Nonhomologous DNA end-joining (NHEJ) is the predominant double-strand break (DSB) repair pathway throughout the cell cycle and accounts for nearly all DSB repair outside of the S and G2 phases. NHEJ relies on Ku to thread onto DNA termini and thereby improve the affinity of the NHEJ enzymatic components consisting of polymerases (Pol μ and Pol λ), a nuclease (the Artemis-DNA-PKcs complex), and a ligase (XLF-XRCC4:Lig4 complex). Each of the enzymatic components is distinctive for its versatility in acting on diverse incompatible DNA end configurations coupled with a flexibility in loading order, resulting in many possible junctional outcomes from one DSB. DNA ends can either be directly ligated or, if the ends are incompatible, processed until a ligatable configuration is achieved that is often stabilized by up to 4 bp of terminal microhomology. Processing of DNA ends results in nucleotide loss or addition, explaining why DSBs repaired by NHEJ are rarely restored to their original DNA sequence. Thus, NHEJ is a single pathway with multiple enzymes at its disposal to repair DSBs, resulting in a diversity of repair outcomes.

Eukaryotic cells have evolved to repair multiple forms of DNA damage to maintain a high level of fidelity between cell divisions. Among types of damage, DNA double-strand breaks (DSBs)3 are particularly detrimental as they can result in insertions, deletions, or chromosomal translocations that are the primary transforming step in many human cancers. Pathological DSBs can arise from both exogenous (e.g. ionizing radiation or reactive oxygen species) or endogenous (e.g. DNA replication errors or incidental action by nuclear enzymes) sources. In some cases, DSBs are required as part of a physiological process, such as the breaks that occur during V(D)J recombination and immunoglobulin heavy chain (IgH) class switch recombination (1). Both pathological and physiological DSBs require efficient processes for repair that result in minimal to no change to the broken chromosome. Repair mechanisms can be largely divided between those that use extensive homology from a sister chromatid or homologous sequence elsewhere in the genome and those that use little to no homology. Both mechanisms require end processing by nucleases, utilization of DNA polymerases, and a final ligation step to complete repair of the broken DNA (Fig. 1). Nonhomologous DNA end-joining (NHEJ) was originally a phrase used to describe a type of illegitimate repair that utilizes little to no long homology (2) (we feel it unnecessary to include the word “canonical” or use the term “c-NHEJ” as we consider NHEJ a stand-alone pathway that does not need to be described in reference to separate alternative end-joining pathways that have their own distinct components). “Nonhomologous” could be misinterpreted as meaning completely homology-independent by a newcomer to the field, but up to 4 bp of microhomology during repair is common for NHEJ, and the term is simply meant to contrast with “homologous” recombination (HR), which can use several hundred base pairs of homology as a template for high-fidelity repair. In NHEJ, the DSB is first recognized by a heterodimer consisting of Ku70 and Ku80 (Ku). The DNA-dependent protein kinase catalytic subunit (DNA-PKcs) has a high affinity for DNA ends, which is even tighter when Ku is bound to that end (3). The nuclease, Artemis, exists in tight complex with DNA-PKcs within the cell and is likely recruited along with DNA-PKcs (4). Nucleotide addition can occur by the Pol X family polymerases, Pol μ and Pol λ. Finally, the DNA ligase IV complex, including XRCC4, XLF, and perhaps PAXX, carries out the critical ligation step for either strand of the DSB.

Importantly, NHEJ is an iterative process, where each of the DNA ends involved in the break can be acted upon by these components multiple times and in a different order (Fig. 51). Other important factors that dictate repair are the differential requirements for the various NHEJ proteins depending on the configuration of the DNA ends, which can include blunt ends, 5′ or 3′ overhangs, or ends containing adducts refractory to processing or ligation. Recent work has begun to systematically examine how various DNA end configurations are processed differently (5, 6). We briefly mention how NHEJ relates to the

---

1. Nicholas R. Pannunzio1, Go Watanabe1, and Michael R. Lieber2
2. From the Departments of Pathology, Biochemistry and Molecular Biology, and Molecular Microbiology and Immunology, Section of Molecular and Computational Biology, Department of Biological Sciences, Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, Los Angeles, California 90033
3. This work was supported by National Institutes of Health grants (to M. R. L.). This is the second article in the Thematic Minireview series “DNA double-strand break repair and pathway choice.” The authors declare that they have no conflict of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This article contains Figs. S1 and S2.

1. Both authors contributed equally to this work.
2. To whom correspondence should be addressed. Tel.: 323-865-0568; E-mail: lieber@usc.edu.

3. The abbreviations used are: DSB, double-strand break; PDB, Protein Data Bank; NHEJ, nonhomologous DNA end joining; Pol, polymerase; HR, homologous recombination; aa, amino acid; SSA, single-strand annealing; TdT, terminal deoxynucleotidyltransferase; 3′-PG, 3′-phosphoglycolate; PNK, polynucleotide kinase; CtIP, C-terminal-binding protein interacting protein; XLF, XRCC4-like factor; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; a-EJ, alternative end-joining; nt, nucleotide.

