115Å apart (**l**ong-**r**ange **c**omplex) termed SRC and LRC respectively. Transition into a short-range complex required the presence of Ku70, Ku80, XLF, XRCC4, L4, DNA-PKcs kinase activity (but not necessarily its presence), and critically L4, but not its catalytic activity (46). Further FRET and biochemical studies from these authors revealed specific roles for Poll mediated fill-in synthesis and TDP1-dependent phosphodiesterase in the SR complex (21).

Interestingly, while Loparo and colleagues worked on single molecule imaging, new structural data from several groups revealed diverse NHEJ complexes consisting of DNA-PK bound to a single DNA end or in complex with other NHEJ factors (XLF, XRCC4/L4, PAXX) to tether two ends together in a dimeric structure. Remarkably, these new NHEJ structures revealed dimeric complexes with DNA ends synapsed in direct contact, in good agreement with the SR complex predicted by the smFRET approaches, as well as multiple distinct iterations of LR synaptic complexes, positioning the two ends ~115Å apart (18, 22, 70).

DNA ends synapsed by the LR complexes can be classified into two groups: 1) Ku80-mediated dimers dominated by *trans* interactions between DNA-PKcs and the extreme C-terminal tail of Ku80, and 2) XLF-mediated dimers dominated by a pair of DNA-PKcs interfaces interacting reciprocally in *trans* along with a centrally positioned XLF dimer interacting with two dimers of XRCC4 (XRCC4-XLF-XRCC4) providing multiple, protein-protein interfaces that stabilize the DNA synapse. Our recent structure-function mutagenesis studies suggest that LR dimer complexes have distinct functions in promoting NHEJ. Specifically, the Ku80-mediated dimer promotes access to DNA ends for nucleolytic end processing, while the XLF-mediated dimer limits nucleolytic processing—matching the “end protection” complex suggested by Loparo and colleagues (20). Our current limited understanding of how NHEJ progresses stems from all of these studies (with all the unique approaches utilized); active areas of research now focus on understanding how LR complexes translate between the two dimer forms and what is required to progress from LR complexes to SR complexes.

It is now well-appreciated that L4 has an essential role in promoting the transition to the short-range complex (10, 21, 46, 73). Recent studies have provided more information as to how L4 contributes to the transition from long-range to short-range end synapsis (20, 70). This function is unique among mammalian DNA ligases. The catalytic domains of the three human DNA ligases contain three motifs: a DNA binding domain (DBD), a nucleotidyltransferase domain (NTD), and an oligonucleotide oligosaccharide fold domain (OBD). DNA **l**igases **III** and **VI**, but not ligase **I** (L1, L3, and L4) also include BRCT repeats (39). In the case of L4, a small region of the polypeptide between its two BRCT domains mediates L4's interaction with its obligate co-factor XRCC4 (9). L4's catalytic core has generally not been observed in most of the recent Cryo-EM structures, with the notable exception of the SR complex when the catalytic domain is seen bound to a DSB (18–20). Of note, single molecule imaging experiments revealed that L4 dynamically interacts with each DNA end independently in the LRC; these interactions are mediated by positively charged residues in the DBD that were observed (in Cryo-EM structures of human L4 in the SRC) to make electrostatic interactions with the DNA backbone, slightly internal to the DNA termini (48). Cortes and colleagues using engineered chimeras of human DNA ligases revealed that chimeric proteins including the BRCT domain of L4 with L1's NTD and OBD, but not L1's DBD produces a fusion protein that supports NHEJ-specific VDJ recombination (85). The finding that fusing L1's entire catalytic core (including its DBD) with L4's BRCT-motifs was insufficient to promote ligation via NHEJ, directly implicates L4's DBD in facilitating L4's non-catalytic function in NHEJ.

More experiments that promote the conclusion that L4 has an important structural role in joining comes from our recent single molecule imaging study showing that whereas L4 is required for stable recruitment of DNA-PK to chromatin after DNA damage, catalytically inactive L4 supports its recruitment. Of note, retention of DNA-PK at DSBs in cells expressing catalytically inactive L4 is markedly longer than active L4; these data suggest that repair facilitated by catalytically inactive L4 is slower than when L4 is active (73).

In this study, we further investigate L4's structural role in promoting end-joining using NHEJ functional assays in genetically engineered cell strains and cryo-electron microscopy (Cryo-EM, data not shown). We confirm that repair facilitated by catalytically inactive L4 is much slower than with wild type L4, and that repair is partially dependent on DNA polymerase theta (Polθ).

Moreover, repair promoted by inactive L4 is highly dependent on the Ku80-mediated LRC. In addition, we report a new version of the Ku80-mediated LRC with density for L4's DBD interacting with DNA-PKcs. Of note, the interface that mediates L4's interaction with DNA-PKcs also facilitates a previously reported *trans* interaction between L4's DBD and Ku70 that tethers DNA ends in the SR complex (18). Functional assays establish the relevance of this L4 DBD/Ku70 interaction for end joining, whereas the interaction of L4's DBD with DNA-PKcs appears to be dispensable for NHEJ. This interface is distinct from the two DNA/protein interfaces studied by Stinson et al (48) that impact L4's tethering of DNA ends in single molecule FRET experiments. We posit that numerous interactions of the DBD with Ku and the two DNA termini are essential in promoting transition from LR to SR complexes where end-joining occurs.

RESULTS

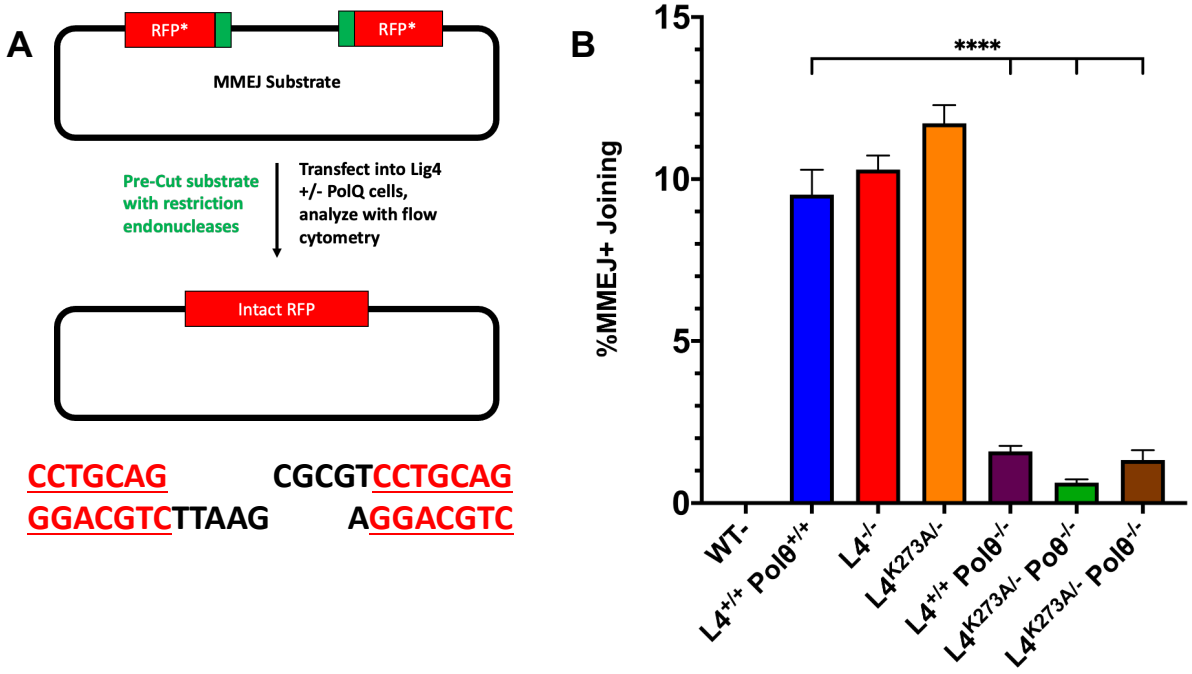


Figure 13: Microhomology mediated end joining is mediated through Polθ independent of NHEJ deficiency. A) Schematic of MMEJ fluorescent reporter assay, with 7 bp of microhomology highlighted and underlined. Substrates were pre-cut with ApoI and MluI (NEB) to generate noncomplementary overhangs, then cotransfected with a ECFP plasmid to assess transfection efficiency. Use of this microhomology is required to restore the full ds-red reading frame to allow MMEJ+ signal. MMEJ+ signal was assessed as % of transfected cells that expressed DsRed. **B)** MMEJ joining activity in edited 293T cell strains, independent of L4 presence or catalytic activity. Statistics performed using an unpaired two-tailed Student's t-test (**** $p > 0.0001$).

Microhomology-mediated end joining (MMEJ) is dependent on Polθ. We recently showed that catalytically inactive L4 promotes substantial end-joining and radioresistance in cell culture models, bolstering earlier reports documenting a non-catalytic role for L4 in NHEJ. Our study suggests that joining in cells expressing inactive L4 [a K273A mutation] is likely mediated by L3, but in the context of an undefined NHEJ complex (10). This cooperation between L4 and L3 is strongly corroborated by a recent collaborative study of catalytically inactive L4 in mice that lack nuclear L3 (14). In a separate publication, analyses of NHEJ factor recruitment to DNA DSBs in living cells revealed that whereas in cells lacking L4, NHEJ factor recruitment (using Halo-Flag tagged Ku70, DNA-PKcs, and XRCC4) is transient (in the case of Ku70 and DNA-PKcs) and absent for XRCC4; but in cells expressing the K273A mutant, although initial NHEJ factor recruitment occurs similarly as in cells expressing active L4, dissociation of NHEJ factors is markedly stalled (73).

The increased retention of NHEJ factors in cells expressing L4 K273A suggests that repair may occur less rapidly than in NHEJ-proficient cells where end-joining is rapid, and (when DNA ends are ligatable and compatible) generally accurate. Sequencing across rejoined DSBs in cells expressing catalytically inactive L4 revealed lower fidelity repair outcomes with an increased dependence on ligating ends across short homologous sequences near the break (1-5 bp, frequently referred to as microhomology). This is characteristic of end-joining facilitated by Polθ, which is notably slower than NHEJ (and therefore only happens at low levels in NHEJ proficient cells) . Additionally, rejoined DSBs from cells expressing L4 K273A share characteristics of joints mediated by Polθ (increased nucleotide loss and increased use of terminal microhomology) (15). To assess the role of Polθ in cells expressing L4 K273A, Polθ was ablated using a CRISPR/Cas9 strategy; the impact of Polθ loss specifically on repair of DSBs via terminal microhomology was tested using a linearized plasmid substrate that requires use of a 7-bp short-sequence homology internal to the break site to restore an RFP open reading frame. All cells proficient in Polθ, regardless of NHEJ proficiency, support basal levels of MMEJ. In contrast, MMEJ joining is remarkably reduced in Polθ ablated cells expressing either wild-type L4 or catalytically inactive L4 (Figure 13B). [This joining substrate reports only on MMEJ mediated joints, explaining the independence from NHEJ activity.] Consistent with previous reports (77, 86, 87), Polθ mediated

end joining functions independently from end joining catalyzed by NHEJ. We conclude that Polθ is required for MMEJ mediated joining, and that catalytically inactive L4 cannot facilitate MMEJ in the absence of Polθ.

