Chapter 2

The Role of Multimerization During Non-Homologous End Joining

Michelle Rubin, Jonathan Newsome and Albert Ribes-Zamora

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/53982

1. Introduction

In order to sustain life, cells must protect their genetic information from the constant threat posed by mutagenic agents such as ultraviolet light, irradiation or reactive oxygen species, as well as from mistakes introduced during the replication of their genomes [1]. To deal with this problem, natural selection has favored a system that repairs the damage caused by these DNA lesions while allowing the highly infrequent but steady production of mutations that constitute the source for adaptive changes during evolution. To repair damaged DNA, cells have developed a myriad of highly specialized pathways that recognize and repair specific types of injuries produced by specific types of mutagenic events [2, 3]. For instance, while base excision repair detects and repairs chemically damaged nucleotides typically produced by oxygen radicals or alkylating agents, nucleotide excision repair is responsible for the removal of thymine dimers caused by ultraviolet light exposure. Additionally, the mismatch repair pathway specializes in fixing errors introduced during DNA replication. More globally, these pathways are part of the DNA damage response (DDR), a signal transduction cascade coordinated by the ATM/ATR kinases in mammalian cells that halts cell cycle progression while DNA is being repaired, and it can trigger apoptosis when the damage is deemed non-repairable [4]. The importance of these pathways is underlined by their high conservation, both in prokaryotes and eukaryotes, and it is emphasized by the role that they play in disease when impaired. Malfunctioning DNA repair pathways are associated with several disorders such as Xeroderma pigmentosum or Nijmegen syndrome as well as with increased cancer risk, as they boost the formation of spontaneous mutations that can lead to tumorigenesis [5].
For any type of cell, one of the more toxic DNA injuries is the double-strand break (DSB). This form of lesion can arise as a consequence of mechanical stress, exposure to irradiation or as a result of a replication fork encountering a single-strand nick [6]. DSBs can induce translocations, aneuploidy and global genome instability that can ultimately render cells either unviable or tumorigenic [5]. To repair a DSB, mammalian cells can take advantage of the presence of homologous chromosomal copies and use homologous recombination (HR) to faithfully amend the break [7]. In the cell cycle phases where identical chromosomal copies are not available, the preferred repair pathway is non-homologous end joining (NHEJ) which seams the two ends of the break with mostly minimal alteration of the DNA sequence [1, 8-10]. NHEJ requires the completion of three major steps: (1) protection and synapsis of both DNA ends, (2) processing of the DNA termini and (3) the final ligation of the ends [9, 10]. The DNA-PK complex, formed by the Ku heterodimer and DNA-PKcs, is responsible for the initial protection and synapses of the ends and recruits other NHEJ factors to the DSB. These factors include the nuclease Artemis and polymerases Polμ and Polλ that will remove and add nucleotides to replace possible damaged bases generated during the breaking process. The final ligation step is performed by DNA ligase IV (LigIV), whose recruitment to DSBs depends on its close association with XRCC4, a process aided by XLF. The regulation of these steps is still not clear but it is known that DNA-PKcs phosphorylates several NHEJ factors and can induce its own removal from DSBs, and that ubiquitination also plays a role in disassembling complexes once the damage has been repaired [11-14]. Furthermore, ATM can phosphorylate NHEJ factors, although their role during NHEJ remains to be elucidated [15].

During evolution, the end protection properties of some NHEJ factors were recycled into protecting the natural ends of chromosomes. How telomeres manage to harbor NHEJ factors to protect their ends while preventing them from triggering end-to-end fusions is still an open question [16]. Later on, NHEJ was again recycled into joining physiologically programmed DSBs that occur during V(D)J recombination and class switch recombination (CSR) in B- and T- lymphocytes. These pathways ensure antigen-binding diversity in antibodies as well as the presence of different antibody isotypes capable of binding different downstream effectors. Consistent with a role of NHEJ in these pathways, mutations in several NHEJ factors are associated with diseases in which the immune system is compromised. For instance, mutations in XLF, DNA-PKcs and Artemis are present in patients suffering from severe combined immunodeficiency or SCID [17].

2. Multimerization of NHEJ factors

For the last two decades, research in NHEJ has mostly focused on the identification of genes involved in this pathway and the dissection of their enzymatic activities [9]. The structures of most NHEJ factors have been solved and this is starting to unravel how NHEJ is regulated throughout the cell cycle [18]. Despite these advances, we are still lacking a clear model of how all these factors assemble at DSBs and whether different
complexes form depending on the type of damage, the location of the break or the cell cycle phase when the injury occurs [10]. One of the emerging themes in the field is the assembly of NHEJ factors as multimers at DSB. This multimerization has been reported for several NHEJ proteins and occurs with varying degrees of complexity. The simplest form of multimerization is seen with DNA-PKcs, where monomers bound to opposing ends of a DSB can dimerize and effectively synapse the DNA break [19]. Similarly, two Ku heterodimers are capable of forming a heterotetramer that can tether the ends of a DSB. More intricate multimerization can be observed in the MRN heterotrimer, a complex composed of Mre11, Rad50 and Nbs1 that can form either heterohexamers, where two molecules of each subunit combine into a larger structure, or multimers of four MRN complexes at DSBs [20]. Most recently, a complex multimerization of NHEJ proteins has been observed in the form of long filaments created by the polymerization of multiple alternating copies of XLF and XRCC4 homodimers [21-23]. Combined, this data suggests that formation of multimers is a constant theme in the assembly of NHEJ proteins at DSBs. Below we review current literature on this topic, identifying questions that remain to be answered while laying out possible new research directions. While NHEJ can be divided into classical NHEJ (c-NHEJ) or alternative NHEJ (a-NHEJ) depending on the factors required for completion, here we focus on c-NHEJ and will refer to it as NHEJ.

