Overview of Nonhomologous DNA End Joining (NHEJ)²

All living cells have mechanisms for repairing double-strand DNA breaks (DSBs). Pathologic (disadvantageous) DSBs arise when the replication fork encounters a nick. Ionizing radiation particles create clusters of reactive oxygen species along their path, and these create DSBs. Reactive oxygen species themselves may cause DSBs. For dividing mammalian cells in culture, 5–10% appear to have at least one chromosome break (chromatid gap) at any one time (1). Hence, the need to repair DSBs arises commonly (2).

There are two primary pathways for the repair of DSBs (Fig. 1). Homologous recombination occurs in late S or G₂ of the cell cycle when the sister chromatid is close in proximity (3). NHEJ is the major pathway for the repair of DSBs because it can function throughout the cell cycle and because it does not require a homologous chromosome (3, 4). Rather, NHEJ involves rejoining of what remains of the two DNA ends, and the mechanism has evolved in a manner that tolerates nucleotide (nt) loss or addition at the rejoining site.

Because of the few nucleotides of resection and random addition necessary to get the two DNA ends into a ligatable configuration, NHEJ is distinctive among major DNA repair pathways for its imprecision. Hence, NHEJ leaves “information scars” at most sites of repair in vertebrates. The positive aspect of NHEJ is that the phosphodiester backbone and structural integrity of the chromosome are restored at sites that would otherwise result in loss of several hundreds of genes on entire chromosomal arms or segments. Attendant with the imprecision of NHEJ is the accumulation of randomly located mutations over time in the genome of each somatic cell of an organism.

Like most DNA repair processes, there are three enzymatic activities required for repair of DSBs by the NHEJ pathway: (a) nuclease(s) to remove damaged DNA, (b) polymerases to aid in the repair, and (c) a ligase to restore the phosphodiester backbone (Fig. 2). In vertebrates, the Artemis-DNA-dependent protein kinase catalytic subunit (DNA-PKcs) complex becomes active as a 5’- and 3’-endonuclease when DNA-PKcs binds to a DSB DNA end. Polymerases μ and λ are two of the known polymerases for NHEJ. A complex of XLF (Cernunnos), XRCC4, and DNA ligase IV composes the ligase for NHEJ.

Key Components and the Order of Their Action in the Mechanism of NHEJ

When a DSB occurs during G₀, G₁, and early S phase, the Ku heterodimer (Ku70/Ku80) appears likely to be the first protein to bind. (During late S or G₂, there may be some competition between NHEJ and homologous recombination, although this is still an active area of investigation (5).) Ku must change conformation once it binds DNA because its interactions with other proteins such as DNA-PKcs are much stronger once the Ku-DNA end complex has formed (6). Ku is capable of interacting with the nuclease (Artemis-DNA-PKcs), the polymerases (μ and λ), and the ligase (XLF-XRCC4-DNA ligase IV). Hence, one can think of Ku as a tool belt protein that can stabilize any of a number of enzymatic activities at a DNA end (Fig. 2). One might assume that the nuclease, polymerases, and ligase function in this order (Fig. 3), and in some of the simpler scenarios, this probably occurs. However, each of these enzymes has a range of flexibility in their function that permits the NHEJ process to go to completion in any of a large number (hundreds) of ways, even when starting with two identical DNA ends. This flexibility accounts for the very diverse number of results, with some ends showing nt loss (1–10 nt, typically) and some joining sites (junctions) showing untemplated nt addition (0–3 nt, typically) (Fig. 2, lower).

The next three sections describe in detail the flexibility of the nuclease, polymerase, and ligase components in NHEJ.
MINIREVIEW: Mechanism of Human Nonhomologous DNA End Joining