This is an open access article under the CC BY license.
other pathways of double-strand break repair, but our major focus is the NHEJ process. Therefore, readers are directed to the other works in this Thematic Minireview series for a detailed explanation of other DSB repair mechanisms.

**Overview of NHEJ in humans and its relationship with other pathways of double-strand break repair**

In human cells, NHEJ appears to repair nearly all DSBs outside of S and G2 cell cycle phases and even about 80% of DSBs within S and G2 that are not proximal to a replication fork (Fig. 1) (7). In late S and G2, HR is another major pathway for DSB repair, relying on more extensive homology tracts as a template for repair (8). When NHEJ is compromised due to the absence of one or more key protein components, the activities of other DNA end-joining pathways that typically involve more extensive end resection become apparent. Greater levels of 5’ end resection expose homologous sequences embedded on either side of a DSB, allowing for stable annealing of 3’ single-stranded DNA (ssDNA) that promotes more efficient joining and ligation (9). Although NHEJ usually requires ≤4 bp of microhomology, the alternative end-joining (a-EJ) pathway (also known as Pol θ-mediated end-joining or microhomology-mediated end-joining) (Fig. 1) (10), which utilizes the additional factors of poly(ADP-ribose) polymerase and DNA Pol θ, requires microhomology that ranges between 2 and 20 bp. Although NHEJ dominates DSB repair in most mammalian somatic cells, Pol θ-mediated events appear at an observable frequency in certain cell types (11), for certain repair events (12), and in some organisms (13). Greater levels of resection can further promote the non-conservative homology-directed repair pathway of single-strand annealing (SSA) that requires >25 bp of homology (Fig. 1) (14–17). Therefore, the mechanisms of NHEJ and HR occur on opposite ends of a spectrum with respect to homology usage with a-EJ and SSA occurring between them on a gradient of increasing levels of DNA end resection and homology usage (6).

A key reason for the dominance of NHEJ is that extensive DNA end resection is prevented by Ku binding (18), and the tight affinity and high abundance of Ku in cells increases the likelihood that Ku is the first protein to bind at a broken DNA end (6) (Fig. 1). A small protein called CYREN (cell cycle regulator of NHEJ (69 aa); also called MRI-2, a sub-peptide of C7orf49 (157 aa)) has been proposed to affect Ku DNA binding (19), although the data on CYREN effects on Ku binding are conflicting (20). Signaling factors appear to be important in controlling resection, as there is evidence that the DNA damage-response protein p53-binding protein 1 (p53BP1) is antagonistic to end resection, acting through a number of effector proteins (21, 22). 53BP1 and mediator of DNA damage checkpoint protein 1 (MDC1) are recruited to DSBs through a num-

**Figure 1. NHEJ in the context of other double-strand break repair pathways.** DNA double-strand breaks (DSBs) can be repaired by NHEJ, alternative end-joining (a-EJ), single-strand annealing (SSA), or homologous recombination (HR). Pathway choice and pathways other than NHEJ are discussed in other Minireviews in this thematic series. The name NHEJ originally arose to distinguish it from repair that requires extensive DNA homology (i.e. HR and SSA). Lengths of terminal microhomology (MH) between 1 and 4 bp are common in NHEJ, a-EJ is also called microhomology-mediated end joining (MMEJ) or Pol θ-mediated end joining (TMEJ). The major difference in the pathways is the requirement for significant DNA end resection. The p53-binding protein 1 (p53BP1) is a chromatin remodeler and a positive regulator for NHEJ. Although Artemis-DNA-PKcs can carry out some nucleolytic resection (typically ≤20 nt), the NHEJ pathway does not require extensive end resection, and the ends are protected from deeper resection by the binding of the Ku heterodimer (Ku70–80) to the DNA ends. By contrast, the C-terminal binding protein-interacting protein (CtIP) and the MRN (MRE11 (meiotic recombination 11)–RAD50–NBS1 (Nijmegen breakage syndrome protein 1)) complexes are involved in extensive 5’ to 3’ resection of regions of the duplex, and this generates stretches of ssDNA at DNA ends for a-EJ, SSA, and HR. SSA typically requires >25 bp of microhomology, whereas the requirement for a-EJ is typically ≤20 bp. Poly(ADP-ribose) polymerase 1 (PARP1) and Pol θ are important for a-EJ. Bloom syndrome RecQ-like helicase (BLM) and exonuclease 1 (EXO1) account for additional resection, and replication protein A (RPA) binds to ssDNA to promote the SSA and HR pathways. RAD52-mediated annealing of homologous sequence is key for the SSA pathway. XPF-ERCC1 cuts the remaining 3’ nonhomologous ssDNA prior to ligation by DNA ligase 1. By contrast, RAD51-mediated strand exchange with its association with BRCA1, BRCA2, and RAD54 is essential for facilitating the HR pathway.
THEMATIC MINIREVIEW: NHEJ for repair of double-strand breaks

...number of modified histone residues and appear to have distinct roles in DSB repair (8, 23, 24). Further work is required to elucidate specifically how 53BP1 recruitment inhibits extensive end resection. Overcoming this barrier to resection, however, is the first step to enable either a-EJ or SSA.