2.1. The Ku heterodimer

Once a break forms, one of the first responders is the Ku heterodimer, an abundant protein (~400,000 molecules per cell) formed by the Ku70 and Ku80 subunits [24]. Ku is an obligate dimer as in the absence of one subunit the other subunit disappears from cell extracts, presumably due to lack of proper folding [25, 26]. Despite showing low sequence similarities, both Ku70 and Ku80 contain nearly identical domains and have very similar secondary and tertiary structures. Both subunits share a vonWillebrand domain (vWA, also referred to as the a/b domain) in their N-termini followed by a central dimerization domain that can also bind DNA [27]. The only divergence between both subunits is the presence of a C-terminus SAP domain (SAF-A/B, Acinus and PIAS) exclusively in Ku70, which is replaced in KU80 by a domain that is involved in recruiting DNA-PKcs to DSBs [24, 27, 28].

The structure of Ku shows a quasi-symmetrical configuration with both central domains dimerizing and forming a DNA binding ring flanked by the two vWA domains on opposite sides of the ring [27]. This creates a toroidal structure with a basket-like shape that can thread onto DNA (Figure 1A). Ku binds duplex DNA ends with great affinity (Kd~10^{-9} M) and in a sequence independent manner, hence its role as the first DSB recognition factor [29-31]. Ku needs at least 14bp to bind DNA and since the DNA binding ring is preformed, Ku requires a free end to associate with DNA [27]. Consistent with this, the affinity of Ku to circular DNA is orders of magnitude lower compared with linear DNA [32]. Similarly, Ku’s affinity to single-stranded DNA (ssDNA) is lower than to double-stranded DNA (dsDNA), which presumably favors HR over NHEJ once resection of ends becomes too extensive to hold Ku [7].
Resolution of the X-ray structure of DNA bound Ku indicated that binding occurs with a preferred orientation that places the Ku70 vWA domain closest to the end and the Ku80 vWA furthest from the end [27]. In the budding yeast *S. cerevisiae*, mutations in α-helix5 (α5), the most outer structure in the vWA domain had opposite effects in each subunit [33]. Whereas Yku80-α5 mutations disrupt Ku’s telomeric silencing function without perturbing DNA repair abilities, mutations in Yku70-α5 impair NHEJ while preserving Ku’s telomeric functions [33]. This suggests that Ku is spatially organized in two faces with distinct roles in NHEJ and telomeric functions, a hypothesis termed the two-face model. In essence, this model states that the inward face, composed mostly of the Ku80 vWA and Ku70 C-terminal domain (CTD), is oriented towards internal tracks of DNA and has telomeric roles. In contrast, the outward face of Ku is the closest to the DNA end and its main role is to engage the nearby DNA end in NHEJ. Consistent with this, mutations in both the Ku70 vWA domain and the Ku80 CTD, the two major components of the outer face, impair NHEJ [33, 34]. The most recent discovery that mutations in human Ku70-α5 also diminish NHEJ suggests that the two-face model may also be conserved in mammalian cells [35].

**Figure 1.** Molecular modeling of Ku and DNA-PKcs dimerization. A Possible Ku tetramer formation through outward face interactions. Each Ku dimer is represented in red and blue, whereas DNA is depicted as black line (Adapted from 1JEY). B Possible Head-to-Head mediated dimerization of two DNA-PKcs molecules (Adapted from 3KGV).

Ku plays multiple roles during NHEJ. Initially, Ku not only binds and detects DSB but also protects the ends from nucleolytic degradation and tilts the choice of DSB repair towards NHEJ and away from HR [7, 36]. Perhaps the most extensive role for Ku is to recruit NHEJ factors to DSB sites. The Ku80 CTD recruits DNA-PKcs to DSBs and binds a long list of NHEJ factors including XRCC4, LigIV, XLF, Polμ, and Polλ [9, 10]. Moreover, Ku’s requirement for the recruitment of DNA-PKcs, XLF and XRCC4 has been demonstrated in vivo [37, 38]. In addition, Ku possesses some catalytic activities and can function as a deubiquitylating enzyme or as a 5’-dRP/AP lyase which suggests that Ku may aid in processing DNA ends before the final ligation [39].