The Artemis-DNA-PKcs Complex in the Nucleolytic Processing Steps of NHEJ

DNA-PKcs can bind to DNA ends with a $K_D$ of $3 \times 10^{-9} \text{ M}$, but this affinity improves to $3 \times 10^{-11} \text{ M}$ at a Ku-DNA end complex (7). Artemis and DNA-PKcs exist as a complex within cells (8), and this complex binds to Ku-DNA end complexes. When Ku moves internally, this permits DNA-PKcs to contact the DNA end, which then activates the serine/threonine kinase activity of DNA-PKcs (9–11). Activation of the kinase activity represents one of the simplest signal transduction systems because it permits DNA-PKcs to phosphorylate itself and Artemis (8, 12). The autophosphorylation of DNA-PKcs causes a conformational change in DNA-PKcs that regulates access by other NHEJ proteins (12–15). This conformational change in DNA-PKcs may alter the conformation of Artemis because now Artemis can function as a 5'- or 3'-endonuclease at overhangs (supplemental Fig. 1) (8, 16). This conformational change in Artemis also permits it to function as an endonuclease at a variety of other single/double-strand DNA structures, including DNA hairpins (8, 17), which turns out to be critical in the gene rearrangement process of the vertebrate immune system (called V(D)J recombination). One DNA-PKcs molecule can autophosphorylate itself in cis or another DNA-PKcs molecule in trans, and the relative ratio of cis- versus trans-phosphorylation is not known (supplemental Fig. 2) (18). A subset of DSBs created by ionizing radiation cannot be repaired without Artemis (19).

Polymerase (pol) X Polymerases in Template-independent and Template-dependent Synthesis Steps of NHEJ

Genetic and biochemical evidence for the role of pol X polymerases exists in yeast and mammalian cells (20, 21). In yeast, Pol4 is the only pol X polymerase. In mammalian cells, the pol X family consists of pol $\beta$, $\mu$, and $\lambda$ and terminal deoxynucleotidyltransferase (TdT). The latter three all contain BRCT domains, and all three function in NHEJ, whereas pol $\beta$ does not. pol $\mu$ and $\lambda$ share the most homology with Saccharomyces cerevisiae Pol4, and both appear to be expressed in mammalian somatic cells. TdT is expressed only in pre-B and pre-T cells. $\text{pol } \lambda$, pol $\mu$, and TdT share a range of structural similarities that correlate with their ratio of template-dependent versus template-independent synthesis (22). TdT carries out tem-
plate-independent synthesis, and pol α is almost exclusively template-dependent. However, pol μ can carry out both template-dependent and template-independent synthesis in the presence of the physiologic divalent cation Mg2+ (23). pol μ and, to a greater extent, pol λ show some degree of template slippage, which can result in the generation of direct repeats, a feature seen at some sites of NHEJ (supplemental Fig. 3). The template-independent addition by pol μ could conceivably result in fold-back at the region of addition, followed by synthesis using the same strand as a template (supplemental Fig. 3, lower reaction). This could result in the generation of inverted repeats. The generation of direct and inverted repeats at a subset of NHEJ events in vivo is called T-nucleotide addition, where “T” stands for “template” (24). The flexibility of pol μ and λ might account for T-nucleotides. TD T functions only during V(D)J recombination to add random nucleotides (called N-nucleotides) so as to increase the junctional diversity during the generation of the antigen receptor repertoire. The template-independent synthesis by pol μ is the likely basis for occasional antigen receptor junctional additions in lymphocytes from animals in which TD T has been knocked out (25, 26).

One might wonder why pol μ, a polymerase present in all somatic cells, would add nucleotides randomly. The evolutionary advantage of this becomes clear when considering a DSB in which there is no terminal microhomology between the two DNA ends. Random addition has the benefit of potentially generating 1 or 2 nt of terminal microhomology, permitting more efficient annealing of the ends and thereby facilitating NHEJ. Hence, use of terminal microhomology occurs in some fraction of NHEJ events when it exists between the two ends, and pol μ may manufacture such terminal microhomology when it does not exist at the two DNA ends. In addition, when 3’ overhangs are involved, pol μ may be more robust at this type of fill-in from a region of minimal end-to-end annealing than any other polymerase (27, 28). Data for pol μ and the XRCC4-DNA ligase IV complex suggest that these two activities can bind a single Ku-DNA complex at one time (29, 30), raising the possibility of coordination of polymerase and ligase steps (29, 30).