Following commitment to NHEJ, the nuclease, polymerase, and ligase components act on the DNA ends until repair is complete. Pathway commitment likely is not final until the strands of the break site are ligated, and if the DSB remains unrepaired, the repeated processing of ends may shift repair to another pathway. Below we provide a brief overview of the types of proteins that are involved in NHEJ and their functions, which applies to nearly all vertebrates.

The nucleases of NHEJ

Direct ligation of broken DNA ends is often impeded due to end incompatibility caused by mismatching overhangs or chemical modifications (Fig. 1). Therefore, following commitment to NHEJ, nucleases are required to process mismatched or modified ends to prepare them for ligation. This typically involves removing short regions of the 5′ or 3′ overhangs by either exonuclease or endonuclease processing to expose short regions of microhomology (≤4 nt) between the strands that can facilitate end joining. Extensive end resection (≥20 nt), which occurs to initiate HR or SSA pathways, is prevented by the presence of Ku, distinguishing the end processing of NHEJ from other DSB repair pathways. When DNA resection is required for NHEJ, DNA-PKcs is recruited in complex with the nuclease Artemis to Ku-bound DNA ends. DNA-PKcs undergoes autophosphorylation and activates Artemis (25, 26). Artemis then gains the ability to cut DNA ends at single-strand–to–double-strand DNA (ss–dsDNA) boundaries, which includes all overhangs and other structures such as gaps, loops, and bubbles that may arise due to mismatches between the two DNA ends being joined (27, 28).

Artemis is a member of the metallo-β-lactamase family of nucleases, containing the conserved metallo-β-lactamase and β-CASP domains. This family of nucleases has the ability to hydrolyze DNA or RNA in various configurations (29). In addition to an intrinsic 5′ exonuclease activity on ssDNA that does not require DNA-PKcs (30, 31), Artemis possesses a DNA–PKcs-dependent endonuclease activity on both 5′ and 3′ DNA overhangs of duplex DNA. Such overhangs often result due to pathological DSBs where breaks on opposite DNA strands occur in very close proximity. Also, Artemis endonuclease activity is essential for the hairpin opening step during V(D)J recombination (following cleavage by recombination activation genes, RAG1 and RAG2), and patients lacking Artemis suffer from severe combined immunodeficiency because of a V(D)J recombination defect in antigen receptor gene assembly (4, 32).

DNA-PKcs interacts with the C terminus of Ku80, which is highly dynamic and flexible (Fig. 2A). The final 12 amino acids of Ku80 are sufficient for interacting with DNA–PKcs (33, 34), but Ku-DNA–PKcs complex formation is very weak unless Ku is bound to a DNA end. The presence of Ku on DNA increases the binding affinity of DNA–PKcs for DNA ends by 100-fold (35). Following binding to the DNA end, DNA–PKcs autophosphorylates, thus activating the endonuclease activity of Artemis (4). This likely occurs when autophosphorylated DNA-PKcs phosphorylates the C-terminal inhibitory region of Artemis (aa 454–458), promoting the dissociation of the inhibitory region from the N-terminal catalytic domain (aa 1–7) (Fig. 2B) (25, 36).

It has been estimated that 20–50% of ionizing radiation-induced DSBs require Artemis for repair (37, 38). One possibility is that the remaining DSBs can be joined without the need of any nuclease (Fig. S2), but considering the number of nucleases present in the cell, it seems likely that other nucleases could be employed at incompatible ends, especially when Artemis is not present. Among those suggested to be involved in DSB repair include APLF, which is an abbreviation for Aprataxin and PNKP-like factor (also known as PALF) (39–41), flap structure-specific endonuclease 1 (FEN1), DNA replication helicase/nuclide 2 (DNA2), and exonuclease 1 (EXO1). In addition to nucleases, the Werner syndrome ATP-dependent helicase/nuclease (WRN) and the Bloom syndrome RecQ-like helicase (BLM) may also be involved in processing of DSB ends by creating a cleavage substrate for several of the aforementioned nucleases (42–44).

Another possible factor is the MRN complex (consisting of MRE11, RAD50, and NBS1), which is important for the resection step of the HR and SSA pathways to generate extensive 3′-terminated ssDNA overhangs. The intrinsic 3′→5′ exonuclease activity of the MRE11 component cannot generate these 3′-terminated overhangs by acting directly at a DNA end and relies on the C-terminal-binding protein interacting protein (CtIP) to stimulate MRN endonuclease activity to incise distal from the break. Next, the 3′ exonuclease activity can degrade DNA from the incision back toward the DNA end, thus creating the 3′-terminated ssDNA overhangs that can further undergo long range resection (e.g. by EXO1 or DNA2–BLM) (45, 46). This processing may have implications for the binding of Ku to DNA ends because MRE11 endonuclease activity occurs upstream of the Ku-bound DNA end.