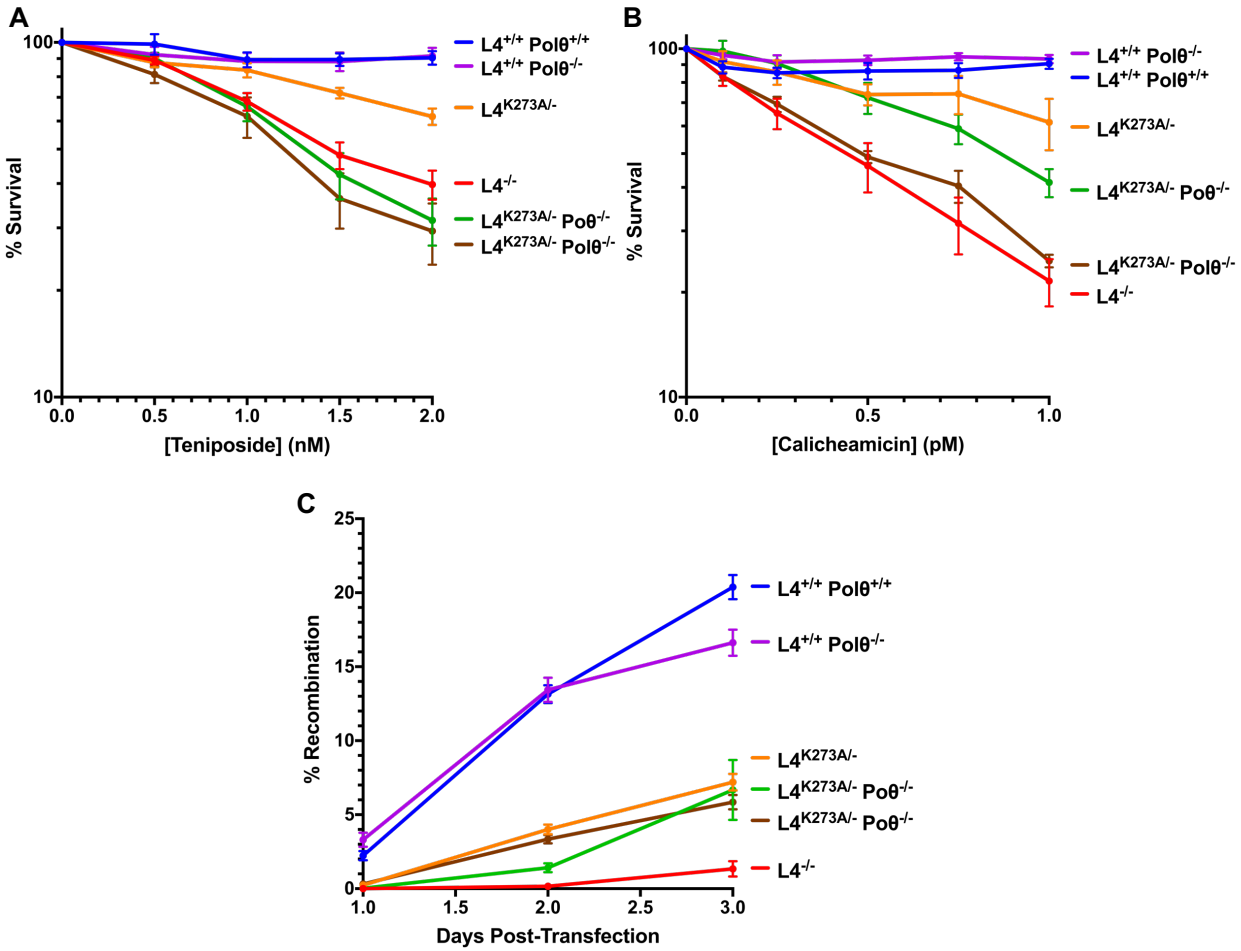


Figure 14: Catalytically inactive L4 supports end joining through multiple alt-EJ pathways. A) Sensitivity of 293T cells to teniposide as assessed by MTT assay. **B)** Sensitivity of 293T cells to calicheamicin as assessed by MTT assay. **C)** Time-course version of a VDJ episomal assay, with fixed cells representing joining after 24, 48, or 72 hours as analyzed by flow cytometry.

In cells expressing catalytically inactive L4, resistance to DSB-inducing agents requires Polθ. Since Polθ promotes substantial amounts of MMEJ, we reasoned that MMEJ may play a more prominent role in the alternative end joining promoted by cells expressing L4 K273A. Cells deficient in Polθ were not sensitive to either calicheamicin (that generates DSBs with 3' overhangs with 3' phosphoglycolate adducts) or teniposide (a type II topoisomerase poison that leaves 5'-peptide adducted DSB overhangs primarily in S phase). However, L4 K273A expressing cells that

lack Polθ, are markedly sensitive to both drugs, approaching the level of sensitivity observed in cells that completely lack L4 (Figure 14A-B). Because Polθ cannot facilitate sufficient end-joining in L4 deficient cells (as evidenced by a synthetic lethal interaction between loss of Polθ and NHEJ), these data suggest that Polθ may promote end-joining in the context of a catalytically inactive L4 complex.

VDJ coding joining facilitated by catalytically inactive L4 is slower than wild type NHEJ and is not dependent on Polθ. Time course assays were performed to assess the rate of VDJ coding end joining (Figure 14C). As shown previously, cells expressing K273A L4 join RAG-induced DSBs (in this case coding ends) much more efficiently than cells lacking L4 (10). However, joining in the K273A cells is delayed compared to joining from cells expressing active L4. Finally coding joint efficiency or rate is not impacted by loss of Polθ in cells expressing either wild type or K273A L4. Altogether, these data suggest that K273A L4 contributes to a variety of repair mechanisms, including both Polθ mediated alt-EJ and L3-mediated joining [as shown by our recent collaborative study (14)].

End-joining facilitated by K273A L4 requires the Ku80-mediated long-range NHEJ complex. As noted above, accumulating data support a new model whereby NHEJ functions by initially synapsing two DNA ends in LR complexes positioning the ends ~115 Å apart. A fraction of LRCs transition to SR complexes that juxtapose ends close enough to facilitate ligation (46). Emerging cryo-EM data have dramatically bolstered and expanded this model providing structural evidence for a variety of distinct LRCs, but only one SRC (18–20). What is lacking is a clear understanding of how complexes transition from one form to another, or if catalytically inactive L4 is dependent on a specific complex. The slow rate of joining in cells expressing K273A L4 as well as the delayed dissociation of NHEJ factors from chromatin in cells expressing K273A L4 might suggest stalling in a particular complex.

To test the impact of L4 inactivity when DNA-PK mutation disrupts normal transition between LR and/or SR complexes, a series of episomal end-joining assays were performed in a 293T K273A cells strain that was ablated for DNA-PKcs expression using a CRISPR strategy. In these assays, the interplay between DNA-PKcs and L4 can be ascertained by complementing the DNA-PKcs deficiency using mutant constructs that disrupt either the Ku80-mediated dimer (4A) or the

XLF-dependent dimer (898/2569) or both (4A/2569, Figure 15). In these experiments VDJ coding and signal joining assays were performed using a previously described hyper-RAG2 mutant that induces substantially higher levels of recombination; this accentuates differences in end-joining capacity. As can be seen, K273A cells support substantial VDJ coding and signal end-joining in cells complemented with wild type DNA-PKcs (although still less than observed in cells with wild type L4). As reported previously, DNA-PKcs mutants disrupting either of the two LR complexes have modest deficiencies in joining in cells with wild type L4 (20). When complementing the DNA-PKcs deficiency with the 898/2569 mutant that disrupts the XLF-dependent dimer, signal and coding end-joining are proportionally reduced in cells with K273A L4, as compared to cells with wild type L4. In contrast, both signal and coding end-joining in cells expressing inactive L4 is virtually ablated in complementation experiments with DNA-PKcs mutants that disrupt the Ku80-mediated dimer (4A and 4A/2569, Figure 15). These data suggest that interactions mediated by the domain swap dimer are required to facilitate joining promoted by inactive L4.

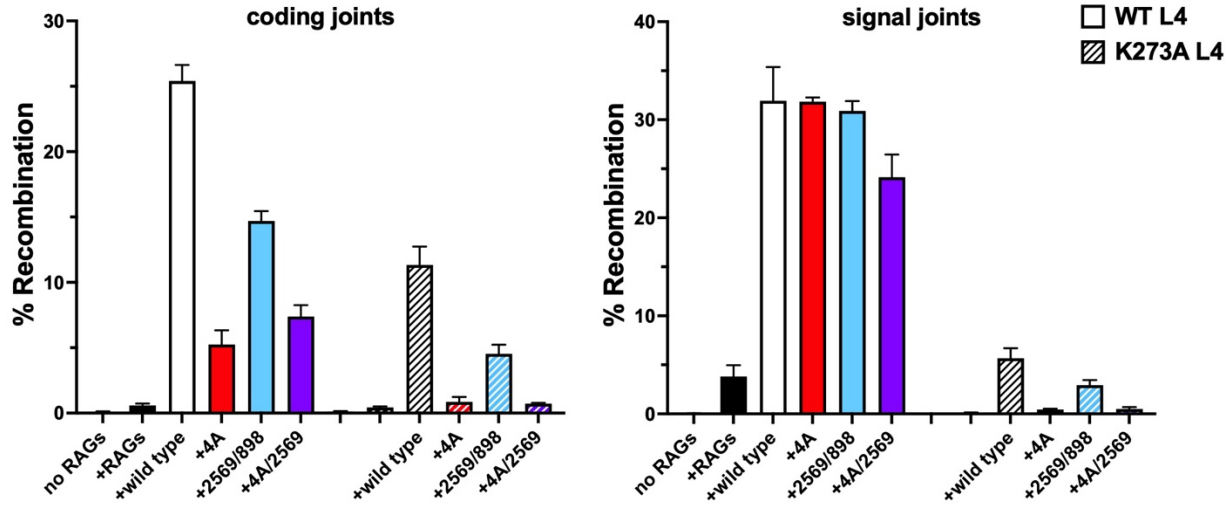


Figure 15: End-joining facilitated by K273A L4 requires the Ku80-mediated long-range NHEJ complex. A-B) VDJ coding end (left) and signal end (right) joining assays in 293T PKcs$^{-/-}$ L4$^{+/+}$ (solid bars) and PKcs$^{-/-}$ L4$^{K273A/-}$ (striped bars).

L4's DNA binding domain (DBD) interacts with the M-HEAT region of DNA-PKcs in the Ku80-mediated dimer. During refinement of recent NHEJ complexes, a density previously not appreciated was identified as a part or L4's DBD (Chloe L. Hall, Steven W. Hardwick, and Amanda K. Chaplin. Data not shown). Of note, L4's DBD has not previously been observed in any of the

different forms of the LRC (18, 19). This interaction appears to be mediated by two highly conserved positively charged residues R136 and K140. L4 R136 interacts with D1440 in DNA-PKcs whereas L4 K140 interacts with DNA-PKcs at residues 1495-1498. There is also a potential interaction of a loop within the FAT domain (3197-3226) that interacts with L4's active site. Finally, there is a potential interaction of L4 K264 with Ku70 T307. This density for L4 and its interaction with DNA-PKcs has only been observed in the Ku80-mediated dimer.

The DNA-PKcs interface that interacts with L4's DBD is not required for efficient NHEJ. To ascertain the functional relevance of this DNA-PKcs/L4 interaction, expression constructs ablating the interacting interfaces in the two proteins were generated. In DNA-PKcs, residues D1440 and E1497 were changed to arginine and lysine respectively (DE>RK). In L4, residues R136 and K140 in the DBD were substituted with aspartic acid (RK>DD). Mutant DE>RK DNA-PKcs was tested in transient episomal end-joining assays in the DNA-PKcs deficient 293T cell line described above. As can be seen, DNA-PKcs DE>RK restores end-joining at or above WT levels (in this case VDJ coding joints) as well as wild type DNA-PKcs (Figure 16B). Because we considered that the tethering of L4 might be relevant to the LR to SR transition (that is possibly altered in cells with inactive L4), end-joining was also tested in the K273A expressing 293T cells with CRISPR-ablated DNA-PKcs (described above, Figure 16C). DE>RK mutant DNA-PKcs restores joining to a similar level as wild type DNA-PKcs, albeit at a reduced level compared to cells expressing active 293T cells. We next generated stable cell strains expressing either wild type, DE>RK, or no DNA-PKcs in the V3 CHO cell strain that lacks DNA-PKcs (Figure 16D). While cells lacking DNA-PKcs are hyper-sensitive to both calicheamicin and teniposide, cells expressing DE>RK mutant DNA-PKcs are similarly resistant to these DSB-inducing agents as wild type DNA-PKcs. We conclude that the interface in DNA-PKcs that interacts with L4's DBD in the Ku80-mediated dimer is not required for cellular resistance to DSB-inducing drugs or for rejoining DSBs in episomal end-joining assays.