**XLF, XRCC4, and DNA Ligase IV in the Ligation Steps of NHEJ**

DNA ligase IV can ligate double-stranded DNA molecules that have compatible overhangs or that are blunt (31). XRCC4 stabilizes the ligase IV protein in cells and improves its joining activity by increasing the efficiency of the adenylation of the ligase IV (31, 32). XLF stimulates the ability of XRCC4-DNA ligase IV to ligate in the presence of the physiologic divalent cation Mg2+ (33).

XRCC4-DNA ligase IV has a remarkable degree of flexibility, and there are several facets to this flexibility (supplemental Fig. 4). First, XRCC4-DNA ligase IV can ligate one strand independent of the other strand (34). For example, if the bottom strand is unligatable because of an uncleaved flap or a 5’-OH group, the top strand can still be ligated. Second, XRCC4-DNA ligase IV can ligate across gaps of even several nucleotides (23). Third, when Ku is present, XRCC4-DNA ligase IV can ligate some incompatible DNA ends that have short overhangs, albeit at lower efficiency than the ligation of ends that share 1 or more nt of terminal microhomology (23, 35). All of these aspects of the flexibility contribute to the range of DNA end configurations that XRCC4-DNA ligase IV can handle in its challenging role as the only ligase optimally suited for the repair of DSBs. Given that the two DNA ends of the DSB are likely to be in close proximity, each DNA end is likely to have a Ku molecule bound and associated with any of the three enzymatic activities, permitting nuclease, polymerase, and ligase action concurrently (supplemental Figs. 5 and 6).

**Evolutionary Advantage of NHEJ Flexibility**

At damage sites due to ionizing radiation, nt have been irreversibly lost at the instant of impact. Therefore, additional loss of a few more nt during the NHEJ repair process is of limited additional consequence; this is especially the case compared with the consequence of not restoring the chromosomal integrity. Tolerance of some level of imprecision has likely permitted the nuclease, polymerases, and ligase of NHEJ to evolve the substantial substrate and catalytic flexibility that they have.

Vertebrates have taken advantage of the imprecision and flexibility of NHEJ in the generation of antigen receptors for the adaptive portion of the immune system. In V(D)J recombination, the imprecision at the V-to-D and D-to-J joining sites (junctions) markedly increases the amount of potential diversity that would otherwise be limited simply to the various combinations of V, D, and J segments.

Recent work has shown that many prokaryotes have an NHEJ system that may be somewhat simpler than the one in eukaryotes, but which serves the same function (36–38). Hence, nearly all living cells have an NHEJ system, illustrating the importance of this general aspect of DNA repair.
MINIREVIEW: Mechanism of Human Nonhomologous DNA End Joining

Is NHEJ a Collection of Subpathways and Are There Alternative End Joining Pathways?

Some DNA end configurations might require only the ligase complex, XLF-XRCC4-DNA ligase IV. Other end configurations might require the polymerase complex followed by the ligase, nuclease, and polymerase can each function without Ku, but at substantially reduced efficiency.

This latter way of viewing the flexibility of NHEJ may explain some in vivo data suggesting a DNA-PKcs-independent pathway or other independencies. In mice lacking one of the NHEJ components, especially DNA ligase IV, the presumption has been that the entire NHEJ pathway is blocked (41). However, the other NHEJ components are still present, and under circumstances in which NHEJ cannot go to completion, enzymes from other pathways might be expected to participate, to the extent that their enzymatic properties permit. For example, in a ligase IV-null cell, if the polymerase activities are able to synthesize for sufficient lengths to separate the top strand nick from the bottom strand nick (with duplex DNA between the nicks), then ligase I or III would be adequate to complete the joining. Outside of the context of an NHEJ mutant cell, this “backup” use of another ligase might be very rare. The rate of these alternatives or backup routes to joining relative to wild-type NHEJ is not clear; however, these alternatives do not appear capable of functioning at the level of NHEJ in response to ionizing radiation or V(D)J recombination carried out by wild-type RAG proteins (42). In class switch recombination, the involved sequences are rich in specific repeats that may permit the overhangs at the two DNA ends to be aligned, extended with polymerase, and ligated by ligase I or III in place of ligase IV (43). The fact that these repeats are not usually used for joining suggests that NHEJ is responsible for the large majority of end joining in class switch recombination.