CtIP is an important regulator of end processing as it not only stimulates MRN but also the long range resection by BLM and DNA2 (44). Importantly, CtIP is phosphorylated and active in S and G,

G2 (47), indicating that cell cycle is another factor that dictates nuclease involvement. Furthermore, the abundance and localization of these nucleases at DSB sites will determine which nucleases are responsible for the most resection at DSBs. Because Artemis is recruited to breaks by DNA–PKcs at the early stages of NHEJ, and because only limited resection occurs, Artemis appears to the primary nuclease for most NHEJ repair events (27).

The polymerases of NHEJ

Members of the Pol X family of polymerases participate in DSB repair by NHEJ. DNA Pol μ and Pol λ are the two members involved in NHEJ in the majority of human cells (48, 49). Each of these polymerases has an N-terminal BRCT domain that allows them to interact with Ku (Fig. 2) (50). Primary cells derived from mice with genetic knockouts of both Pol μ and Pol λ exhibit little or no sensitivity to ionizing radiation, although knockouts in cell lines can have some deficit in DSB repair in some assays (51, 52). Pol μ and Pol λ can...
incorporate both dNTPs and rNTPs (48, 49), with any incorporated ribonucleotides subsequently removed by base excision repair (53). Importantly, both Pol μ and Pol λ can incorporate nucleotides in a template-dependent or template-independent manner (51), although template-independent insertion by Pol μ appears stronger than that of Pol λ (54, 55). Both of these polymerases appear to be able to use an unstable primer-template junction, such as would exist during intermediate stages of NHEJ. The activity of these polymerases further explains the high level of diversity that can occur at NHEJ junctions and demonstrates that although resection is one way of generating short stretches of homology between broken DNA ends, template-independent nucleotide addition of one or both broken DNA ends is another.
THEMATIC MINIREVIEW: NHEJ for repair of double-strand breaks

DNA polymerase β (Pol β) is another member of the Pol X family, but it lacks a BRCT domain (56), which is a likely reason why it is not involved in NHEJ. The final known member of the Pol X family is terminal deoxynucleotidyltransferase (TdT). TdT is only expressed in early B- and T-lymphocytes, making it most relevant to the NHEJ repair that occurs during V(D)J recombination, where it has a major role in promoting immunoglobulin diversity. DNA polymerases outside the Pol X family are able to incorporate nucleotides during NHEJ, but only in a template-dependent manner (16, 17, 57–59).

The ligase complex of NHEJ

DNA ligase IV (Lig4) functions exclusively in NHEJ, making it a central component of the repair process. Lig4 acts in complex with the X-ray repair cross-complementing 4 (XRCC4) enzyme (9), which stimulates Lig4 enzyme activity in biochemical assays (60). Loss of either Lig4 or XRCC4 severely compromises NHEJ. Several other factors have also been implicated for efficient ligation. A screen for XRCC4-interacting factors yielded the XRCC4-like factor (XLF; also known as Cernunnos), a 33-kDa protein with weak sequence homology and structural similarity to XRCC4 (61–63). The N-terminal head domain of XLF interacts with the N-terminal head domain of XRCC4 (62) allowing XLF to complex with XRCC4-Lig4 (Fig. 2). This interaction would presumably stabilize the juxtaposition of the DNA ends prior to covalent ligation, but this is still an area of active investigation. Another protein found to have structural similarity to XRCC4 is the 22-kDa protein PAXX (paralog of XRCC4 and XLF) (64, 65). The C terminus of PAXX (aa 199–201) interacts with Ku, and similar to XLF mutants, PAXX mutants are more sensitive to ionizing radiation and DSB-inducing agents (Fig. 2) (66, 66, 67).

Accessory proteins of NHEJ: Tyrosyl DNA phosphodiesterase 1, polynucleotide kinase, and aprataxin

Although the above proteins can carry out a majority of the NHEJ reactions, some circumstances require the activity of other proteins to chemically modify DNA ends to make them suitable for repair. For example, tyrosyl DNA phosphodiesterase 1 (TDP1) is the only identified enzyme that can specifically process the 3’-phosphoglycolates (3’-PG) that can form as a by-product of up to 10% of ionizing radiation-induced DSBs (68, 69). Ends with 3’-PG adducts are unligatable and must be removed for NHEJ to proceed. Polynucleotide kinase (PNK) and aprataxin are two more factors that may be enlisted in DSB repair by NHEJ. Human PNK possesses both kinase and phosphatase activity. Phosphorylation by PNK is necessary when a 5’ end lacks a phosphate group, and the phosphatase activity is important for removing 3’ phosphates that can arise following some types of oxidative damage (70). Aprataxin is employed when Lig4 initiates but does not complete a covalent join, resulting in an aborted ligation product where the AMP group remains covalently bound to the 5’ strand of one of the DNA ends. In this case, the deadenylation reaction catalyzed by aprataxin is required to remove the AMP group (71). Following phosphorylation of XRCC4 by CK2, both PNK and aprataxin can bind to XRCC4 via their forhead-as-