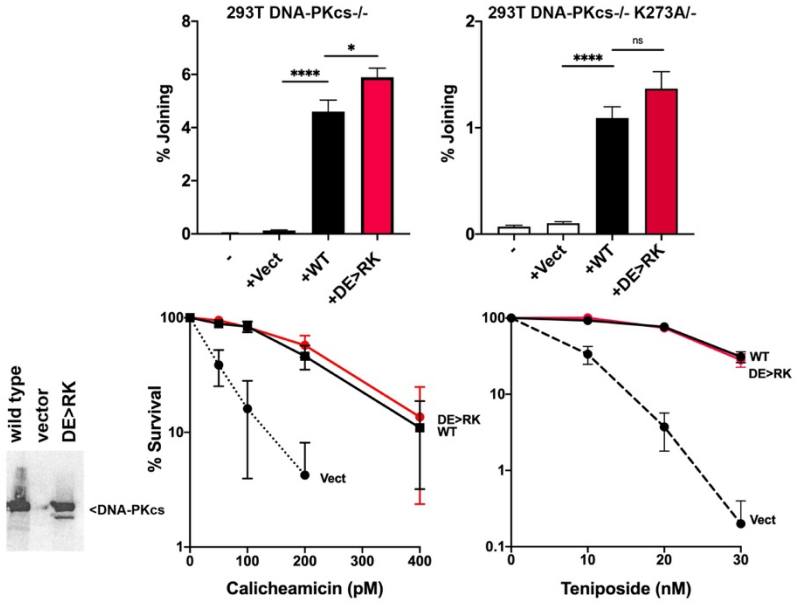


Figure 16: The DNA-PKcs interface that interacts with L4's DBD is not required for efficient NHEJ. A) 293T Western with transient complementation of Vect, WT, DE>RK in both $Lig4^{+/+}$ and $Lig4^{K273A*/-}$ **B)** 293T joining assays, VDJ coding PK22 left c195-14 right **C)** Western assessing DNA-PKcs expression in clonal populations of CHO-V3 cells complemented with empty vector (vect), WT DNA-PKcs (WT) or DNA-PKcs DE>RK (DE>RK) **D)** Colony formation assays testing sensitivity to Calicheamicin (left) and teniposide (right).

The L4 interface that interacts with DNA-PKcs is important for efficient NHEJ in cells with active L4, but essential for NHEJ in catalytically inactive L4. Expression vectors encoding L4 with only the RK>DD substitutions were prepared; in addition, the RK>DD substitution was combined with additional L4 mutations targeting its catalytic activity [including the K273A mutation that ablates the major adenylation site, and also K273A with five additional lysine>arginine substitutions (5xK) that ablate a small level of "back-up" adenylation observed in *in vitro* experiments with the K273A mutant] (10). To test whether the DBD interface that can interact with DNA-PKcs impacts joining, episomal end-joining assays as described above were performed. Consistent with our previous study, in the absence of L4, virtually no VDJ coding end joining is observed (Figure 17A-C); in contrast, wild type L4 (WT), as well as catalytically inactive L4 mutants (K273A, 5xK mutants) all facilitate substantial coding end joining. When L4 is active, the RK>DD mutant promotes substantial coding end joining, albeit reduced compared to active L4. However, when the RK>DD mutation is combined with inactive L4, joining is virtually ablated.

It is well-established that the RAG1/2 endonuclease directs its DSBs into the NHEJ pathway by a poorly understood mechanism (reference). To address whether the RK>DD mutant impacts joining of DSBs that are not exclusively repaired by NHEJ, joining of DSBs induced by both TelN (a DNA hairpin forming phage protelomerase) and I-SceI (an endonuclease that produces noncomplementary, 4 base pair 3' overhangs) were studied. We previously found that cells expressing inactive L4 can join non-RAG induced DSBs more efficiently than RAG DSBs (10). Similar to VDJ-coding hairpins, cells expressing the RK>DD mutant are modestly deficient in joining TelN induced DSBs; the deficiency is exacerbated when the RK>DD substitution is combined with loss of catalytic activity (Figure 17D). Interestingly, cells expressing the mutant combining inactive L4 with the RK>DD substitution are also severely deficient in joining non-hairpin I-SceI breaks despite the increased level of joining reported previously with inactive L4 (12).

Stable complementation of wild type and mutant L4 was performed utilizing a PiggyBac retrotransposon system to stably introduce L4 into L4 deficient 293T and U2OS cell lines. To test whether cells expressing the RK>DD mutant are also more sensitive to agents that induce genomic DSBs, stable cell strains expressing wild type or mutant L4 were treated with calicheamicin and teniposide. Cells expressing inactive L4 (either K273A or 5xK) or the RK>DD mutant L4 are more resistant to calicheamicin and teniposide than cells lacking L4. Consistent with the end-joining assays, L4 mutants that combine the RK>DD substitution with mutations that disrupt catalytic activity are as sensitive to calicheamicin as cells lacking L4 (Figure 17B). We conclude that R136 and K140 are critical for end-joining; the observation that the impact of mutation of these residues potentiates deficits in joining mediated by catalytically inactive L4 suggests that interactions mediated by these residues are important for L4's non-catalytic role in repair via NHEJ.

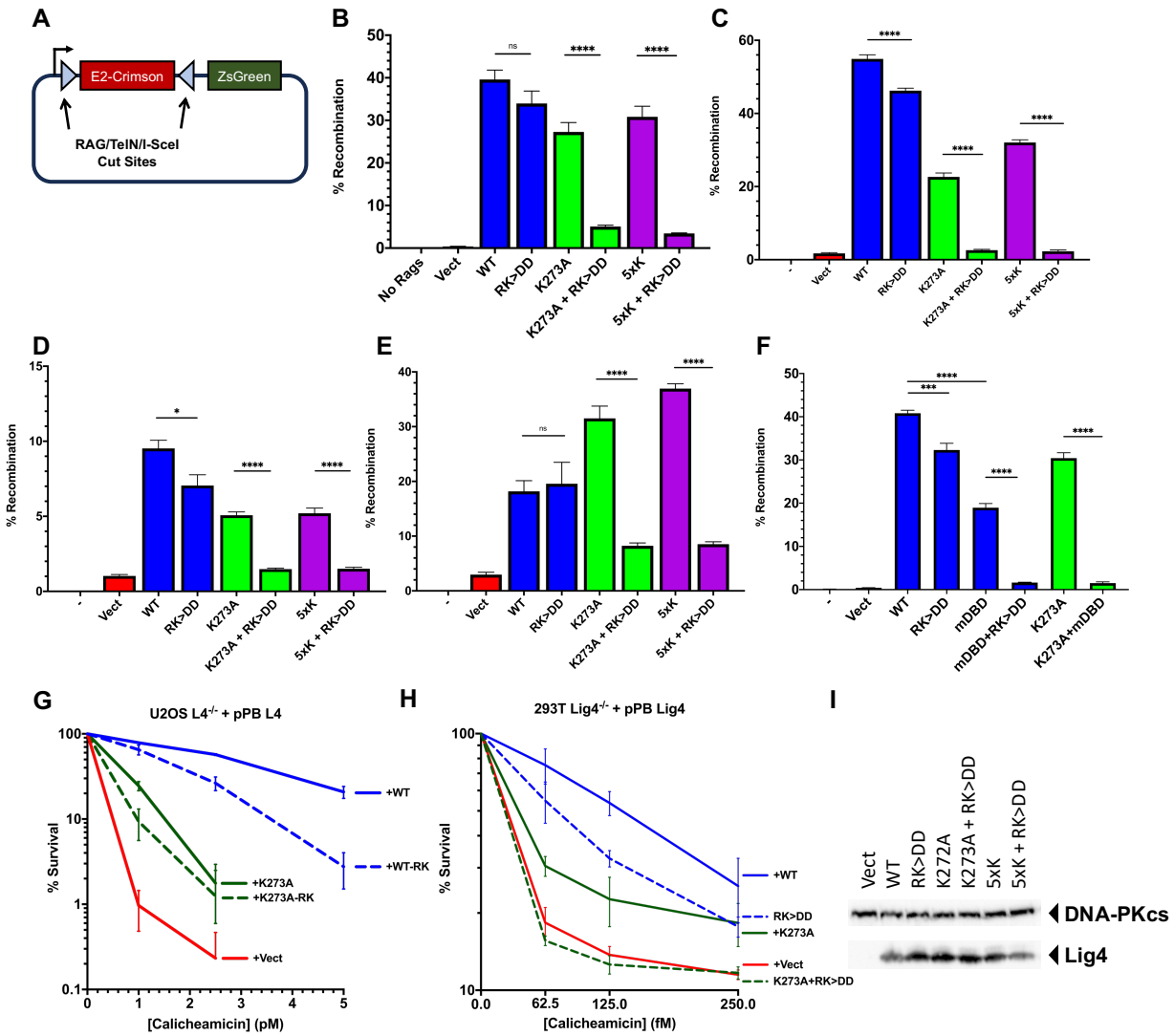


Figure 17: The L4 interface that interacts with DNA-PKcs is important for efficient NHEJ in active L4, but essential for NHEJ in catalytically inactive L4. **A)** Diagram of the joining assay substrate, with cut sites . Substrates were cotransfected into cells along with and without enzyme (+/-) and appropriate L4 expression vector, then red/green fluorescence via flow cytometry at 72 hours. **B)** Results from the VDJ coding- and **C)** signal-end substrates when cotransfected with Rag1 and hyper mutant Rag2 along with the corresponding L4 expression vector. **D-E)** TelN/I-SceI substrates cotransfected with endonuclease (-) or with appropriate endonuclease and corresponding L4 expression vector. **F)** VDJ coding-end joining assays as before testing additional mutations in the L4 DBD **G)** Sensitivity of 293T L4-/- cells complemented with PiggyBacc-expressed L4 mutants to calicheamicin as assessed by MTT assay, compared to 293T $L4^{+/+}$ (termed 293T WT). **H)** Sensitivity of PiggyBacc-edited U2OS L4 cell strains as assessed by clonogenic survival assays. **I)** Immunoblot showing transient L4 expression in a 293T L4 -/- cell strain with DNA-PKcs levels as a control.

The L4 interface that interacts with DNA-PKcs in the Ku80-mediated dimer forms a highly conserved salt-bridge interaction with Ku70 in the short-range complex. Multiple Sequence

Alignments performed by COBALT revealed substantial conservation of R136 and K140, from numerous vertebrate species (Figure 18). To clarify the discrepancies of mutating these two interacting interfaces in L4 and DNA-PKcs, we searched published structures (18, 19, 22) of multimeric NHEJ complexes for interactions between L4 and structural factors that could explain how disruption of R136 and K140 might impact NHEJ. R136 and K140 are only structured in two available structures of NHEJ complexes: the interaction with DNA-PKcs described above, and in the short-range complex reported by He and colleagues (18). In this publication, L4 residues 134-143 were observed interacting with the von Willebrand domain of Ku70. Specifically, R136 in L4's DBD bridges directly with D192 in Ku70 in a *trans* interaction (Figure 17). We posit that L4's ability to tether two DNA ends is mediated by numerous interactions including 1) protein/DNA interactions involving two patches of positively charged residues described by Loparo and colleagues (48), 2) protein/DNA interactions with L4's catalytic site, 3) the *trans* interaction between L4 R136 and K140 with Ku70 D192 (18), and finally 4) the capacity to form the Ku80-mediated dimer. Whereas disruption of any one of these interactions results in only a partial loss of end-joining and radio-resistance, loss of two results in severe deficits in both.