The ability of the Ku heterodimer to self-associate and create higher order multimers was already apparent in early studies. Seminal work from Cary *et al* using gel filtration
chromatography showed that while recombinant Ku exists as a single heterodimer in solution, the addition of 24bp DNA fragments promotes the appearance of Ku multimers whose sizes correspond to that of a heterotetramer [40]. More importantly, using a combination of electron microscopy (EM) and atomic force microscopy (AFM), they visualized Ku-dependent end-to-end bridging events involving either ends of two DNA fragments or loops within a single DNA fragment. Ku was found forming higher order multimers in the junction of those events, which suggests that Ku multimerization is responsible for its end bridging properties. To note, this was not the result of non-specific aggregation of Ku as multimers could only be detected with DNA-bound Ku and not with free Ku molecules. The ability of Ku to synapse two ends was later confirmed by Ramsden et al using a mix of biotinylated DNA fragments with $^{32}$P-radiolabeled dsDNA [41]. When streptavidin beads were used to pull down biotinylated DNA, researchers could recover radiolabeled DNA only when recombinant Ku was also present in the mix. This co-precipitation could not be explained by a single Ku molecule binding and stabilizing the junction of two DNA fragments with complementary ends as similar results were obtained when DNA fragments with non-compatible ends were used. This result suggests that synapses were achieved through the interaction of at least two Ku molecules each bound to a differently labeled DNA. More recently, DNA-bound Ku heterotetramers have been demonstrated as supershifts in electromobility shift assays (EMSA) and EM studies have visualized formation of end bridges using recombinant yeast Ku protein, indicating that multimerization-dependent synapses may be evolutionarily conserved [42].

Recent work with the Ku80 CTD suggests that Ku heterotetramerization may occur through the interaction of two outward faces [42]. This long and flexible domain interacts with the Ku core domain and upon binding to DNA it undergoes a conformational change that relocates it to the outward face [43]. Researchers have now shown that the Ku80 CTD can dimerize and thus, can putatively mediate Ku-Ku interactions across a DSB enabling the tethering of two DNA ends [42]. In fact, Ku proteins bearing Ku80 CTD truncations have reduced ability to form heterotetramers as shown by reduced supershift EMSA signals compared with wild type [42]. Ku80 CTD truncations impair NHEJ, but this result cannot be fully attributed to reduced heterotetramer formation as the Ku80 CTD is also involved in recruiting DNA-PKcs to DSBs. Intriguingly, a mutation in the outward face located Ku70-α5 also impairs NHEJ, although the effect of this mutation on Ku heterotetramerization remains to be investigated [35].

While the presence of Ku multimers of higher order than heterotetramers has been detected in EM and AFM using recombinant Ku proteins, its functional significance remains to be elucidated and evidence for its existence in living cells is lacking [40, 44-46]. A single heterotetramer is sufficient to create a synaptic complex across DSB and it is difficult to envision how higher order Ku complexes may aid in this process. Since Ku can slide towards internal tracks of DNA, one possibility is that multiple Ku molecules threaded into a single DNA end could form filaments held by interactions between inward and outward faces. However, fluorescence anisotropy studies do not support this model [47].
2.2. The DNA-PKcs

With over 400 kDa, the DNA-PKcs is one of the largest kinases in mammalian cells. Along with ATM and ATR, it belongs to the phosphatidylinositol-3-OH kinase (PI3K)-related kinase (PIKK) family that preferentially phosphorylates serines and threonines followed by a glutamine [48]. Although DNA-PKcs can bind directly to DNA, during NHEJ it is recruited to DSBs by the Ku80 CTD flexible domain, which increases the affinity of DNA-PKcs to DNA by 100 fold [49-51]. Therefore, assembly of the DNA-PK complex only occurs at DSBs where it induces Ku displacement one helix turn away from the end and positions DNA-PKcs at the very tip of the break [44]. The DNA-PK complex performs two major roles during NHEJ: it forms a synaptic complex across DSBs and serves as a scaffold for the recruitment of all other NHEJ factors [52]. DNA-PKcs is responsible for the recruitment of Artemis to DSBs, which provides the NHEJ machinery with a variety of end processing activities including 5’ endonuclease, 3’ endonuclease and hairpin opening [53]. In addition, DNA-PKcs directly binds XRCC4 and stimulates the ligase activity of XRCC4/Ligase IV complex [52, 54, 55]. Both dimerization and XRCC4 interaction induces DNA-PKcs kinase activity which is known to phosphorylate several NHEJ factors such as Ku, Artemis, XRCC4, LigIV and XLF, although the role of these phosphorylation events in NHEJ remains to be elucidated [56, 57]. More importantly, DNA-PKcs possesses over 15 autophosphorylation sites that become phosphorylated after formation of the synaptic complex and that are involved in releasing DNA-PKcs from Ku and DSB [11]. Consistent with this, non-autophosphorylatable mutations or kinase-dead DNA-PKcs mutants still localize to DSBs but are retained longer at sites of DNA damage [43].

Given its large size, the complete structure of DNA-PKcs has been elusive at the atomic level. Single particle cryo-EM, small-angle X-ray scattering (SAXS) experiments and more recently, the crystal structure at 6.6 angstroms resolution, have shown that multiple N-terminus HEAT repeats, encompassing ~66 helices, form a ring-like structure with a gap at one end (Figure 1B) [58-60]. This structure is usually referred to as the ‘palm’ region and it also encloses a globular DNA binding domain, although a clear picture of how DNA-PKcs interacts with DNA is missing. The top of the palm houses the so-called ‘crown’ or ‘head’ that includes the globular C-terminus kinase domain, along with FAT and FATC domains. Also missing is the exact location of the Ku80 CTD interaction and the details of how the whole Ku heterodimer is accommodated by the DNA-PKcs structure to create the DNA-PK complex [61]. Several pieces of evidence indicate that DNA-PKcs undergoes conformational changes as a result of autophosphorylation [43, 62, 63]. SAXS analysis has detected a phosphorylation-driven conformational change that repositions the head with respect to the palm [43], whereas the crystal structure suggests that auto-transphosphorylation widens the gap at the end of the palm and facilitates disassembly of DNA-PKcs from Ku and DSBs [58].