Optimal NHEJ component utilization is influenced by DNA end configuration

NHEJ is a single pathway, but the DNA end configurations at a given DSB determine which NHEJ components are most important for efficient ligation. In other words, NHEJ has several enzymes at its disposal, but it does not need to engage all of them unless presented with certain DNA end configurations (5). Even the core NHEJ components may load and act in various combinations, highlighting the flexibility of NHEJ and explaining the diversity of repair products generated for the very same DSB configuration and DNA end sequence. This model is supported by many structural and biochemical studies demonstrating the different routes DNA end processing can take to reach a ligatable joint (Fig. 3). The stability of this ligatable joint is greatly enhanced when base pairing of ssDNA from either side of the break can occur via microhomology, although for NHEJ this microhomology need not be extensive, as even a single base pairing (even a non-Watson-Crick base pairing) will increase the stability enough to improve ligation efficiency a few fold over what is observed for NHEJ at blunt ends (74). In some cases, simple breathing of the DNA ends that exposes a complementary base pair between two broken ends may be adequate for repair, whereas in other cases more extensive processing by nucleases and polymerases may be required (16, 17).

The iterative nature of NHEJ means that multiple components can act on a single DSB during multiple rounds of processing (Fig. 3). Nucleases can remove nucleotides from a DNA end, with Pol μ subsequently adding nucleotides to that very same DNA end. Similarly, XRCC4-Lig4 can successfully ligate one DNA strand of a DSB only to have Artemis-DNA-PKcs reverse this by cleaving the newly ligated strand at the DNA gap generated by the ligation. Therefore, use of one set of components is not mutually exclusive to the use of other components, and all are active and in play as long as a DSB remains incompletely repaired.

Blunt-end ligation by Ku-XRCC4-Lig4

Biochemical studies have demonstrated that Ku is required for the efficient joining of blunt DNA ends lacking microhomology by NHEJ. When a ligatable joint is formed using exposed microhomology, however, Ku may not be necessary, indicating that Ku becomes less important as ends are able to form a thermodynamically stable joint through terminal base pairing (55). Ku is highly abundant in cells and has a high affinity for DNA ends ($K_u = 6 \times 10^{-10}$ M), allowing it to quickly respond to a break and promote the binding of XRCC4-Lig4 to the DNA ends (75). The C terminus of Lig4 contains two BRCT domains that allow it to bind to two Ku complexes, conceivably one attached to each of the DSB ends (76). The region between these two BRCT domains of Lig4 carries the interaction domain
that binds a homodimer of XRCC4 (Fig. 2D) where the 2 to 1 ratio of XRCC4 to Lig4 further stabilizes the bridging between the two DNA ends (77–79). The Ku-DNA complex is able to efficiently bind and thereby recruit other NHEJ components. An iterative processing occurs to make the two broken DNA ends optimal for ligation. Several types of processing performed by the Artemis-DNA-PKcs complex or DNA polymerases are shown in the white boxes along the large green circle. It would be difficult to represent all the possible DNA end configurations and every type of enzymatic processing in one figure; therefore, this depiction is not meant to be comprehensive but is merely to highlight some of the possibilities with the key components for each process indicated in parentheses. Any of these processes can occur to either end of a break in any order and multiple times. Once XRCC4-Lig4 is able to successfully ligate across a break, an intermediate with one strand ligated can form. Ligation of the second strand will complete repair. Alternatively, the gapped intermediate generated by ligating one strand has two ss–dsDNA boundaries, and Artemis-DNA-PKcs can cut at either boundary to generate a new DSB, thereby returning the ends to the iterative processing step where they can undergo further alterations.

The nucleases of NHEJ can process multiple DNA end configurations

Artemis has been implicated as the major nuclease involved in NHEJ when such activity is required. Although the role Artemis plays in DNA hairpin opening during V(D)J recombination is well-characterized, its role in NHEJ is now beginning to be understood. Recent biochemical studies have revealed that the ligation of incompatible overhangs is strongly stimulated in the presence of the Artemis-DNA-PKcs complex. Therefore, Artemis is recruited to process various DNA overhangs at broken DNA ends to promote formation of a stable ligatable joint. This makes sense when one considers that DNA hairpins are structurally similar to DNA overhangs, due to a sterically-constrained hairpin tip that results in only transient base pairing of the terminal base pairs (4 nt), thus creating a ss–dsDNA boundary (85). This ability of Artemis to act at ss–dsDNA boundaries gives it the flexibility to process a number of DNA end configurations.

The endonuclease activity of the Artemis-DNA-PKcs complex can remove both 5’ and 3’ DNA overhangs to create DNA end structures that can be ligated by the XRCC4-Lig4 complex (50, 86) (Fig. 3). At 5’ overhangs, Artemis cuts directly at the ss–dsDNA boundary, but when processing 3’ overhangs and DNA hairpins, Artemis preferentially leaves a 4-nt 3’ overhang. Long 5’ and 3’ overhangs can also be endonucleolytically pro-
cessed by Artemis, and this may be useful to make microhomology embedded within the overhang available for annealing to create a stable ligation joint (5). These observations suggest a model in which Artemis-DNA-PKcs binds to the ss–dsDNA boundary to occupy 4 nts along the single-stranded segment at the boundary followed by nicking on the 3′ side of the 4 nts (27).