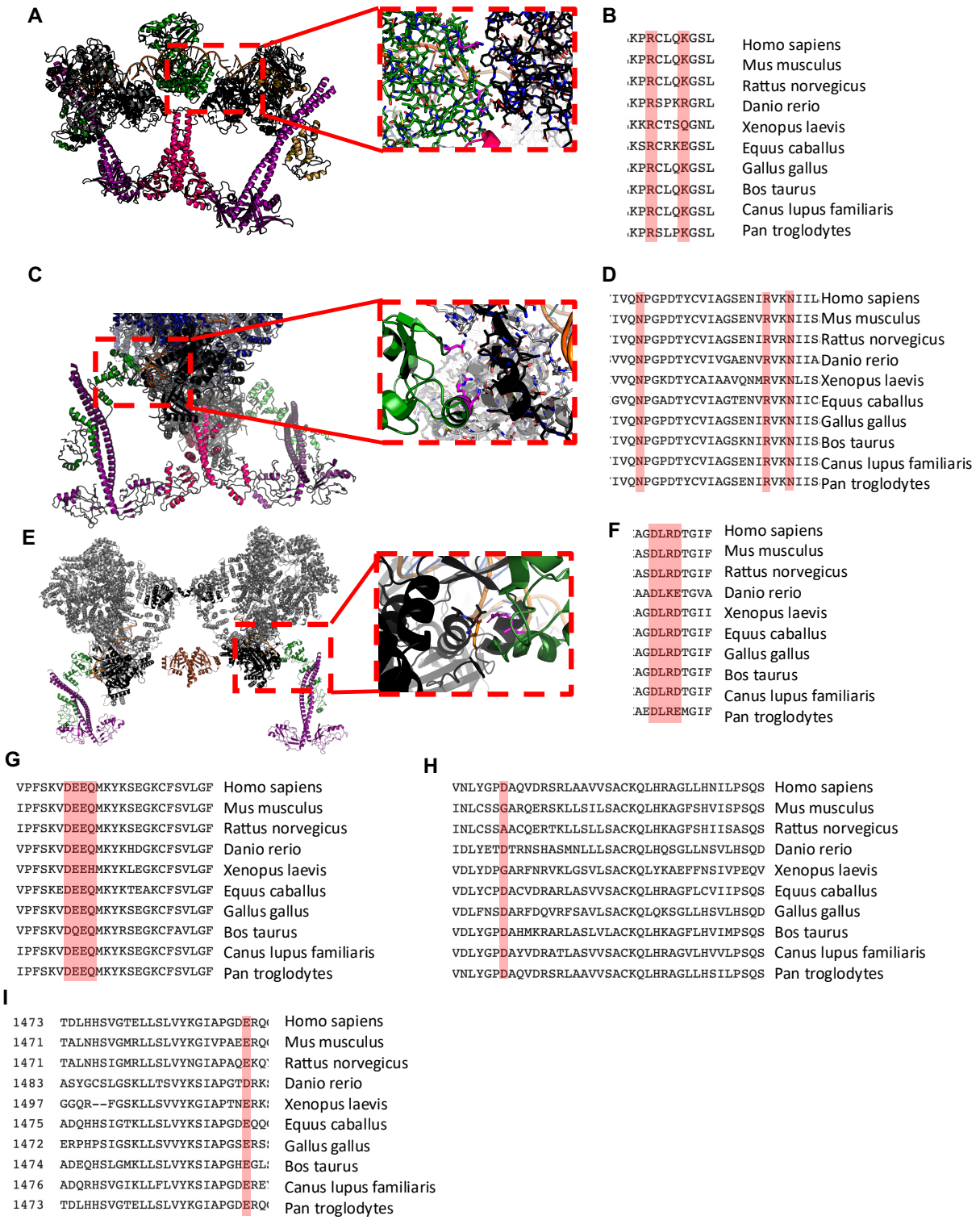


Figure 18: Identifying conserved shared interfaces between NHEJ factors. A) Cryo-EM structure of the Short-Range NHEJ dimer (PDB: 7LSY) highlighting an interface between the ligation-positioned Lig4 molecule (green) and Ku70 (black). Lig4 residues R136 and K140 are highlighted in magenta and form the basic side of a salt bridge with Ku70. Also pictured Ku80 (black), …

Figure 18 (cont'd)
XRCC4 (Purple), XLF (pink), and the BRCT-structural domain of a second Lig4 molecule (gold). **B)** Species conservation of Lig4 orthologs in 10 vertebrate species aligned to human Lig4 K134–L143. **C)** Cryo-EM structure of the XLF-dependent DNA-PK long-range NHEJ dimer highlighting an interaction between Lig4 and Ku70/Ku80. Zoomed in diagram showing Lig4 N692, R708, N711 (magenta) interacting with Ku80. **D)** Species conservation of Lig4 orthologs in 10 vertebrate species aligned to human Lig4 K134–L143. **E)** Shared interface between Ku80 and Lig4-BRCT in the Ku80-mediated dimer with PAXX/XRCC4/L4 (PDB 8BHY, Lig4 N711 highlighted in magenta). **F)** Species conservation of Ku70, with D192-D196 highlighted in red. **G)** Species conservation of Ku80 D327-Q330 "DEEQ motif" (red) aligned to human Ku80. **H)** Conservation of DNA-PKcs D1440 (red). **I)** Conservation of DNA-PKcs E1497 (red).

DNA-interacting residues in the L4 DBD work cooperatively with the Ku70 *trans* interface to join ends. A recent study by Loparo and colleagues revealed that L4 dynamically interacts with ends prior to SRC synapsis, mediated by 4 basic residues in its DBD that directly interact with DNA in cryo-EM structures. Mutating these interactions significantly impaired interactions with DNA and disrupted stable SRC formation. We hypothesized that combining the DNA-interface mutant with catalytically inactive L4 would impair its ability to promote end joining, similar to K273A + the RK>DD mutation targeting the L4 DBD-Ku70 interface. To address this, we designed an expression vector mutating the equivalent residues in human L4 (K28E, K30E, R32E, K162E, R163E, K164E, together termed mDBD) and tested the ability to join VDJ coding-end substrates (Figure 17F). In cells expressing catalytically active L4, mDBD decreases joining by ~50%, indicating that disrupting DNA-binding significantly (but not completely) impacts the ability of NHEJ to promote joining. Combining mDBD with either RK>DD or L4 K273A completely ablates all joining. This suggests that 1) ligation via L4 is structurally supported both by protein-DNA interactions and protein-protein interactions between L4's DBD and Ku70's VWA, and 2) that the alternative end-joining stimulated by catalytically inactive L4 is dependent on its structural interactions that stabilize the SRC.

DISCUSSION

A single protein interface may have multiple functions, potentially supporting flexibility in complex formation or driving transitions between distinct complexes. Here we report functional significance of a L4 protein-protein interaction recently identified in a new form of the Ku80 mediated LRC. To address the functional relevance of this interaction, a mutational approach was

utilized to ablate the interface by substituting residues in both DNA-PKcs and L4. In contrast to many of our previous studies where mutation of either side of a single interface generally results in similar cellular phenotypes, mutations designed to ablate either the L4 or DNA-PKcs side of the interface had very different cellular impacts. Of note, ablating the L4 surface impacted function, and this impact was enhanced when L4's catalytic activity was also blocked. After examining published structures of NHEJ complexes, we observed that the L4 residues involved (R136/K140) also mediate a *trans* interaction with Ku70 in the NHEJ short-range complex. This study suggests a model where the L4 interface that interacts with DNA-PKcs in the Ku80-mediated dimer may eventually transit (either directly or indirectly) into a more consequential interaction with Ku70 that we posit contributes to tethering DNA ends prior to ligation.

Of note, we have observed a similar dual function of another interface in DNA-PKcs. The four highly conserved lysine residues in the M-HEAT region of DNA-PKcs that facilitate both cis and trans interactions with Ku80's C-terminal acidic peptide (20), also interact with a distinct interface in the recently described LR-ATP complex reported by He and colleagues that likely represents a transitional LR complex (54). Finally, we have recently observed via cryo-EM experiments, Ku-bound to chromatinized DSBs (unpublished data, Amanda Chaplin). In this structure, we observe a specific interface in Ku80 that interacts with the H3 nucleosome subunit. However, this same interface in Ku80 also mediates its interaction with the BRCT domains of L4 observed in many NHEJ complexes (both long- and short-range). One explanation for these observations is that single interfaces that facilitate assembly of different complexes may dictate how transitions between NHEJ complexes proceed.

Lig4 plays an essential structural role in tethering DNA ends prior to ligation. Single molecule imaging studies from Loparo and colleagues demonstrate that interactions of L4's DBD are critical for progression from long-range to short-range complexes (48). More specifically, critical DNA-interacting residues in the L4's DBD dramatically reduced L4-DNA interactions in the LR complex and subsequent SR complex formation, indicating that this noncatalytic interaction between L4 and the break serves as an essential step preceding SR synapsis. Their data suggests a model where L4 dynamically interacts with ends protected by DNA-PKcs, and L4's binding of the two ends promotes progression to the SR complex. Our recent report established that catalytically

inactive L4 delays release of Ku, DNA-PKcs, and XRCC4 from chromatin after DNA damage in living cells (73); from these data, we posit that the catalytic site of L4 also promotes progression to short-range complexes. The description here of another L4 DBD interaction (R136 and K140 in L4 with D192 in Ku70) that functions collaboratively with the catalytic site may also impact L4's ability to promote progression to short-range complexes underscores the central role of L4 in this transition. These data prompted an examination of the same L4/DNA end interaction in the cellular assays used in this study. Altogether, these studies suggest that L4 may serve as a molecular “sensor” in the long-range complex, promoting transition into the ligation-competent short-range complex only if both DNA ends are appropriate substrates for ligation. This "sensor" function is dependent not only on L4's interaction with DNA ends, but also L4's interaction with Ku70, and L4's catalytic site.

L4 is unique among eukaryotic ligases in its characteristic as a single-turnover enzyme (51); thus, it is quite possible that the L4 molecule that seals one strand does not necessarily seal the second strand. One possibility is that rapid ligation of one strand and NHEJ-complex mediated end-tethering act cooperatively to support ends. Conversion of a DSB to a single strand break through one-step ligation would dramatically reduce the threat of the lesion and the conversion may allow NHEJ factors to “make way” for other DNA repair pathways to resolve the second strand’s break. In sum, these data underscore how the molecular dynamics of NHEJ complexes are highly dependent on L4 and its structural interactions.

Inactive L4 requires the Ku80-mediated dimer for function. We have shown previously that the two forms of long-range synaptic complexes have distinct functions and are in equilibrium. The observation that cells expressing K273A L4 have prolonged DSB-induced chromatin association, suggested to us that long-range to long-range or long-range to short-range transitions might be impacted by loss of L4's catalytic activity. This prompted an examination of whether either of the two LR complexes are functionally critical in cells expressing K273A L4. Our data demonstrate show that the Ku80-mediated dimer is essential for K273A L4 to promote end-joining, and we suggest that K273A L4 promotes alternative end-joining in the context of the Ku80 mediated dimer.

L4's essential non-catalytic role that facilitates end-joining is mediated by distinct interactions in NHEJ short-range complexes and involves distinct alternative joining pathways. NHEJ complexes in living cells expressing K273A L4 have significantly different characteristics than cells expressing wild type L4. NHEJ factors in complex with K273A L4 display prolonged association with chromatin in response to damage, suggesting that although catalytically inactive L4 facilitates NHEJ complex formation, progression of those complexes is delayed (73).

Clearly, L4's structural interactions with other NHEJ factors, and not just its catalytic interactions with DNA are critical for promoting end-joining. But how does catalytically inactive L4 promote joining? The data presented here and previously demonstrate that distinct end-joining mechanisms are facilitated by the presence of inactive L4. We have shown previously that K273A robustly facilitates L3 mediated joining *in vitro*; these data are strongly corroborated by our recent collaborative study from Yu and colleagues demonstrating a synthetic lethal interaction between catalytically inactive L4 and loss of nuclear L3 in mice (14). Here we show that unlike loss of L4, inactivation of L4 is not synthetically lethal with Polq ablation. Moreover, in K273A L4 expressing cells, whereas VDJ joining is minimally impacted by loss of Polq, both MMEJ and resistance to DSB-inducing agents are strongly impacted by loss of Polq. From drug sensitivity and episomal joining assays, we show that alternative end-joining stimulated by catalytically inactive L4 is dependent on stable promotion of the SRC. Furthermore, since 293T $L4K^{273A/-}$ $Pol\theta^{-/-}$ and cells expressing L4 K273A + RK>DD are comparably sensitive to DSB inducing agents, we can infer that the high degree of SRC stability with catalytically inactive L4 reported by multiple groups (73) is responsible for promoting alt-EJ in the context of catalytically inactive L4. An emerging consensus in the end-joining field is that there are multiple alternative end-joining pathways. The data presented here suggest that catalytically inactive L4 promotes repair by several of these pathways, likely thorough its ability to promote sustained end-processing in the SRC.

Conclusions: All together, these studies further bolster a perhaps incontrovertible conclusion that L4 has multiple distinct structural roles in NHEJ. What is lacking is an understanding of 1) the molecular mechanism of transition between the two forms of long-range complexes, 2) the transition from long- to short-range complexes and 3) the stoichiometry of DNA ligases in the short-range complex—particularly for difficult ends that are not rapidly ligated.