During NHEJ, DNA-PKcs multimerization is limited to the dimerization of the two DNA-PKcs located at opposite ends of a DSB to create a synaptic complex. Early studies demonstrated the ability of DNA-PK to mediate co-immunoprecipitation of biotinylated DNA fragments with radiolabeled probes [64]. In agreement, initial EM experiments visualized the ability of DNA-PKcs to circularize DNA fragments and detected synaptic complexes
whose size was consistent with that of a DNA-PKcs dimer. While DNA-PKcs was sufficient to create a synaptic complex, these were significantly more abundant when Ku was present [64]. Importantly, end-to-end bridges still occurred in the presence of kinase inhibitors, indicating that autophosphorylation events were not required for synaptic complex formation [64]. Further single particle EM, cryo-EM and SAXS studies have visualized dimers of DNA-PKcs molecules that form in a concentration dependent manner and in a process that is highly enhanced in the presence of Ku [43, 60, 63]. These techniques detected two types of DNA-PKcs dimers with different orientations depending on the DNA molecules used. In the presence of 40bp Y-shaped DNA fragments, DNA-PKcs dimers formed in a palm-to-palm fashion whereas, in the presence of a 40bp hairpin DNA, DNA-PKcs dimers had the opposite orientation and formed through head-to-head interactions [43, 60, 63]. These two kinds of DNA-PKcs dimers were observed in the presence or absence of Ku. The reason for the existence of two DNA-PKcs dimer subspecies is not clear but authors speculate that the Y-shape DNA induced orientation may be caused by the binding of two DNA-PKcs molecules to the same DNA fragment, suggesting that dimers generated by head-to-head interactions may be the only ones capable of bridging two DNA fragments [43]. Corroborating this hypothesis, DNA-PKcs dimers with a head-to-head orientation were the only type of dimers observed in the absence of DNA [43]. Overall, the current model proposes that Ku recruits DNA-PKcs to sites of DNA damage where it dimerizes through head-to-head interactions, creating synaptic complexes across DSBs [10]. DNA-PKcs dimerization at breaks stimulates auto-transphosphorylation, which in turn induces conformational changes that disassemble the DNA-PK complex and promotes its timely release from DNA ends [11].

2.3. The MRN complex

The association of the conserved Mre11 and Rad50 subunits along with Nbs1 (a protein whose functional homolog in S. cerevisiae is Xrs2), makes up the mammalian MRN complex, also known as the MRX complex in yeast [65-68]. This complex plays vital roles in multiple DNA repair pathways, including HR and NHEJ, and is responsible for the co-activation of the DDR in the presence of DNA injury [68-70]. Analysis of the domain composition and enzymatic activities of each subunit suggest possible mechanistic roles for the MRN complex during DNA repair. Rad50 is a member of the SMC protein family whose members play roles in chromosome condensation and cohesion [71, 72]. A key feature of these proteins is the presence of long coiled-coil domains that can fold on themselves via an antiparallel manner, bringing the N- and C-terminus in close proximity [73]. In Rad50, folding of the coiled-coil domain permits the reconstitution of a bipartite ATP-binding cassette (ABC)-type ATPase globular domain made of N-terminal Walker A and C-terminal Walker B nucleotide binding motifs [74, 75]. In other complexes, binding and hydrolysis of ATP by similar ABC-ATPase domains, mediate large conformational changes that can be transmitted to other members of the complex [71, 76]. Crystallography and SAXS data support a similar role for the Rad50 ABC-ATPase domain in the MRN complex [74, 77, 78].

As is the case for Rad50, Mre11 can bind DNA and possesses a specific region capable of capping DNA ends [20, 79]. Mre11 contains a phosphoesterase domain in its N-terminus
that endows the MRN complex with ssDNA endonuclease and 3’ to 5’ dsDNA exonuclease enzymatic activities. Mre11 is the only subunit that interacts with all components of the complex as it also binds Nbs1 whose main function is to recruit DNA repair factors to DSB. For instance, the Nbs1 N-terminus is responsible for bringing ATM and ATR to DNA damage locations and hence, acts as a DSB sensor during the initiation of the DDR [80, 81]. Similarly, the Nbs1 N-terminus contains BRCT and FHA domains that bind and recruit CtIP (Sae2 in *S. cerevisiae*), an important nuclease during DNA repair, to DSBs [28, 82] as well as MDC1 [83, 84] and ATR[85], and the WRN helicase [86].

The MRN complex is at the center of the decision process that governs whether a DSB is repaired by NHEJ or HR [87, 88]. The current model indicates that, in the presence of DNA injury, recruitment of CtIP to DSB provides the MRN complex with the 5’ to 3’ exonuclease activity necessary to chew away part of the DNA ends and create an initial ~50-100nt ssDNA 3’ overhang [18, 82, 89-92]. This overhang is a poor binding substrate for Ku but an ideal substrate for the HR initiation factor RPA and thus, it favors DSB repair by HR over NHEJ [32, 93]. Since CtIP activity and recruitment to DSBs is dependent upon CDK phosphorylation of both CtIP and Nbs1 during S and G2 phases [93-95], the lack of these post-translational modifications during G1 prevents overhang formation at DSB ends in this phase and tilts the choice of DNA repair pathway towards NHEJ. Therefore, and according to this model, NHEJ remains active throughout the cell cycle but is overpowered by HR during S and G2 due to CDK-dependent 3’ overhang formation by the MRN/CtIP complex. This system ensures that HR is only active when an identical copy of the damaged DNA is available.