In addition to overhangs, evidence also shows that when blunt DNA ends breathe between a closed, fully hydrogen-bonded state to an open, partially hydrogen-bonded state, they form ss–dsDNA boundaries upon which Artemis can act (27). Repair of such ends is relevant as blunt DNA ends may be generated by chemotherapeutic agents, reactive oxygen species, or ionizing radiation (87). Furthermore, breathing allows the Artemis-DNA-PKcs complex to resect into the duplex to generate short overhangs that can form microhomology (5), explaining why even NHEJ of blunt ends can display nucleotide loss at repair junctions. Still, the fact that Artemis-DNA-PKcs does not strongly stimulate the ligation of blunt-ended DNA suggests that even though Artemis-DNA-PKcs is able to resect at blunt ends, these ends are usually joined directly without resection (5, 27).

The versatility of Artemis to act at many different types of ends leads to a unifying model explaining the essential structural features of all DNA substrates at which Artemis functions. Although it may appear that Artemis has the ability to recognize a number of different structures, in fact it is one structure, an ss–dsDNA boundary, that is recognized in a variety of different forms. 5′ and 3′ overhangs, hairpins, and blunt ends in an open state all have potential regions of ss–dsDNA that can act as contact points for Artemis (28). The Artemis active site can then act within the single-stranded portion of the overhang or the hairpin to achieve hydrolysis of the phosphodiester backbone. Although this model must await the elucidation of a DNA–Artemis structure, it explains the diversity of cutting patterns of Artemis.

Besides the role in processing DNA overhangs, Artemis appears to be necessary for removing damaged DNA from broken ends (Fig. 3). When ionizing radiation-induced DSBs bear a 3′-PG terminus (88–90), for example, these DNA ends are unable to undergo ligation because this step requires a 3′-hydroxyl on one end and a 5′-phosphate on the other. TDP1 is able to remove these 3′ modifications; however, TDP1 mutant cells are only marginally radiosensitive compared with Artemis mutants, and it has been demonstrated biochemically that the Artemis-DNA-PKcs complex is able to process these ends (91, 92). This suggests that Artemis can work with or in place of TDP1 to repair the large number of DSBs that can occur following radiation exposure.

The finding that the C-terminal region of Artemis (aa 485–495) interacts with the N-terminal head domain of Lig4 (Fig. 2B) (93–95) adds a further dimension to the role Artemis may play in NHEJ. Although the DNA-PKcs–independent 5′ exonuclease activity has been described, recent data show that Artemis has a DNA-PKcs–independent 3′ endonuclease activity stimulated by XRCC4-Lig4 (96). The interaction between Lig4 and the C-terminal regulatory region of Artemis may recruit Artemis and alter the protein conformation, permitting endonuclease activity without the need for activation by DNA-PKcs. In addition to its crucial role in ligating a stable joint intermediate, the extreme radiosensitivity of Lig4 mutants may be due to its ability to stimulate or recruit various NHEJ components to a DSB.

The DNA polymerases of NHEJ work to create a stable ligation joint

DNA polymerases can serve two important roles in NHEJ: fill-in synthesis of gaps and nucleotide addition to broken DNA ends. Both processes can enhance formation of a stable intermediate for ligation by XRCC4-Lig4. The DNA polymerases Pol μ and Pol λ are recruited to the DNA end by interaction of their N-terminal BRCT domain with the Ku-DNA complex (Fig. 2C) (50). Pol μ primarily adds nucleotides in a template-independent manner, whereas Pol λ primarily has template-dependent polymerase activity, although limited template-independent activity has been reported (54). Pol μ, and also TdT, carries a protein domain, loop 1, that affects association with a DNA template through hydrogen bonding and allows for template-independent nucleotide addition (56, 97).

The template-dependent activity of Pol λ is mostly required when long ssDNA ends are annealed with terminal microhomology, leaving a gap. Fill-in synthesis of this gap will further stabilize the annealed intermediate and promote the ligation (52, 55). When the 3′ overhangs are mismatched and therefore unable to form an annealed intermediate, Pol λ has little effect on NHEJ because there is no DNA template to act upon (5).

Pol μ strongly promotes the ligation of incompatible 3′ overhangs in reactions containing only the Ku-XRCC4-DNA ligase 4 complex (55). By adding nucleotides to the ends of these overhangs in both template-dependent and template-independent mechanisms, Pol μ generates regions of microhomology for subsequent annealing and ligation (55). Nucleotide addition can occur on 3′ overhangs as short as 1 to 2 nts (52). In biochemical reactions containing Artemis, the joining of two mismatched 3′ overhangs is strongly stimulated by Pol μ, promoting the formation of terminal microhomology with limited processing by Artemis (Fig. 3) (5). Interestingly, sequencing of NHEJ junctions reveals that if the ends are compatible, meaning they already share microhomology, nucleotide addition by Pol μ does not occur or is limited (5). This illustrates once again that ends capable of forming a thermodynamically stable intermediate are ligated efficiently without having to recruit additional factors.