MATERIALS AND METHODS

Cell culture and genome editing 293T and U2OS cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies) and Roswell Park Memorial Institute 1640 Medium (RPMI), respectively, supplemented with 10% fetal bovine serum (Atlanta Biologicals, GA), 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U/ml penicillin, 100 g/ml streptomycin and 25 ng/ml Gibco Amphotericin B (Life Technologies).

Transfections of 293T and U2OS cell lines were performed using unsupplemented RPMI media, polyethyleneimine (PEI, 1 µg/mL Polysciences), and plasmid DNA at a ratio of 100 µL RPMI:1 µg DNA:2 µg PEI. First, plasmid DNA and RPMI were well-mixed in a 1.5 mL microcentrifuge tube followed by addition of PEI. Transfection mixes were shaken briefly to thoroughly mix contents, then incubated at room-temperature for up to 30 minutes to allow PEI:DNA complexes to form before adding the reaction volume to freshly plated cells

Genome editing

PolQ deletion in 293T cell strains was performed via CRISPR/Cas9 by targeting cells with a pair of gRNAs spaced 2.1 kb apart in the polq locus to create a large deletion. Each gRNA was cloned into a pCas-2A-Puro (Addgene 62988), then 1 µg of each plasmid was transfected into a 293T $L4^{-/-}$ and $L4^{K273A/-}$ cell strains. Clonal populations were grown out for one week and screened via outside-outside PCR (to identify clones with large deletions at the polq locus). Candidates were then checked for presence of remaining WT allele.

pPB stable L4 complementation: Transfected 1.0E6 cells with 1.5 µg pPB-Puro-Lig4 payload and 1.5 µg pBase transposase plasmids. Cells were allowed to recover for 48 hours before selection with 1-2 µg/mL puromycin. Bulk populations were cultured in the presence of puromycin for the duration of the experiments.

Cell Survival assays: U2OS L4-/- cell strains complemented with PiggyBacc L4 mutants were assessed for resistance to DSB inducing drugs using colony formation assays. Briefly, 200-300 cells of each mutant strain were plated into 6 cm tissue culture dishes with increasing concentrations of drug and allowed to grow for 9 days. Plates were stained with 2% crystal violet solution (VWR) and counted. A minimum of 3 Replicates was performed for each strain at each concentration. 293T+L4 mutant cell strains were assessed using MTT-metabolic assays as a

marker for cell viability. 1.6E4 cells were plated in wells of a 24-well plate at increasing concentrations of calicheamicin in 1 mL of DMEM media. After 4 days of drug treatment, cells were treated with 1 mg/mL MTT (Goldbio) for 1 hr at standard culture conditions. MTT was removed and cells were dissolved in 200 µL 100% DMSO and absorbance was sampled on a plate reader at 570 nm wavelength. A well without cells treated with media + MTT was used as a background control subtracted from all absorbance values.

Episomal Joining Assays: 293T L4$^{-/-}$ cells were transfected following our PEI protocol with a mixture of previously reported episomal joining substrates (10, 11, 88, 89), endonucleases, and plasmid expression vectors encoding L4 mutants in a 1:1:1 ratio (by plasmid mass), then cultured for 72 hours before paraformaldehyde fixation before running on an Attune CytPix flow cytometer. MMEJ assays were similarly transfected into 293T cells with 1.0 µg linearized DNA substrate and 0.1 µg pECFP transfection marker, with 2.0 µg PEI, 100 µL RPMI and fluorescence analyzed 72 hours later gating for CFP+ transfected cells, with the subpopulation of dsRed+ cells representing MMEJ+ outcomes.

ACKNOWLEDGMENTS

We would like to thank Daniel Vocelle and the MSU Flow Cytometry core for operating and maintaining the Attune CytPix, supported by the Equipment Grants Program, award #2022-70410-38419, from the U.S. Department of Agriculture (USDA), National Institute of Food and Agriculture (NIFA).

REFERENCES

1. Waters,C.A., Strande,N.T., Wyatt,D.W., Pryor,J.M. and Ramsden,D.A. (2014) Nonhomologous end joining: A good solution for bad ends. *DNA Repair*, **17**, 39–51.

2. Loparo,J.J. (2023) Holding it together: DNA end synapsis during non-homologous end joining. *DNA Repair*, **130**, 103553.

3. Watanabe,G. and Lieber,M.R. (2023) The flexible and iterative steps within the NHEJ pathway. *Progress in Biophysics and Molecular Biology*, 10.1016/j.pbiomolbio.2023.05.001.

4. Amin,H., Zahid,S., Hall,C. and Chaplin,A.K. (2024) Cold snapshots of DNA repair: Cryo-EM structures of DNA-PKcs and NHEJ machinery. *Progress in Biophysics and Molecular Biology*, **186**, 1–13.

5. Goff,N.J., Mikhova,M., Schmidt,J.C. and Meek,K. (2024) DNA-PK: A synopsis beyond synapsis. *DNA Repair*, **141**, 103716.

6. Rothkamm,K., Kruger,I., Thompson,L.H. and Lobrich,M. (2003) Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol*, **23**, 5706–5715.

7. Lees-Miller,J.P., Cobban,A., Katsonis,P., Bacolla,A., Tsutakawa,S.E., Hammel,M., Meek,K., Anderson,D.W., Lichtarge,O., Tainer,J.A., *et al.* (2021) Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. *Progress in Biophysics and Molecular Biology*, **163**, 87–108.

8. Stinson,B.M., Moreno,A.T., Walter,J.C. and Loparo,J.J. (2020) A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell*, **77**, 1080-1091.e8.

9. Deshpande,R.A., Myler,L.R., Soniat,M.M., Makharashvili,N., Lee,L., Lees-Miller,S.P., Finkelstein,I.J. and Paull,T.T. (2020) DNA-dependent protein kinase promotes DNA end processing by MRN and CtIP. *Sci Adv*, **6**, 0922.

10. Deshpande,R.A., Marin-Gonzalez,A., Barnes,H.K., Woolley,P.R., Ha,T. and Paull,T.T. (2023) Genome-wide analysis of DNA-PK-bound MRN cleavage products supports a sequential model of DSB repair pathway choice. *Nature communications*, **14**.

11. Chang,H. and Lieber,M. (2016) Structure-Specific nuclease activities of Artemis and the Artemis: DNA-PKcs complex. *NUCLEIC ACIDS RESEARCH*, **44**, 4991–4997.

12. Graham,T.G.W., Walter,J.C. and Loparo,J.J. (2016) Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular Cell*, **61**, 850–858.

13. Chen,S., Lee,L., Naila,T., Fishbain,S., Wang,A., Tomkinson,A.E., Lees-Miller,S.P. and He,Y. (2021) Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature*, 10.1038/s41586-021-03458-7.

14. Chaplin,A.K., Hardwick,S.W., Liang,S., Kefala Stavridi,A., Hnizda,A., Cooper,L.R., De Oliveira,T.M., Chirgadze,D.Y. and Blundell,T.L. (2021) Dimers of DNA-PK create a stage for DNA double-strand break repair. *Nat Struct Mol Biol*, **28**, 13–19.

15. Seif-El-Dahan,M., Kefala-Stavridi,A., Frit,P., Hardwick,S.W., Chirgadze,D.Y., Maia De Oliviera,T., Britton,S., Barboule,N., Bossaert,M., Pandurangan,A.P., *et al.* (2023) PAXX binding to the NHEJ machinery explains functional redundancy with XLF. *Science Advances*, **9**, eadg2834.

16. Buehl,C.J., Goff,N.J., Hardwick,S.W., Gellert,M., Blundell,T.L., Yang,W., Chaplin,A.K. and Meek,K. (2023) Two distinct long-range synaptic complexes promote different aspects of end processing prior to repair of DNA breaks by non-homologous end joining. *Molecular Cell*, 10.1016/j.molcel.2023.01.012.

17. Goff,N.J., Brenière,M., Buehl,C.J., de Melo,A.J., Huskova,H., Ochi,T., Blundell,T.L., Mao,W., Yu,K., Modesti,M., *et al.* (2022) Catalytically inactive DNA ligase IV promotes DNA repair in living cells. *Nucleic Acids Research*, 10.1093/nar/gkac913.

18. Mikhova,M., Goff,N.J., Janovič,T., Heyza,J.R., Meek,K. and Schmidt,J.C. (2024) Single-molecule imaging reveals the kinetics of non-homologous end-joining in living cells. 10.1101/2023.06.22.546088.

19. Sallmyr,A., Rashid,I., Bhandari,S.K., Naila,T. and Tomkinson,A.E. (2020) Human DNA ligases in replication and repair. *DNA Repair*, **93**, 102908.

20. Grawunder,U., Zimmer,D. and Lieber,M.R. (1998) DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. *Current Biology*, **8**, 873–879.

21. Chaplin,A.K., Hardwick,S.W., Stavridi,A.K., Buehl,C.J., Goff,N.J., Ropars,V., Liang,S., Oliveira,T.M.D., Chirgadze,D.Y., Meek,K., *et al.* (2021) Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular Cell*, **0**.

22. Stinson,B.M., Carney,S.M., Walter,J.C. and Loparo,J.J. (2024) Structural role for DNA Ligase IV in promoting the fidelity of non-homologous end joining. *Nat Commun*, **15**, 1250.

23. Malashetty,V., Au,A., Chavez,J., Hanna,M., Chu,J., Penna,J. and Cortes,P. (2023) The DNA binding domain and the C-terminal region of DNA Ligase IV specify its role in V(D)J recombination. *PLOS ONE*, **18**, e0282236.

24. Medina-Suarez,D., Han,L., O'Reily,S., Liu,J., Wei,C., Brenière,M., Goff,N.J., Chen,C., Modesti,M., Meek,K., *et al.* (2024) DNA ligases 3 and 4 cooperate in vivo to facilitate DNA repair and organism viability.

25. Chan,S.H., Yu,A.M. and McVey,M. (2010) Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in Drosophila. *PLOS Genetics*, **6**, e1001005.

26. Wyatt,D.W., Feng,W., Conlin,M.P., Yousefzadeh,M.J., Roberts,S.A., Mieczkowski,P., Wood,R.D., Gupta,G.P. and Ramsden,D.A. (2016) Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell*, **63**, 662–673.

27. Yousefzadeh,M.J., Wyatt,D.W., Takata,K., Mu,Y., Hensley,S.C., Tomida,J., Bylund,G.O., Doublié,S., Johansson,E., Ramsden,D.A., *et al.* (2014) Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. *PLOS Genetics*, **10**, e1004654.

28. Chen,S., Vogt,A., Lee,L., Naila,T., McKeown,R., Tomkinson,A.E., Lees-Miller,S.P. and He,Y. (2023) Cryo-EM visualization of DNA-PKcs structural intermediates in NHEJ. *Science Advances*, **9**, eadg2838.

29. Chen,X., Ballin,J.D., Della-Maria,J., Tsai,M.-S., White,E.J., Tomkinson,A.E. and Wilson,G.M. (2009) Distinct kinetics of human DNA ligases I, IIIα, IIIβ, and IV reveal direct DNA sensing ability and differential physiological functions in DNA repair. *DNA Repair*, **8**, 961–968.

30. Meek,K. (2020) Activation of DNA-PK by hairpinned DNA ends reveals a stepwise mechanism of kinase activation. *Nucleic Acids Research*, **48**, 9098–9108.

31. Neal,J.A., Xu,Y., Abe,M., Hendrickson,E. and Meek,K. (2016) Restoration of ATM Expression in DNA-PKcs–Deficient Cells Inhibits Signal End Joining. *J.I.*, **196**, 3032–3042.

32. Deriano,L., Stracker,T., Baker,A., Petrini,J. and Roth,D. (2009) Roles for NBS1 in Alternative Nonhomologous End-Joining of V(D)J Recombination Intermediates. *MOLECULAR CELL*, **34**, 13–25.

Chapter 4: Discussing the two-stage synaptic model of NHEJ.

By

Noah J. Goff

Conclusions and Future Directions

INTRODUCTION

My goal with this chapter is to place each of the previous chapters more firmly in the context of the evolving synaptic model of NHEJ, highlight some outstanding questions in the field, and to propose experiments involving DNA Ligase IV to address those questions. The overarching themes of my dissertation research center around how DSB repair pathways—in particular end joining pathways—use structural features to promote repair while minimizing mutations. Of course, we are not the only group actively working on this problem, and I've highlighted several notable studies released over the course of my PhD studies that contribute to our most recent model of NHEJ and how it interacts with alternative end joining pathways.