Despite favoring HR, the MRN complex also plays essential roles during NHEJ when HR is inhibited. Yeast defective for any MRX subunit are inviable in the presence of a single HO–induced break that can only be repaired by NHEJ [96]. Similarly, mammalian cells depleted of Mre11 display reduced end-joining activities [97], impaired NHEJ [69, 98] and, in the case of B-lymphocytes, markedly reduced CSR [99]. Intriguingly, yeast nuclease-dead Mre11 mutants can carry end-joining activity near wild type levels [100] and B lymphocytes only show mild defects in CSR in the presence of a Mre11 mutation lacking nuclease activity [99]. These results indicate that the MRN complex mostly plays a signaling and structural role during NHEJ, and that its end processing capabilities are dispensable or can be compensated by other nucleases. Scanning force microscopy (SFM) and AFM have demonstrated the ability of the MRN complex to create long range bridges across DNA molecules [101, 102] implicating DNA end tethering as the most likely structural role for MRN during NHEJ.

Formation of higher order multimers is essential to our understanding of the mechanistic roles of MRN during NHEJ. The MRX complex assembles as a heterohexamer where two Mre11 molecules bind simultaneously to two Rad50 and two Nbs1 subunits (Figure 2A) [70, 103]. A combination of SAXS with X-ray crystallography has shown that, through interactions of the N-terminus globular domain, Mre11 form very stable dimers capable of forming bridges between two DNA molecules [20, 104]. Rad50 can also form dimers through two different dimerization regions located at opposite poles of the molecule [74]. At the end of the coiled-coil region, Rad50 contains a C_{XXC} motif that can dimerize through the formation of two zinc-hook (Zn-hook) domains that lock in a single Zn(2+) ion [103]. At the opposite pole, two globular
ABC-ATPase domains bound to ATP can dimerize in head-to-tail fashion between N- and C-terminal domains, trapping two ATP molecules in the process [74]. Moreover, Mre11 dimers bind ABC-ATPase dimers forming the so-called M2R2 head region. In this disposition, a heterohexamer has a circular shape formed by two coiled-coil regions as semicircles are united at one end by the Zn-hook domain and by interactions within the M2R2 head at the other end (Figure 2C). Mre11 also contains a helix-loop-helix domain in its C-terminus that extends away from the N-terminal globular region and binds the base of the Rad50 coiled-coil region in the vicinity of the ABC-ATPase domain, further reinforcing the interaction between Mre11 and Rad50 dimers (Figure 2A). In contrast, Nbs1 does not form dimers nor does it contribute to heterohexamer-ization. Recent studies indicate that the MRN heterohexamer presents two distinct configurations [77, 78, 105, 106]. In the absence of ATP binding, the MRN complex adopts an ‘open’ conformation where Mre11 dimers localize in between the two Rad50 ABC-ATPase domains, preventing their dimerization (Figure 2A). In this ‘open’ configuration, Rad50 can only dimerize through the Zn-hook domain. Upon ATP binding, a conformational change allows displacement of the Mre11 dimers and dimerization of two Rad50 ABC-ATPase domains (Figure 2B). This ‘close’ configuration is substantially more rigid and may promote DNA binding by Rad50 [77, 78, 105, 106]. Subsequent ATP hydrolysis disrupts ABC-ATPase mediated Rad50 dimerization and stimulates Mre11 nuclease activity [107].

Figure 2. The MRN complex undergoes ATP-Driven conformational changes. A Molecular structure of the open state where Mre11 dimer is depicted in red and Rad50 in blue (Adapted from 3QG5). B Closed conformation of the MRN complex. (Adapted from 3QF7 and 3THO). C Tethering of DSB ends by different multimerization states of the MRN complex.
In the MRN complex, multimerization regulates formation of DNA bridges, with increasing order of multimers providing longer range tethering capabilities. Short range bridges can be achieved by single heterohexamers, where each Mre11 subunit is bound to a different DNA end (Figure 2C) [103]. In contrast with Ku heterotetramers and DNA-PKcs dimers, where multimerization occurs after the assembly of subunits located at each end of a DSB, short range bridges mediated by single heterohexamers must be achieved without disruption of the MRN complex as heterohexamers are predicted to be pre-assembled before binding DNA ends. Therefore, single heterohexamers may only bridge ends that are in close proximity (~100 angstroms) as to allow simultaneous binding of each end to the MRN complex without disrupting the M2R2 head [20, 103]. The role that DNA binding activity of the Rad50 dimers may play in synapse formation by heterohexamers remains to be elucidated but SFM has shown that Rad50 is also able to bind and tether DNA molecules in the absence of Mre11 [108].