The most commonly discussed alternative end joining pathway to NHEJ is called microhomology-mediated end joining (MMEJ), which is a somewhat confusing name, given that a subset of NHEJ events also uses 1–4 nt of terminal microhomology (44). MMEJ requires terminal microhomology (whereas microhomology is optional in NHEJ), and the lengths of the microhomology (often >5 nt) are somewhat longer than those used in NHEJ. The ligase for a subset of MMEJ events is ligase I, but that for others is ligase IV in yeast (44). In mammalian cells, there is evidence for ligase III involvement (45), illustrating that MMEJ, and perhaps other alternative routes to joining, may represent merely the possibilities that can arise when the full set of NHEJ components is not present, but that are not normally of substantial physiologic relevance. This seems likely given that the types of pathologic DSBs that arise naturally, such as with ionizing radiation and oxidative damage, would almost never generate the pairs of DNA ends that are required for MMEJ or many other alternative routes, viz. ends that share several nt of terminal microhomology. The ligase III-dependent joining was shown recently to decrease markedly in nondividing cells (46), which is exactly the opposite of what evolution would select for in an end joining pathway. Therefore, in cells mutated for NHEJ, other enzymes may participate, but with a requirement of uncommon extents of and reliance upon microhomology. Rather than calling such events MMEJ, it would be less confusing to call these events ligase IV-independent NHEJ or Ku-independent NHEJ, etc.

Concluding Comments and Future Issues

NHEJ arose early in evolution and is present in many prokaryotes and all eukaryotes. NHEJ is imprecise at the local sequence level, but efficient at restoring chromosomal structural integrity. The enzymatic aspects of NHEJ are becoming clearer, but still include mechanistic uncertainties. At a site of a DSB, it is not yet clear what chromatin modifications are required to permit NHEJ and how these chromatin changes are achieved. Coordination between the completion of the DSB repair and the cell cycle is another area for continuing study. During DNA replication, it is not yet clear what determines whether a given DSB is repaired by NHEJ versus homologous recombination. Therefore, there is still much to be learned about this relatively recently described DNA repair pathway.

Acknowledgments—I thank Drs. K. Schwarz, D. Shibata, and X. Cui for comments on the manuscript.

REFERENCES

1. Lieber, M. R., and Karanjawala, Z. E. (2004) Nat. Rev. Mol. Cell Biol. 5, 69–75
2. Li, H., Vogel, H., Holcomb, V. B., Gu, Y., and Hasty, P. (2007) Mol. Cell. Biol. 27, 8205–8214
3. Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., and Takeda, S. (2006) DNA Repair 5, 1021–1029
4. Moore, J. K., and Haber, J. E. (1996) Mol. Cell. Biol. 16, 2164–2173
5. Adachi, N., Ishino, T., Ishii, Y., Takeda, S., and Koyama, H. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12109–12113
6. Lieber, M. R., Yu, K., and Raghavan, S. C. (2006) DNA Repair 5, 1234–1245
7. West, R. B., Yaneva, M., and Lieber, M. R. (1998) Mol. Cell. Biol. 18, 5908–5920
8. Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. (2002) Cell 108, 781–794
9. Hammarsen, O., DeFazio, L. G., and Chu, G. (2000) J. Biol. Chem. 275, 1541–1550
10. Chu, G. (1997) J. Biol. Chem. 272, 24097–24100
11. Yoo, S., and Dynan, W. S. (1999) Nucleic Acids Res. 27, 4679–4686
12. Meek, K., Gupta, S., Ramsden, D. A., and Lees-Miller, S. P. (2004) Immunol. Rev. 200, 132–141
13. Uematsu, N., Weterings, E., Yano, K., Morotomi-Yano, K., Jakob, B., Taucher-Scholz, G., Mari, P. O., van Gent, D. C., Chen, B. P., and Chen, D. J. (2007) J. Cell Biol. 177, 219–229
14. Jovanovic, M., and Dynan, W. S. (2006) Nucleic Acids Res. 34, 1112–1120
15. Goodarzi, A. A., Yu, Y., Riballo, E., Douglas, P., Walker, S. A., Ye, R., Harer, C., Marchetti, C., Morrice, N., Jeggo, P. A., and Lees-Miller, S. P. (2006) EMBO J. 25, 3880–3889