DISCUSSION

Chapter 1: Evolving models of non-homologous end joining. The core of this chapter is a review of recent NHEJ literature that's contributed to an updated model of end synapsis and processing. This work spans early single molecule imaging (SMI) studies performed by Loparo and colleagues (21, 46), early Cryo-EM work from Tom Blundell, Yuan He, and Amanda Chaplin (18, 19), and an introduction to the role of DNA Ligase IV (L4). I've also reproduced sections of a review that I co-wrote with another graduate student, Mariia Mikhova, and our advisors Kathy Meek and Jens Schmidt. The entire review is quite large (too large to reproduce in total for this work) and focuses on the role of DNA-PKcs in the new synaptic models of NHEJ. The sections included were both primarily drafted by me and highly pertinent to understanding evolving models of NHEJ—a core throughline of both my dissertation work and other contributions during my PhD. There have been several other reviews from leading groups focused on other aspects of recent end joining research—including excellent reviews of recent Cryo-EM data (52, 53), and end-synapsis (80).

Chapter 2: Catalytically inactive DNA ligase IV promotes DNA repair in living cells. This project began as a straightforward CRISPR-Cas9 gene editing experiment coupled with an easy set of lab-standard reporter assays to onboard a first-year graduate student. We never expected to be the core project of my PhD. The initial 293T mutagenesis work in early spring of 2020, immediately before the onset of the COVID-19 pandemic. The central research question—does catalytically inactive L4 promote end joining in cell cultures—was inspired by papers from

Loparo and colleagues proposing a 2-stage synaptic model of NHEJ, where DNA ends were "protected" in a long range complex prior to recruitment of downstream factors *(21, 46)*. Importantly, the authors discovered that L4's catalytic function was dispensable for this result *in vitro,* indicating L4 held important noncatalytic function that could support otherwise highly regulated end processing. While other studies had suggested a non-catalytic role for L4 both *in vitro* and in yeast cells, it was unclear whether DSB repair in more complex eukaryotic cells would support any NHEJ after loss of L4's core catalytic function. Furthermore, the studies only tested the role of L4 in promoting processing mediated by DNA polymerase lambda and the phosphodiesterase TDP1.

In addition to studying NHEJ, our lab also has research projects into immunoglobulin development through VDJ recombination, which produces NHEJ-targeted DNA hairpin structures as recombination intermediates. Given the additional targeting of hairpins to NHEJ, we were quite surprised when the initial 293T joining assays revealed robust hairpin joining in cells lacking L4 catalytic activity. Notably, the initial *in vitro* work from Loparo and colleagues did not test DNA hairpin substrates in the biochemical xenopus with catalytically inactive L4. While revising the paper, we found a 2013 paper from Wilson and colleagues (90) that provided early evidence for catalytically inactive L4 supporting end joining in yeast. We were completely unaware of this paper at the time, but it suggests a remarkable conservation of NHEJ phenotypes across eukaryotes—even those lacking DNA-PKcs.

The primary takeaways from my first paper were: 1) that L4's structural presence does serve as a regulatory step in canonical NHEJ, 2) loss of L4 catalytic activity in the shortrange complex decreases the fidelity of end joining, and 3) that the L3-XRCC1 complex acts collaboratively with the core NHEJ factors *in vitro* to stimulate end-joining. To date, the role of a L3-NHEJ connection serves as a major extant question in the two-stage iterative model of NHEJ. An in-review (at time of writing) paper from Dr. Kefei Yu's laboratory established a catalytically inactive $L4^{K273S/K273S}$ and $L3^{nuc-/-}$ mouse lines, with each being viable through birth and adolescence. Notably, $L3^{nuc-/-}$ $L4^{K273S/K273S}$ double mutant offspring are nonviable, providing additional *in vivo* evidence that there is compensation from L3 following loss of L4's catalytic activity.

Chapter 3: New insight into how the DNA binding domain of DNA Ligase IV facilitates end joining independent of its catalytic activity. This project originated both as a direct follow-up to my first publication, but also a structure-directed investigation of new data from Amanda Chaplin and Steve Hardwick's group at the University of Leicester and Cambridge University to investigate interfaces discovered in new Cryo-EM data. It is my current area of active research. We have shown that disrupting end-tethering substantially impairs L4 K273A-mediated end joining, with only modest impacts on joining in the presence of catalytically active L4. I propose that tethering ends in the SRC occurs through four mechanisms: 1) the XRCC4:XLF:XRCC4 bridge linking each Ku70/80 heterodimer, 2) the *trans* L4 DBD:Ku70 interaction (disrupted by the RK>DD mutation) that anchors one end of the complex with the ligation positioned break, 3) the L4-DNA interaction recently reported by Loparo and colleagues (48), and 4) rapid ligation of one strand early following transition into the shortrange complex to covalently link DNA ends. As seen in chapter 3, disruption of any two of these interfaces appears to be sufficient to ablate end joining, either mediated by L4 catalytic activity or through promotion of alt-EJ pathways (as is the case with catalytically inactive L4).

In addition to investigating interfaces that could mediate L4's structural role in NHEJ, we also sought to investigate backup repair factors (in addition to L3-mediated joining) that promote end joining after failed NHEJ. Based on my observation that cells expressing catalytically inactive L4 overutilizes microhomology in L4 K273A cells (Chapter 2), we hypothesized that DNA-Polymerase Theta (Polθ) may play an outsized role in supporting NHEJ in cases where the ligation complex fails. To test, I used a CRISPR/Cas9 genome editing approach to knock out Polθ from my 293T $L4^{K273A*/-}$ cells, then tested their response to a MMEJ specific reporter substrate, NHEJ-specific VDJ Coding joint substrate, and sensitivity to DSB inducing calicheamicin and teniposide. Of note, I was unable to isolate any 293T $L4^{-/-}$ $Polθ^{-/-}$ double knockout cell strains despite multiple gRNA transfections. Based on previous experiments in the lab as well as other reports (77, 91, 92), we believe that Polθ serves as the primary alternative end joining factor following loss of NHEJ and as such double knockouts are nonviable—or nearly nonviable.

This hypothesis is supported by my experiments indicating that 293T $L4^{K273A/-}$ $Polθ^{-/-}$ cells are hypersensitive to DSB inducing agents calicheamicin and teniposide (Chapter 3). Anecdotally,

these cell cultures also divide much slower, comparable to L4$^{-/-}$ cultured under the same conditions. There is neither a growth disparity nor DSB drug sensitivity observed with loss of Polθ on its own. Despite the hypersensitivity to genomic DSBs, additional loss of Polθ appears to have no impact on VDJ coding, and only a modest deficiency in joining TelN and I-SceI substrates. A combination of two factors could explain this: 1) signal from episomal joining assays is measured as a binary on/off per cell and is not necessarily a fair measure of total break repair capacity for a given substrate so long as minimal joining is performed; 2) there may be an alternate repair mechanism that can join simple overhangs (like those generated by I-SceI and Artemis-opened hairpins) but not more complex adducts caused by calicheamicin and teniposide. I suggest that L3 may serve in this role based on Chapter 2 and the mouse models developed by Dr. Kefei Yu's laboratory.

Additional contributions to work that haven't been reproduced for my dissertation: In addition to the work presented in chapters 2 and 3, I have made several substantial contributions to other projects that have resulted in authorship, both published and in press. I've chosen to focus my dissertation on my Ligase IV projects, and as such I've not reproduced excerpts from these manuscripts (a full list of my contributions to papers can be found in Appendix 1).

Ku80 Linker Mutants: First, I'd like to highlight my contributions to testing the Ku80 C-terminal domain mutations published in Buehl et al. 2023 (20). This project began shortly after returning from COVID-19 lockdowns, with our lab working collaboratively with several structural biologists to use mutational strategies to test the functional relevance of synaptic NHEJ complex structures recently identified using Cryo-EM. These structures revealed a *trans* interaction between Ku80's acidic extreme C-terminal domain (CTD) and a basic patch on DNA-PKcs, mediated by an unstructured flexible linker in Ku80 that is highly conserved in length (but not primary sequence). The Ku80 CTD had previously been identified as a mediator of end joining (5–7), and mutating the DNA-PKcs residues that mediate this interaction (termed 4A) revealed a significant defect in nucleolytic end processing. We hypothesized that the highly conserved length of the flexible linker (93) observed in the Cryo-EM structures may be functionally relevant in promoting the *trans* interaction. It was observed by Dr. Wei Yang (NIH) that the Ku80 CTD could also interact with DNA-PKcs in *cis* if the linker stretches to its maximum length, and

she suggested that adding or removing residues may differentially impact function by biasing formation of the *cis* or *trans* interactions. We observed that lengthening the linker by 6-15 residues substantially reduced the amount of joining (>2 fold). Interestingly, shortening the linker 3 or 6 residues significantly increased NHEJ function both in episomal joining assays and in resistance to calicheamicin and etoposide (a Top2 poison). Further shortening the linker by 10, 13, or 15 residues progressively impaired NHEJ, with the shortest linker length mimicking complete deletion of the CTD. We concluded that the hypothesized *cis* interaction may have an inhibitory effect on synapsis, and there is a minimum linker length required to promote synapsis and NHEJ function. It's currently unclear why the flexible linker is conserved at its current length (as opposed to 3-6 AA shorter), it may be due to lower repair fidelity, or some other deficiency not identified in our testing.

Single Molecule Imaging of NHEJ Factors: The second contribution I would like to highlight is to Mariia Mikhova's fantastic study into the nuclear dynamics of NHEJ factors using live-cell Single Molecule Imaging (73), in particular my work generating L4 knockouts in their U2OS cell strains with halo-tagged NHEJ factors. These cells were then used to assess changes in Ku70, DNA-PKcs, XLF, and XRCC4 dynamics with complementation of L4 expression vectors for the revisions of her (in press) paper. Notable findings include: 1) XRCC4 is primarily localized to the nucleus but does not interact with chromatin in the absence of L4, 2) There appears to be a mechanism to actively remove Ku and DNA-PKcs if L4/XRCC4 recruitment fails, and 3) in the presence of catalytically inactive L4 all four tagged NHEJ factors have a greatly elongated window (compared to WT L4) where they interact with chromatin—including DNA-PKcs. This final point is particularly surprising given that DNA-PKcs departs chromatin significantly earlier than other factors. We originally hypothesized that DNA-PKcs departure corresponded with transition into the short-range complex—matching the single Cryo-EM structure reported by Yuan He and colleagues (18), but the *in vitro* work indicates that there is no defect in short-range complex formation or in end processing in the presence of catalytically inactive L4 (21, 46). In my opinion, it's likely that the Cryo-EM short-range complex structure is not necessarily the only possible conformation of NHEJ factors in the experimentally observed short-range complex. To wit, I believe that it may be possible for DNA-PKcs to remain present in some short-range

complexes, but it notably does depart the break earlier than other factors— possibly after single strand ligation to allow more room for processing the second strand.

Ongoing argument over the role of DNA-PKcs and an alternate model of synapsis: DNA-PKcs had previously thought to primarily be dominant in NHEJ in vertebrates— particularly in those that predominantly use VDJ recombination for immunoglobulin diversification. Central to this hypothesis was the absence of DNA-PKcs from several model organisms: most notably budding yeast (*S. cerevisiae*), *C. elegans*, and fruit flies (*D. melanogaster*). A study from Lees-Miller and colleagues searched genomic sequencing data across a wide range of species for potential DNA-PKcs orthologs looking to build a deeper understanding of how NHEJ evolved (35). Remarkably, the authors discovered candidate DNA-PKcs orthologs in genomes sequenced from a broad range of plant, fungi, and animal species—with high conservation of some well-studied features (including the ABCDE phosphorylation patch). This startling result suggests that DNA-PKcs is not a unique regulator of NHEJ that arose due to programmed recombination—but rather a more universal regulator of core end joining across nucleated life-forms.

There is an alternate model for NHEJ synapsis and ligation—primarily championed by Eli Rothenberg (NYU) and Michael Lieber (USC). They conclude—through their imaging approaches using purified NHEJ proteins—that DNA-PKcs is dispensable for most NHEJ function and tethering may occur through XRCC4-XLF filaments coating intact sections of DNA (they observe filaments of purified LX4 complexes coating dsDNA near a DNA end with super resolution microscopy). In their model, Ku70/80 recognizes one end, promotes filamentation upstream of the break, then Ku binds to a *trans* filament (bound to the opposite DNA end) to initiate synapsis, then proceeds to bring the ends into a short-range synapse (measured as direct FRET+ signal in their assays).