Longer range bridges can be achieved through the formation of higher order multimers where two heterohexamers, each bound to a different end, combine to form a structure capable of tethering DNA molecules as far as 1200 angstroms apart [20, 103]. This configuration has been confirmed by AFM and EM and consists of two M2R2 heads separated by two long coiled-coil regions held together by two Zn-hooks (Figure 2C) [102, 103]. Consistent with this, AFM has demonstrated that upon DNA binding, heterohexamers extend their two coiled-coil regions in a parallel fashion that disrupts Zn-hook mediated dimerization within the heterohexamer and favors formation of Zn-hook interactions with other heterohexamers [102]. These assemblies have been proposed to mediate long distance tethering of homologous sequences during HR and to hold DNA ends in close proximity during NHEJ, preventing them from going astray and facilitating DSB repair by the rest of the NHEJ machinery [70, 109]. Consistent with this, in yeast, loss of the Rad50 CxxC Zn-hook motif abolishes DNA repair while replacing it with a FKBp homodimerization domain has no major effect [110]. Similarly, truncations of the coiled-coil region in Rad50 impairs DNA repair [109]. Since heterohexamers could potentially form Zn-hook interactions with more than one heterohexamer at a time, further higher order arrangements have been proposed to form multiple interactions to secure bridges across DNA molecules, a possibility whose biological significance remains to be investigated.

2.4. XLF and XRCC4

Although neither of them have any intrinsic enzymatic activity, XLF (also known as Cernunnos) and XRCC4 are involved in the final ligation step catalyzed by LigIV. XRCC4 binds tightly to LigIV and drives its localization to sites of DNA damage, whereas XLF stimulates the ability of the XRCC4-LigIV complex to ligate DSBs 20-200 fold, especially in the presence of non-cohesive ends [111-114]. XRCC4 is also capable of binding several other NHEJ factors including XLF, DNA-PKcs and the Ku heterodimer, while known interactions for XLF include XRCC4 and the Ku heterodimer. In addition, both XLF and XRCC4 can bind DNA, although their localization to sites of DNA damage is dependent on their interaction with members of the DNA-PK complex. While DNA-PKcs can recruit the XRCC4-LigIV complex
to DSBs, Ku is capable of bringing both XRCC4-LigIV and XLF to sites of DNA damage in the absence of DNA-PKcs [37, 38, 54, 115, 116].

Both XLF and XRCC4 are obligated homodimers that have very similar structures where the presence of an N-terminal globular domain, or 'head', is followed by a coiled-coil region that mediates homodimerization [112, 117, 118]. In this disposition, the two head domains on each homodimer face each other with opposite orientation. Given its heterogeneity and flexible nature, the structure of the C-terminal domain for both XRCC4 and XLF remain to be resolved. Nevertheless, XRCC4 SAXS analysis is consistent with the C-terminal domain folding backwards and interacting with the head domain [119]. Likewise, XLF's structure resolution revealed a fold back in the coiled-coil region that shortens the helix and creates a kink that likely positions the C-terminal domain in close proximity to the head domain [118]. The presence of discrete regions in both XRCC4 and XLF correlates with the spatial organization of their interactions with NHEJ factors. While their N-terminal domains mediate XRCC4-XLF interaction, the XRCC4-LigIV interaction maps to the XRCC4 coiled-coil region and the C-terminal domain of XLF interacts with Ku [22, 38, 116, 119-122]. In addition, both XRCC4 and XLF can bind DNA. EMSA analysis has identified the upper part of the XRCC4 coiled-coil domain and the XLF C-terminus as their respective DNA binding regions [22, 23, 118].

Besides forming homodimers, XRCC4 also exists as tetramers and higher order multimers in solution [119, 123]. Although coiled-coil mediation of tetramerization was initially proposed, SAXS analyses have demonstrated that tetramerization is mostly mediated by the interaction of two N-terminal domains in a way that leaves the stalks of each dimer pointing in opposite directions [119, 124, 125]. These head-to-head interactions can also drive formation of XRCC4 filaments, as detected by SAXS [119]. Interestingly, while full length XRCC4 mostly exists as tetramers and filaments in solution, truncation of the C-terminus makes homodimers the predominant XRCC4 form, suggesting that the C-terminus also contributes to tetramerization and filament formation. The presence of a LigIV BRCT region responsible for binding the XRCC4 coiled-coil region also made XRCC4 filaments unstable, indicating that the XRCC4-LigIV complex does not exist as part of a filament and suggesting that under physiological conditions, XRCC4 remains in multiple configurations [119]. Given the fact that mammalian cells contain six times more XRCC4 molecules than LigIV, XRCC4 filaments may constitute a protein reservoir that can readily be mobilized in the presence of DNA damage. The ability of XLF to form higher order assemblies in solution suggests that XLF filaments may also exist in cells [118].

In addition, XRCC4 can also form filaments through its interaction with XLF. SAXS analysis, EM, SFM and crystallography have all detected long filaments of alternating XRCC4 and XLF molecules bound through head-to-head interactions (Figure 3) [21-23, 119, 126]. In this conformation, XRCC4 and XLF stalks are both oriented towards the same direction, albeit with a 30-degree offset from each other. Furthermore, two filaments can intertwine through XRCC4-XRCC4 interactions to form a left-handed helix with a ~220 angstrom diameter where head domains reside in the interior while coiled-coil regions stick out to the exterior. Higher order multimers where several filaments constitute a thicker fiber have also been
Importantly, mutations in residues directly involved in XRCC4-XLF interaction not only disrupt filament formation but also disrupt NHEJ and render cells radiosensitive, indicating that XRCC4-XLF filaments are functionally relevant structures during DSB repair [23, 126]. Furthermore, during NHEJ, XRCC4-XLF filament formation is likely to be regulated as in vitro experiments have shown that DNA-PKcs dependent phosphorylation of XRCC4 and XLF disassembles XRCC4-XLF filaments [126].