XLF and PAXX stimulate ligation by the XRCC4-Lig4 complex

XLF and PAXX are the most recently characterized NHEJ factors shown to support ligation by the Lig4 complex. Both XLF and PAXX share structural similarity with XRCC4 (62, 64). Individual XLF and PAXX mutants display only a mild phenotype, but XLF PAXX double mutants are synthetically lethal in mice and reduce V(D)J recombination in human B-lymphocytes (98–101), suggesting that although they may be redundant, at least one is necessary for efficient repair by NHEJ. The main purpose of XLF and PAXX appears to be in providing additional structural support to stabilize two DNA ends, thereby enhancing the ability of XRCC4-Lig4. This likely occurs.
Figure 4. Structural aspects of NHEJ. Select reported three-dimensional structures of NHEJ components are shown. Protein structures are shown in ribbon representation except for DNA-PKcs and an XRCC4-XLF filament, which are shown in surface representations. A, crystal structure of Ku70/80 heterodimer alone at 2.7 Å (PDB code 1JEQ) is shown in the middle (109). Ku70 and Ku80 are shown in red and yellow, respectively. B, these toroidal Ku70/80 proteins bind to broken dsDNA ends to form a Ku70/80-DNA complex (solved at 2.5 Å) (PDB code 1JEQ) (109). C, crystal structure of DNA-PKcs at 4.3 Å is shown in the center (PDB code 5LUQ) (84). The DNA-PKcs is color-coded as follows: N terminus (blue); circular cradle (green); head comprising FAT region (purple); kinase (yellow); FRB (orange); FATC (light pink). D, DNA-PKcs binds the Ku70/80-DNA to form a DNA-PK complex. A 6.6 Å cryo-EM structure of DNA-PK holoenzyme is shown (PDB code 5Y3R) (111). E, structure of Artemis has not been reported yet. F, crystal structure of the catalytic region of DNA ligase IV (DBD-NTD-OBD) in complex with an Artemis fragment (aa 485–495) was solved at 2.4 Å (PDB code 3W1B) (94). The Artemis fragment is shown in orange and interacts with the DNA-binding domain (DBD), which is shown in violet. The nucleotidyltransferase domain (NTD) is shown in cyan. The catalytic lysine (Lys-273), which forms a covalent AMP-lysine intermediate, is shown as a sphere, and a possible αP0 is attached to the lysine. The OB-fold domain (OBD) is shown in blue. G, crystal structure of the complex of XRCC4 homodimer and the BRCT repeats of ligase IV (Lig IV) at 2.4 Å is shown (PDB code 3II6) (107). Each XRCC4 molecule is shown in cyan and green. Two BRCT domains are shown in red. H, crystal structure of XRCC4(1–224)/XRCC4(1–157) complex (both are homodimers) at 3.94 Å is shown (PDB code 3WRJ) (110). The XRCC4 homodimer is shown in cyan and green. The XRCC4 homodimer is shown in yellow and orange. I, this XRCC4-XLF complex can form filaments, shown in the same color scheme at the left top corner, which might bridge DNA ends. J, crystal structure of PAXX homodimer at 3.45 Å is shown in cyan and purple (PDB code 3WTF) (64). Note that Ku70/80 bound on the DNA end can recruit XRCC4-ligase IV complex, and Ku70/80 also directly interacts with and recruits XLF and PAXX through their C termini. Also note that structures of the Pol X family polymerases are not shown here due to a space limitation. The figure was created using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

in a subset of NHEJ repair reactions where the two broken ends are incompatible and lack the thermodynamic stability provided by annealed microhomology.

Homodimers of XLF bind directly to XRCC4 via an N-terminal head domain (102). This head domain also allows XLF to interact with the Ku-DNA complex (103). In biochemical reactions containing only Ku and the XRCC4-Lig4 complex, XLF was shown to only stimulate the ligation of short, incompatible 3’ overhangs (55). In another study, however, XLF was shown to promote the ligation of all mismatched and noncohesive overhangs in the presence of Ku, DNA-PKcs, and XRCC4-Lig4 (104). Although it is possible that DNA-PKcs could affect XLF interactions, it is also possible that differences in the DNA substrates used in each study affect the outcome because the study involving DNA-PKcs used ~3 kb of linearized plasmids and the other used fragments of ~70 bp. Although further study is required to fully understand the major role of XLF in NHEJ, it seems that XLF promotes annealing of at least some incompatible substrates.