Notably, addition of DNA-PKcs reduces the number of FRET+ foci (decreasing their "pairing efficiency" quantification) but does substantially increase the size of each focus. The authors claim that the presence of DNA-PKcs only results in pairing of "clusters" of DSBs, and that their relatively low measurement of "normalized pairing efficiency" points to DNA-PKcs being dispensable for "simple" DSBs (60). I propose an alternate hypothesis where DNA-PKcs can act to synapse the DNA-stem loop structures used to cap their fluorescent oligonucleotides (the authors confirmed that purified Ku70/80 + LX4 was unable to synapse these hairpin loops, but never ran

the control in the presence of DNA-PKcs). The stem-loop structures closely resemble hairpin DNA intermediates present in VDJ recombination, and the resulting aggregate of tagged DNA donors would perfectly match the larger foci seen in all reactions included with DNA-PKcs. Furthermore, data from Mariia Mikhova's paper indicate that there is likely no role for XRCC4-XLF filaments given XRCC4 does not associate with breaks in the absence of L4 (73).

In my opinion, most of their reasoning regarding the role of DNA-PKcs is circular; the authors cite the data from *Reid et al. 2015* (60) as justification to not test if DNA-PKcs promotes synapsis in subsequent papers (42, 58, 59). They then use their entire body of work to argue against DNA-PKcs's involvement in most NHEJ events. Their experimental design is unable to independently visualize localization of the acceptor fluorophore in the absence of a donor in proximity for FRET (less than 100 Å with their fluorophores), therefore making it impossible to clearly observe long-range synapses. This approach is less powerful than the Loparo group's studies and misses synapsis consistent with the LRC (where ends are not held close enough for the FRET+ signal Rothenberg and colleagues utilized as their metric for synapsis). In their measurements, they observe that addition of purified Ku70/80, XLF, and L4-XRCC4 is sufficient to synapse ends using a "Normalized Frequency" of high-FRET foci. The authors argue for a model of NHEJ resolving "simple DSBs" without presence or activity of DNA-PKcs. Instead, they argue synapsis occurs through long XLF-XRCC4-L4 filaments that they observe "coating" DNA.

Studies on the timing of NHEJ are a major component of the new two-stage synaptic model: A recent report from Loparo and colleagues addressed the dynamics of L4 at the instant of transition into the short-range complex. They developed a three-color imaging system to simultaneously DNA end synapsis (consistent with the SRC) as well as DNA-protein interactions. Their data suggest that two molecules of L4 are recruited to a DSB prior to complex formation and both dynamically interact with DNA ends. Prior to short-range synapsis, one LX4 complex vacates the LRC seconds before they observe DNA-DNA interactions consistent with SRC formation (48). Unfortunately, the authors only tested substrates generated with blunt ligation-compatible DNA ends that require no additional end-processing. I believe it's likely that one molecule of L4 may act to catalyze single strand break repair—particularly given that L4 is pre-adenylated in a catalytically charged state it seems likely that this reaction may progress very

quickly. Testing a wider range of end joining substrates in the three-color smFRET experiment—particularly a substrate that requires short-range complex targeted end processing—would be highly informative.

Additional open questions remain regarding the role of DNA-PKcs in supporting the short-range complex. The initial report of a short-range complex observed without DNA-PKcs by Cryo-EM (18), coupled with preliminary data showing DNA-PKcs leaving chromatin earlier than other NHEJ factors in a SMI study (16) suggested to us that DNA-PKcs dissociated from the break site after initial synapsis and end processing. One striking result: DNA-PKcs remains bound to breaks for significantly longer with catalytically inactive L4. It's unclear whether DNA-PKcs's retention at breaks is mediated through a prolonged long range complex (something unsupported by the *in vitro* smFRET data) (21, 46) or if there is a conformation of the short-range complex that retains DNA-PKcs. I suggest the following model to explain the discrepancy in data: DNA-PKcs remains to stabilize the short-range complex until L4 can catalyze repair of one end. Finalizing catalysis allows for L4 and DNA-PKcs to dissociate from the break, leaving behind a significantly less hazardous break that can be repaired by the remaining NHEJ machinery (XRCC4, XLF, and Ku all remain bound for significantly longer than DNA-PKcs). This turnover of DNA-PKcs would allow it to "supervise" repair, then either recycle to bind another DSB or allow more room for extensive end processing on the remaining SSB (particularly in a case following ligation opposite another DNA lesion: gap, mismatch, aberrant base, etc.).

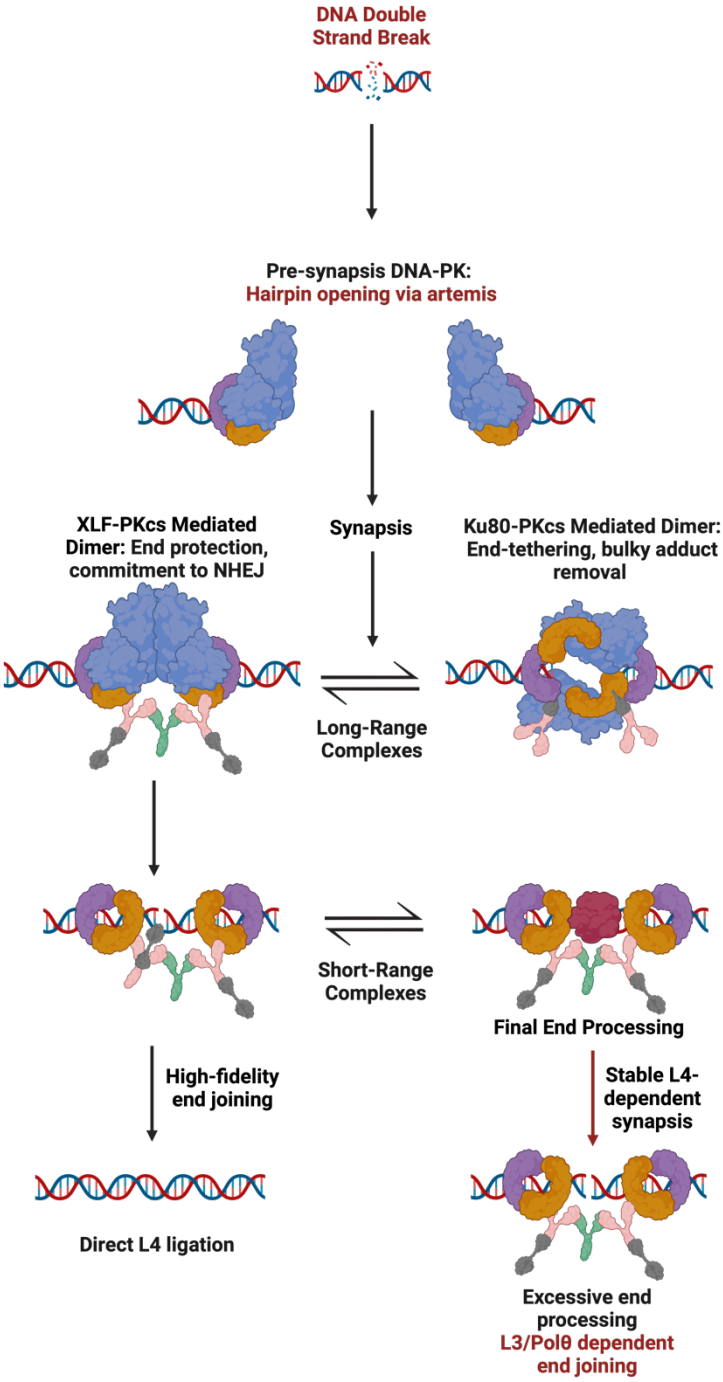


Figure 19: **NHEJ progresses through a two-stage synaptic mechanism.** Normal NHEJ progresses from break recognition (DNA-PK monomers) to LRC formation. Depending on end chemistry, either the end-protection XLF mediated dimer (for ends with small adducts/overhangs ends, middle left) or the Ku80-mediated dimer (ends with bulky adducts, middle right). The NHEJ synapse transitions into a SRC following recruitment of LX4 and XLF (and possibly formation of the XLF-mediated dimer). Loss of L4 catalytic activity promotes long-term stable SRC synapsis, leading to excessive end processing followed by Polθ/L3 mediated end joining (bottom right).

Figure 19 (cont'd)
Notably, loss of tethering activity (either of the mDBD or RK>DD mutations discussed in chapter 3) in the SRC in combination with losing L4 catalytic activity results in no L4-mediated alt-EJ. Furthermore, catalytically active L4 combining DNA end binding mutations (mDBD) with the L4-Ku70 interface mutant (RK>DD) ablates nearly all NHEJ.

Areas of future L4 research: While I believe the work in this dissertation outlines an important contribution to our understanding of NHEJ, there are several open questions regarding L4's role in end joining that I would like to see addressed in the future. First, additional work utilizing many of the newly identified mutants in Chapter 3 could explain why targeting the shared interfaces impact NHEJ so severely. I propose performing a suite of single-molecule imaging studies—identical to the calicheamicin "time course" studies using transiently complemented L4 that I contributed to in Mariia Mikhova's study (73)—particularly to study nuclear dynamics of XRCC4 and DNA-PKcs in the short-range complex. I hypothesize that disrupting residues important for stable long-term synapsis should shorten the time that the XRCC4-L4 complex remains bound to breaks, resulting in a shortened window for chromatin association. Conversely, the time-window for DNA-PKcs chromatin association should increase as failed repair events require rebinding and synapsis before beginning the pathway again.

To support this model, I suggest that future researchers should perform deep sequencing across repaired breaks to investigate repair outcomes in a disrupted shortrange complex. Our current model of NHEJ predicts that the high fidelity of joining routine breaks can be attributed to highly regulated access to DNA ends prior to ligation. It would be very interesting to know if a disruption in short-range complex stability (which has a modest impact on joining of most breaks, chapter 3) also disrupts fidelity of repair outcomes. Furthermore, combining L4 K273A with the RK>DD mutant (disrupting the L4-Ku70 interaction in the SRC) ablates nearly all joining. In my 2022 paper, we suggested that prolonged synapsis in the short-range complex could contribute to excessive end processing and the increased rate of deletions. Perhaps the repair outcomes in these combination mutants are both less efficient at repairing breaks but higher fidelity than the forms of alternative end joining.

I also suggest that a future graduate student utilize the Ku70/DNA-PK/XLF/XRCC4 halo + L4 to test mutations identified by other groups. The SMI approach in these cells allows

conclusions regarding the timing and regulation of NHEJ that are currently unaddressed. The recent L4 DBD-DNA interacting mutants reported in Stinson et al. 2024 (48) clearly disrupt short-range complex formation *in vitro*, but there are important differences between biochemical systems and live-cell imaging. For example, depletion of XRCC4-L4 complex (LX4) from frog extracts has minor impacts in the frog extract system has no discernable impact on LR complex formation (46); however, there are clear major impacts on DNA-PKcs and Ku70's association with chromatin in live cells, hypothesized to be dissolution of the LRC if LX4 is unable to associate with breaks (73). To this end, I propose utilizing the DNA-PKcs-, Ku70-, and XRCC4-halo L4-/- cell strains to test the L4 DBD-DNA interacting mutants published in *Stinson et al. 2024* (48). I predict that targeting the proposed L4-DNA interaction may trap NHEJ complexes in the LR complex in a manner that protects from NHEJ-abortion, resulting in prolonged association of all three tagged NHEJ factors.

Beyond imaging canonical NHEJ factors in the context of L4 mutants, it also would be interesting to image alternative end joining factors both 1) in the presence of catalytically inactive L4 to assess timing of alt-EJ after failed NHEJ and 2) in the presence of fully functional L4 to determine whether traditional "backup" factors respond to DSB induction. There is an open question from chapters 2 and 3 whether XRCC1-L3 can respond to DSBs and act directly within the short-range complex *in vivo*. I briefly attempted a version of this experiment by transiently transfecting $L4^{-/-}$ U2OS cells with expression vectors containing L4 K273A and L3-Halo (data not shown). Unfortunately, there were inconsistencies with L3-Halo expression levels, and I was unable to perform any experiments to test response to calicheamicin. A more careful experimental setup—likely generating a genomically edited halo-L3 $L4^{-/-}$ cell strain may be required to fully address this question.