**Figure 3.** Sequential interactions between the N-terminal domains of XRCC4 and XLF can create multimeric filaments. XRCC4 is colored yellow and XLF is colored teal (Adapted from 3RWR).

Given the ability of XRCC4 and XLF to bind DNA, it is likely that the XRCC4-XLF complex forms nucleoprotein filaments. In agreement, EMSA experiments in the presence of XRCC4 and XLF can detect supershifts consistent with formation of large nucleoprotein complexes [23, 126]. How DNA interacts with XLF-XRCC4 filaments is not clear, but EMSA supershifts are lost in the absence of the XLF C-terminus domain but not when the XRCC4 DNA binding region has been removed, suggesting that the XLF C-terminus plays a greater role than XRCC4 in nucleoprotein filament formation. Recently, HDX studies have revealed that the interface between XRCC4 and XLF in the filament may accommodate DNA, although the details of this interaction remain to be elucidated [22]. These nucleoprotein filaments are highly reminiscent of Rad51 filaments that form during HR and suggest that XRCC4-XLF filaments may also coat dsDNA ends to protect and prepare them for processing and ligation. It is also possible that they facilitate DNA repair by ‘peeling’ away nucleosomes from DSBs and making DNA ends more accessible to other NHEJ factors. In addition, due to their length and ability to bind DNA, XRCC4-XLF filaments can form bridges across DSBs, as demonstrated by their ability to mediate co-immunoprecipitation of two different DNA fragments [23, 126]. Importantly, conditions that disrupt XRCC4-XLF filament formation, such as mutations in the XRCC4-XLF interface, presence of LigIV BRCT domains, lack of the XLF C-terminal domain or DNA-PKcs phosphorylation, also prevent DNA bridging in vitro [23, 126].

While LigIV uses its BRCT domains to bind the XRCC4 coiled-coil region, its catalytic domain interacts with the XRCC4 N-terminus domain and therefore, the presence of LigIV bound to XRCC4 is not compatible with XRCC4-XLF filament formation [127]. These data suggest that XRCC4 may be present in different configurations at DSB: as an XRCC4 fila-
ment, as part of an XRCC4-XLF nucleofilament and as a separate XRCC4-LigIV complex. Since LigIV can only interact with one head of the XRCC4 homodimer, it is also possible that the other head may be free to start polymerization of a LigIV-free XRCC4-XLF filament. In this conformation, a single XRCC4-LigIV at the tip of a DSB may cap an XRCC4-XLF filament that extends inward and away from the end [22]. Other possibilities include DNA-PK acting both as the DNA end cap of an XRCC4-XLF filament and as the recruiter of an XRCC4-LigIV complex or a single filament that expands across the DSB, allowing other NHEJ factors to reach the DNA ends [22, 126]. Further experiments are required to discern among these possibilities and to investigate how and when XRCC4-XLF filaments assemble, and to investigate the additional functions that they play during NHEJ.

3. Conclusions and new directions

The evidence presented here strongly indicates that multimerization of Ku, DNA-PKcs, the MRN complex and the XRCC4/XLF complex play crucial functions during NHEJ. In contrast, while more than one molecule of other NHEJ components like Artemis, LigIV, Polμ and Polλ are likely to be present at DSBs, neither the presence of higher order multimers nor a functional role for the accumulation of their subunits at breaks has been demonstrated. In other DNA repair pathways, several examples of multimerization can also be found. For instance, during HR the RPA complex polymerizes along ssDNA ends forming filaments that protect ends from degradation and are readily substituted by Rad51 filaments to catalyze strand exchange. Other notable examples include WRN, a member of the RecQ helicases involved in DNA damage response and telomere maintenance, which contains a coiled-coil region that serves as a multimerization domain to create trimers and hexamers required for full protein function and BLM, a member of the same protein family that functionally exists as homohexameric rings [128, 129]. Recent studies have demonstrated that CtIP dimerization is also required for its recruitment to DSB and subsequent HR [130].

During NHEJ, the most prevalent role for multimerization is to ensure the formation of protein bridges across a DSB. When bound to both ends of a break, either DNA-PKcs, the Ku heterodimer or the MRX complex can form multimerization-driven synapses that hold the two ends together and facilitate DNA damage repair. In addition, XRCC4-XLF filaments can also form bridges across two DNA fragments and may contribute to end synapses. How NHEJ factors assemble at DSB and the stoichiometry of such assemblies remain to be elucidated. The high redundancy of NHEJ factors capable of bridging ends may reflect the presence of different subcomplexes that are formed depending on the type or location of the damage. Alternatively, different NHEJ proteins may be involved in synapsing DNA ends at different steps during NHEJ. For example, initial Ku-mediated DNA bridges may be disrupted and replaced by DNA-PKcs as recruitment of DNA-PKcs to DSBs is known to displace Ku internally away from the ends [131]. Further investigations on how multimerization influences NHEJ are likely to provide insights not only on the stoichiometry of NHEJ complexes at DSBs but also on the different progression steps of the DNA repair process. An important aspect to consider is the diversity that exists among NHEJ bridging proteins with respect to the distance between DNA
ends once the bridge is formed. While multimers of MRX complex can establish long-range bridges between DNA ends, DNA-PKcs and the Ku heterodimer bridges are limited to short-range synapses. It is possible that different synaptic complexes may modulate the separation of DNA ends during different NHEJ steps to allow the timely access of DNA processing factors while, at the same time, holding the two ends together. In addition, there may be differences in the strength by which each synaptic complex holds the two DNA ends. For instance, Ku mediated synapses can only be observed at high concentrations of Ku protein, which may partially explain why its detection was missed in several studies. The weakness of Ku mediated synapses may facilitate its replacement by putatively stronger synapses via DNA-PKcs. Future experiments are needed to delineate other possible transitions between synapses of different strength during each NHEJ step and to establish whether different DNA bridges can occur simultaneously at the same DSB.