Genetic studies in mice complement these biochemical findings as it was found that an XLF DNA-PKcs double knockout is synthetically lethal. Interestingly, a Ku70 knockout rescues this synthetic lethality (105). Similar to a study showing that a Ku80 deletion rescues the lethality of a Lig4 knockout (106), this demonstrates that several NHEJ factors are epistatic to Ku. Loss of both XLF and DNA-PKcs must severely impair the ability to repair a DSB by NHEJ. Further genetic studies and analysis of DSB repair junctions in these deficient mice will provide more information as to the critical role of XLF.

Like XLF, PAXX also forms homodimers, and its C terminus has been found to associate with Ku (Fig. 2D) (64, 65). In reactions containing only Ku and the XRCC4-Lig4 complex, PAXX was shown to promote the ligation of two blunt ends (64). In some cases, XLF and PAXX may work together to stabilize DNA ends. In reactions containing Ku, XRCC4-Lig4, and XLF, PAXX promoted the ligation of a blunt end to a 3’ overhang (65). Interestingly, a more recent biochemical study showed that if Artemis and Pol μ are included, PAXX does not stimulate NHEJ for 3’ overhangs, but it does for 5’ overhangs (5), indicating that the role of PAXX may be to stabilize substrates that cannot generate microhomology by end processing or nucleotide addition.

THEMATIC MINIREVIEW: NHEJ for repair of double-strand breaks
THEMATIC MINIREVIEW: NHEJ for repair of double-strand breaks

Structural biology of NHEJ

There has been remarkable progress in determining progressively higher resolution structures of some of the NHEJ proteins or, at least, portions of them (Fig. 4) (48, 49, 64, 84, 94, 107–112). Readers should also refer to detailed reviews about the structural aspects of the interactions of ligase IV with XRCC4, XLF, and Artemis (113, 114). However, we still lack a convincing comprehensive view of how the enzymatic components are positioned at a single DNA end or at a pair of DNA ends. We also do not know the relative position of each component relative to most of the others in a large multiprotein complex during NHEJ.

Recently, a cryo-EM structure of DNA-PK was reported at 6.6 Å by docking the available crystal structures of DNA-PKcs (at 4.3 Å) (84) and Ku70/80-DNA complex (at 2.5 Å) (109). This finally allowed positioning of DNA-PKcs relative to the Ku70/80-DNA complex. However, statistics on the structural analysis must be much improved, and the position of quite a number of side chains of DNA-PKcs (>500 amino acids) is still questionable (84, 111, 115). Moreover, we still do not know how the C-terminal domain of Ku80 interacts with DNA-PKcs, and a structure of Artemis has not yet been reported. Furthermore, because some of the reported structures lack their C-terminal portions (e.g. Ku, XRCC4, XLF, and PAXX), an understanding on how these flexible regions work as full-length molecules will be critical for understanding the function of these complexes (64, 109–111).

Higher order structures have also been proposed for some NHEJ components. For example, the Lig4 complex, including XRCC4, XLF, and sometimes PAXX, has been proposed to form a sleeve around the DNA duplex (116–118), but the precise geometry is still not clear. It will be interesting to determine how such models will include Ku, DNA-PKcs, Artemis, and the polymerases μ and λ.

In many ways, the major future questions will require increasing reliance on structural insights.

Concluding comments

DNA DSBs are potentially lethal events that must be repaired in a manner that does not compromise genome integrity. NHEJ is the major pathway that repairs DSBs in mammalian cells. DSBs can occur due to various pathological or physiological events; however, the configuration of the DNA ends at breaks is not uniform. Therefore, NHEJ must be highly flexible so that it can deploy multiple enzymes to process the various types of DNA ends it may encounter. Biochemical and genetic studies have provided mechanistic insight into which NHEJ proteins are utilized, depending on the DNA end configuration. Two blunt DNA ends may only require Ku and XRCC4-Lig4 for joining, whereas incompatible 3’ ends may require processing by Artemis-DNA-PKcs, and incompatible 5’ ends may require XLF or PAXX for additional structural support. Time is likely a critical factor as the longer a break remains, the more accessory NHEJ factors may be recruited to a break in an attempt to repair it.

Many attempts have been made to subdivide the NHEJ pathway based upon the diversity of joining products that occur. However, this diversity of products highlights the flexibility of the NHEJ pathway. Repair by NHEJ does not mean a precise join because the activity of Artemis can lead to nucleotide loss, and the activity of Pol μ can lead to nucleotide gain. Also, the term “nonhomologous” was not meant to imply a total lack of homology usage in repair, as up to 4 nts of microhomology is typical for NHEJ repair. Instead, it was only meant to distinguish NHEJ from HR, which can use several hundred base pairs of homology during repair. Still, NHEJ is far from being completely understood, as evidenced by the discovery of new factors (PAXX) and new activities of known factors (Artemis). Continued research in this area will help elucidate why NHEJ is the dominant repair pathway in mammals and reveal more factors that contribute to DSB repair.

Acknowledgment—We apologize for any work that we have overlooked.

References


THEMATIC MINIREVIEW: NHEJ for repair of double-strand breaks


The synthetic lethality between murine DNA repair factors XLF and DNA-PKcs is rescued by inactivation of Ku70. *DNA Repair* 57, 133–138