Beyond imaging approaches, I think there is significant work to be done understanding both what processes promotes NHEJ's high fidelity outcomes and what factors determine pathway choice between NHEJ and mutagenic alternative end joining. Experiments with a substrate that assesses joining via 7 bp of microhomology (Chapter 3) suggest that a certain subset of DSBs are joined via a highly mutagenic Polθ-dependent mechanism—even in the presence of NHEJ. Using high-depth Illumina sequencing, both my work and work from other labs

(21, 20, 94, 10) indicates that NHEJ+ cells *generally* repair breaks with extremely high fidelity—or at least consistently. Further experiments looking at mutational frequency from resolved chromosomal and episomal breaks in the Polθ +/- cells, particularly the $L4^{K273A*/-}$ cells to examine the quality of repair outcome following break induction. Preliminary sequencing results are inconclusive, VDJ coding substrates were a poor choice.

While several papers have addressed the fidelity of NHEJ in small scale sequencing studies, I believe decreasing sequencing costs will enable larger whole-genome sequencing projects to investigate how mutations in NHEJ impact repair fidelity genome-wide. To do this, I propose sequencing whole-genomes of cells treated +/- DSB inducing agents in NHEJ mutant cell lines. This could become a valuable complementary approach to smaller targeted experiments like the presented targeted sequencing and would deepen our understanding of how mutations in NHEJ factors impact genomic stability—particularly those which have been identified as functionally significant in lab or prevalent in the clinic. Additionally, newer sequencing techniques—particularly Oxford Nanopore Technologies can inform not only per-base data but also provide more accurate representations of genomic translocations and copy number variations for repetitive regions.

FINAL THOUGHTS

First and foremost, thank you for reading through this dissertation, I hope you appreciate both the amount of work that's been done across the field over the past five years and my contributions to understanding L4's role in the evolving NHEJ model. Graduate school during the COVID pandemic was not easy, between lockdowns limiting access to the lab (a significant hurdle when first isolating and screening clones from my initial L4 CRISPR experiment) to supply shortages delaying experiments. Over the course of my PhD, I was most surprised by the literature revealing how prevalent DNA-PKcs across eukaryotes. When I first met Kathy in early September 2019 (before nearly all of the most recent studies into NHEJ's synaptic mechanism), I recall talking about how the mysteries of DNA-PKcs expression levels in humans compared to its absence in many other well studied organisms. This left major doubts about how widely we could apply our studies beyond vertebrates. Learning the true breadth of cross species DNA-PKcs conservation—from plants to fungi to insects and beyond—points to some deeper throughline of highly

regulated end joining repair and deepens my appreciation for NHEJ's role across eukaryotes. To that end, I hope that research into DNA Ligase IV's role in promoting end-joining continues in groups at MSU and beyond.

REFERENCES

1. Graham,T.G.W., Walter,J.C. and Loparo,J.J. (2016) Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular Cell*, **61**, 850–858.
2. Stinson,B.M., Moreno,A.T., Walter,J.C. and Loparo,J.J. (2020) A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell*, **77**, 1080-1091.e8.
3. Chen,S., Lee,L., Naila,T., Fishbain,S., Wang,A., Tomkinson,A.E., Lees-Miller,S.P. and He,Y. (2021) Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature*, 10.1038/s41586-021-03458-7.
4. Chaplin,A.K., Hardwick,S.W., Stavridi,A.K., Buehl,C.J., Goff,N.J., Ropars,V., Liang,S., Oliveira,T.M.D., Chirgadze,D.Y., Meek,K., *et al.* (2021) Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular Cell*, **0**.
5. Amin,H., Zahid,S., Hall,C. and Chaplin,A.K. (2024) Cold snapshots of DNA repair: Cryo-EM structures of DNA-PKcs and NHEJ machinery. *Progress in Biophysics and Molecular Biology*, **186**, 1–13.
6. Vogt,A., He,Y. and Lees-Miller,S.P. (2023) How to fix DNA breaks: new insights into the mechanism of non-homologous end joining. *Biochemical Society Transactions*, 10.1042/BST20220741.
7. Loparo,J.J. (2023) Holding it together: DNA end synapsis during non-homologous end joining. *DNA Repair*, **130**, 103553.
8. Chiruvella,K.K., Liang,Z., Birkeland,S.R., Basrur,V. and Wilson,T.E. (2013) Saccharomyces cerevisiae DNA Ligase IV Supports Imprecise End Joining Independently of Its Catalytic Activity. *PLOS Genetics*, **9**, e1003599.
9. Stinson,B.M., Carney,S.M., Walter,J.C. and Loparo,J.J. (2024) Structural role for DNA Ligase IV in promoting the fidelity of non-homologous end joining. *Nat Commun*, **15**, 1250.
10. Wyatt,D.W., Feng,W., Conlin,M.P., Yousefzadeh,M.J., Roberts,S.A., Mieczkowski,P., Wood,R.D., Gupta,G.P. and Ramsden,D.A. (2016) Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell*, **63**, 662–673.
11. Merker,L., Feller,L., Dorn,A. and Puchta,H. Deficiency of both classical and alternative end-joining pathways leads to a synergistic defect in double-strand break repair but not to an increase in homology-dependent gene targeting in Arabidopsis. *The Plant Journal*, **n/a**.
12. Zhou,J., Gelot,C., Pantelidou,C., Li,A., Yücel,H., Davis,R.E., Färkkilä,A., Kochupurakkal,B., Syed,A., Shapiro,G.I., *et al.* (2021) A first-in-class polymerase theta inhibitor selectively targets homologous-recombination-deficient tumors. *Nat Cancer*, **2**, 598–610.
13. Buehl,C.J., Goff,N.J., Hardwick,S.W., Gellert,M., Blundell,T.L., Yang,W., Chaplin,A.K. and Meek,K. (2023) Two distinct long-range synaptic complexes promote different aspects of end processing prior to repair of DNA breaks by non-homologous end joining. *Molecular Cell*, 10.1016/j.molcel.2023.01.012.

14. Inagawa,T., Wennink,T., Lebbink,J., Keijzers,G., Florea,B., Verkaik,N. and van Gent,D. (2020) C-Terminal Extensions of Ku70 and Ku80 Differentially Influence DNA End Binding Properties. *INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES*, **21**.

15. Singleton,B.K., Torres-Arzayus,M.I., Rottinghaus,S.T., Taccioli,G.E. and Jeggo,P.A. (1999) The C Terminus of Ku80 Activates the DNA-Dependent Protein Kinase Catalytic Subunit. *Mol. Cell. Biol.*, **19**, 3267–3277.

16. Weterings,E., Verkaik,N.S., Keijzers,G., Florea,B.I., Wang,S.-Y., Ortega,L.G., Uematsu,N., Chen,D.J. and van Gent,D.C. (2009) The Ku80 Carboxy Terminus Stimulates Joining and Artemis-Mediated Processing of DNA Ends. *MCB*, **29**, 1134–1142.

17. Chen,S., Lees-Miller,J.P., He,Y. and Lees-Miller,S.P. (2021) Structural insights into the role of DNA-PK as a master regulator in NHEJ. *GENOME INSTAB. DIS.*, **2**, 195–210.

18. Mikhova,M., Goff,N.J., Janovič,T., Heyza,J.R., Meek,K. and Schmidt,J.C. (2024) Single-molecule imaging reveals the kinetics of non-homologous end-joining in living cells. 10.1101/2023.06.22.546088.

19. Lees-Miller,J.P., Cobban,A., Katsonis,P., Bacolla,A., Tsutakawa,S.E., Hammel,M., Meek,K., Anderson,D.W., Lichtarge,O., Tainer,J.A., *et al.* (2021) Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. *Progress in Biophysics and Molecular Biology*, **163**, 87–108.

20. Reid,D.A., Keegan,S., Leo-Macias,A., Watanabe,G., Strande,N.T., Chang,H.H., Oksuz,B.A., Fenyo,D., Lieber,M.R., Ramsden,D.A., *et al.* (2015) Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proceedings of the National Academy of Sciences*, **112**, E2575–E2584.

21. Reid,D.A., Conlin,M.P., Yin,Y., Chang,H.H., Watanabe,G., Lieber,M.R., Ramsden,D.A. and Rothenberg,E. (2017) Bridging of double-stranded breaks by the nonhomologous end-joining ligation complex is modulated by DNA end chemistry. *Nucleic Acids Research*, **45**, 1872–1878.

22. Conlin,M.P., Reid,D.A., Small,G.W., Chang,H.H., Watanabe,G., Lieber,M.R., Ramsden,D.A. and Rothenberg,E. (2017) DNA Ligase IV Guides End-Processing Choice during Nonhomologous End Joining. *Cell Reports*, **20**, 2810–2819.

23. Zhao,B., Watanabe,G., Morten,M.J., Reid,D.A., Rothenberg,E. and Lieber,M.R. (2019) The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis. *Nat Commun*, **10**, 3588.

24. Luedeman,M.E., Stroik,S., Feng,W., Luthman,A.J., Gupta,G.P. and Ramsden,D.A. (2022) Poly(ADP) ribose polymerase promotes DNA polymerase theta-mediated end joining by activation of end resection. *Nat Commun*, **13**, 4547.

25. Goff,N.J., Brenière,M., Buehl,C.J., de Melo,A.J., Huskova,H., Ochi,T., Blundell,T.L., Mao,W., Yu,K., Modesti,M., *et al.* (2022) Catalytically inactive DNA ligase IV promotes DNA repair in living cells. *Nucleic Acids Research*, 10.1093/nar/gkac913.

APPENDIX: A list of authorship contributions to other papers.

This is a list of authorship contributions to academic papers as described in Chapter 4. All information updated as of 11-25-2024, presented in reverse chronological order of acceptance for publication.

1. Medina-Suarez, D., Han, L., O'Reily, S., Liu, J., Wei, C., Brenière, M., **Goff, N.J.**, Chen, C., Modesti, M., Meek, K., Harrington, B., Yu, K. (2024). DNA ligases 3 and 4 cooperate in vivo to facilitate DNA repair and organism viability. Nucleic Acids Research, In Press.
2. Mikhova, M.**, Goff, N.J.,** Janovič, T., Heyza, J.R., Meek, K., Schmidt, J.C. (2024) Single-molecule imaging reveals the kinetics of non-homologous end-joining in living cells. *Nature Communications,* in press.
3. Pascarella, G., Conner, K.N., **Goff, N.J.**, Carninci,P., Olive,A.J. & Meek,K. (2024) Compared to other NHEJ factors, DNA-PK protein and RNA levels are markedly increased in all higher primates, but not in prosimians or other mammals. *DNA Repair, 142*, 103737.
4. **Goff, N.J.**, Mikhova, M., Schmidt, J.C., & Meek, K. (2024). DNA-PK: A synopsis beyond Synapsis. *DNA Repair, 141*, 103716.
5. Buehl, C. J., **Goff, N.J.,** Hardwick, S.W., Gellert, M., Blundell, T.L., Yang, W., Chaplin, A.K., & Meek, K. (2023). Two distinct long-range synaptic complexes promote different aspects of end-processing prior to repair of DNA breaks by non-homologous end joining. *Molecular Cell 83(5),* 698-714*.*
6. **Goff, N. J.**, Brenière, M., Buehl, C. J., de Melo, A. J., Huskova, H., Ochi, T., Blundell, T, Weifeng, M., Yu, K., Modesti, M., & Meek, K. (2022) Catalytically inactive DNA ligase IV promotes DNA repair in living cells. *Nucleic Acids Research, 50(19)*, 11058–11071.
7. Liu, L., Chen, X., Li, J., Wang, H., Buehl, C. J., **Goff, N. J.**, Meek, K., Yang, W., & Gellert, M. (2022). Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Molecular Cell, 82(1)*, 177-189.
8. Chaplin, A. K., Hardwick, S. W., Stavridi, A. K., Buehl, C. J., **Goff, N. J**., Ropars, V., Liang, S., De Oliveira, T. M., Chirgadze, D. Y., Meek, K., Charbonnier, J.B., & Blundell, T. L. (2021). Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular Cell, 81(16)*, 3400-3409.