It could also be insightful to dissect the multimerization state of NHEJ proteins at the end of chromosomes. Telomeres use the shelterin complex to protect the natural chromosome end from being acted upon as a DSB, which could result in deleterious chromosome end-to-end fusions and generalized genomic instability [132]. Surprisingly, telomeres also harbor several members of the NHEJ machinery such as Ku, the MRX complex and DNA-PKcs [16]. How these proteins are prevented from engaging in NHEJ at telomeres is not fully understood but it is possible that interactions with shelterin components not only recruit NHEJ proteins to telomeres but also impair their multimerization. Therefore, delineating how NHEJ proteins interact with sheltering components could provide information on how their DNA repair properties are blocked at telomeres.

Another untapped area of investigation is the role that NHEJ multimerization may play during V(D)J recombination and CSR. These programmed physiological cuts generate different DNA end substrates that support the formation of different subcomplexes depending on their end processing needs. While Ku and LigIV are sufficient to join blunt signal ends, ligation of coding ends necessitates the action of DNA-PKcs, Artemis, XRCC4 and XLF to open the hairpin formed by the RAG1/RAG2 complex. It is possible that multimerization requirements between these two substrates are different and thus, it would be insightful to test the effect that multimerization impairing mutations, like those on Ku80-CTD or in the XRCC4-XLF interaction region, have on V(D)J recombination and CSR. These studies may potentially reveal differences in the requirement for multimerization between programmed DSBs and radiation-induced DNA breaks.

3.1. Therapeutical uses

Due to the accumulation of mutations that they produce, defects in DNA repair mechanisms are associated with the development of several types of cancer. For instance, between 5-15% of hereditary breast, ovarian or pancreatic cancer contain mutations in HR genes whereas 3-4% of familial colon cancers contain mutations in mismatch repair genes [133, 134]. On the other hand, during tumor progression, cancerous cells become ever more dependent on DNA repair mechanisms to prevent their genome instability from inducing cell death. As a result, overexpression of DNA repair genes is frequently found in advanced stage cancers. For example,
DNA-PKcs is overexpressed in nasopharyngeal, colorectal and non-small cell lung carcinomas and its level of expression correlates with advanced tumor stages [135]. The use of chemo- and radio-therapy to treat tumors exacerbates this effect by further selecting cancerous cells with overactivated DNA repair mechanisms that can deal with the newly inflicted DNA damage, especially DSB, the most toxic type of lesion produced by these treatments.

A recent study has shown that Ku and XRCC4 expression can be used to predict the effectiveness of chemo- and radiotherapy in hypopharyngeal cancers. Tumors with lower Ku70 and XRCC4 expression correlated with higher survival rates after treatment [136]. Similar high correlations were obtained when studying DNA-PKcs and Mre11 expression in tumors treated with radiotherapy [137]. These results exemplify how strategies aimed at impairing NHEJ could radiosensitize tumor cells, increase treatment efficacy and improve patients’ outcomes. For instance, targeting DNA-PKcs with small molecule inhibitors (SMI) has hypersensitized cells to ionizing irradiation, and it has successfully delayed tumor growth in mice treated with radiotherapy [138].

The emerging role of multimerization during NHEJ raises the possibility of radiosensitizing cancerous cells by means of preventing multimerization of NHEJ factors. Therapeutic reagents designed to block important sites for multimerization are likely to impair NHEJ and thus enhance the sensitivity of cancerous cells to radiation and possibly others to DNA damaging chemicals. In this context, it would be paramount to investigate how multi-merization of NHEJ factors differs at telomeres and at DSB. These differences could be exploited to design reagents that block NHEJ without affecting their telomeric roles. This is particularly relevant for the Ku heterodimer, as human cell lines lacking Ku expression quickly die due to massive telomere loss [139]. A reagent that impairs Ku’s NHEJ without affecting its telomeric functions could radiosensitize tumor cells without compromising the viability of healthy cells. Similarly, possible differences between multimerization of NHEJ factors at sites of DNA damage with respect to physiologically programmed cuts during V(D)J and CS recombination could be used to generate molecular targeted therapeutic reagents that radiosensitize cancer cells without adversely affecting the patient’s immune system.

Acknowledgements

We would like to express our sincere gratitude to Dr. Alison Bertuch, Dr. Sandra Indiviglio, Dr. Jill Dewey and Katherina Alsina for their insightful revisions to this manuscript.

Author details

Michelle Rubin, Jonathan Newsome and Albert Ribes-Zamora*

*Address all correspondence to: ribesza@stthom.edu

Biology Department, University of St.Thomas, Houston, Texas, USA
References


[107] Lim HS, Kim JS, Park YB, Gwon GH, Cho Y. Crystal structure of the Mre11-Rad50-ATPgammaS complex: understanding the interplay between Mre11 and Rad50. Genes Dev. 2011 May 15;25(10) 1091-1104.


