Repair of Accidental DNA Double-Strand Breaks in the Human Genome and Its Relevance to Vector DNA Integration

Noritaka Adachi1,2,*, Shinta Saito1 and Aya Kurosawa1

1Graduate School of Nanobioscience, Yokohama City University, Yokohama 236-0027, Japan
2Advanced Medical Research Center, Yokohama City University, Yokohama 236-0004, Japan

HR and NHEJ: Two Major Pathways for Repair of Chromosomal DNA Double-Strand Breaks

Efficient repair of chromosomal DNA damage is crucial for cells to maintain genome integrity. DNA double-strand breaks (DSBs) are the most severe type of DNA lesions that can be caused by various exogenous and endogenous mechanisms, such as ionizing radiation, reactive oxygen species, topoisomerase poisons, or replication errors [1]. DSBs, if left unrepaird or mis-repaired, lead to cell death or chromosomal aberrations [2,3]. Human cells have evolved two fundamentally different mechanisms for repairing chromosomal DSBs, homologous recombination (HR) and non-homologous end-joining (NHEJ) [4]. NHEJ not only repairs accidental (non-physiological) DSBs, but is also essential for repairing physiological DSBs that arise in the process of V(D)J recombination in B and T lymphocytes and class switch recombination in mature B cells [3].

A wide variety of proteins have been identified thus far that contribute to the HR and NHEJ machineries [5]. HR is a highly complicated process of DNA transaction, in which Rad51 protein plays an essential role in DNA strand exchange with the aid of several other proteins such as Rad54, Brca2, Rad52, and Rad51 paralogs [6,7]. For HR to occur, DSBs should be processed (i.e., end-resected) to produce a long 3’-overhang single-stranded DNA [8,9], and recent studies have identified a number of proteins involved in end resection or its regulation; among these, Mre11 and CtIP play essential roles in the initial step of end resection [9-11]. In contrast to HR, NHEJ is thought to be a rather simpler process that requires, at least biochemically, only four proteins (two protein complexes); specifically, Ku, a heterodimer of Ku70 and Ku80, initiates an NHEJ reaction by binding to the ends of a DSB, and the DNA ligase complex composed of Xrc4 and Ligase IV (Lig4) seals the ends to complete repair [3]. In most cases, however, many other proteins do participate in NHEJ-mediated repair to trim the DSB ends, which are typically non-ligatable or non-compatible. These additional NHEJ factors involve DNA-PKcs, Artemis, XLF, and DNA polymerase µ/λ; DNA-PKcs and Artemis have evolved in higher eukaryotes and do not exist in yeasts [3,12]. In addition to the classical pathway of NHEJ, recent evidence indicates the existence of a more error-prone mechanism of NHEJ called alternative end-joining that plays a role in DSB repair [3,13]. Alternative end-joining is Ku/Lig4 independent and the precise mechanism remains largely unclear, although PARP1, Ligase III, and several factors involved in end resection (to initiate HR) have been implicated in DSB repair via alternative end-joining [14-16].

Which DSB repair pathway is beneficial for cells to preserve genome integrity? NHEJ (the classical NHEJ pathway) repairs broken DNA ends with little or no homology and is often associated with nucleotide loss, whereas HR allows for accurate repair of DSBs with the use of homologous DNA sequence, usually located on a sister chromatid [3,4,12]. Such difference in accuracy between the two pathways, however, does not mean that HR is superior to NHEJ in maintaining integrity of human genomes, which contain lots of repetitive DNA sequences [4]. For example, an HR reaction between Alu sequences in a cell would cause deleterious consequences and hence must be prohibited [17,18]. Thus, human somatic cells preferentially use NHEJ to repair accidental DSBs; in particular, in G0/G1 phase of the cell cycle, DSB repair is only performed by NHEJ, and HR is inert. Both NHEJ and HR can work, however, in S to G2 phases when DNA replication has been completed and the sister chromatid is available [19]. Thus, how and which pathway is chosen for repair of a DSB(s) has been a critical issue in the DNA repair field, and there has been a debate [4]. Recent evidence suggests that Ku-bound DSBs, where end resection does not occur, are directed to NHEJ, while end-resected DSBs, to which Ku cannot bind, are channeled to HR (or alternative end-joining) [20-24]. Thus, in addition to the end binding protein Ku, various factors that regulate end resection are involved in DSB repair pathway choice [16,25-30]. Apparently, the type of DSB is also a determinant of pathway choice [31,32]; for example, replication-associated one-ended DSBs are preferentially repaired by HR, while topoisomerase II-mediated DSBs are almost exclusively repaired by NHEJ [33,34]. Interestingly, however, it appears that cells do not always choose a proper pathway to deal with induced DSBs. In fact, absence of NHEJ gives a growth advantage to cells accumulating replication-associated DSBs [34,35], although this may simply reflect the fact that NHEJ is basically the first choice to repair any type of those DSBs that naturally allow Ku-binding [4].

Impact of DSB Repair Deficiency on Targeted and Random Integration

Gene targeting via HR provides the definitive tool in analyzing gene function. For gene targeting to be successfully achieved, the target genome sequence should be replaced with the vector DNA (i.e., targeting vector), not with the sister chromatid. The principal limitation of conventional gene-targeting technology is the extremely low efficiency of HR-mediated targeted integration, which occurs at least 2-3 orders of magnitude less frequently than random integration [36], as depicted in Figure 1A.

Random integration is a phenomenon in which a transfected DNA molecule(s) are inserted into (random sites of) the host genome via non-homologous recombination. It has been generally assumed that random integration results from the repair of spontaneous chromosomal DSBs caused by endogenous factors. Indeed, we have recently shown that DNA topoisomerase IIa and reactive oxygen species (ROS) are such endogenous factors responsible for causing DNA damage that leads to random integration of transfected DNA in human cells [37]. Transient inhibition of topoisomerase IIa significantly increases...
random integration [38]; conversely, siRNA-mediated knockdown of topoisomerase IIα reduces random integration [37]. Cells continuously cultured under 3% oxygen conditions after DNA transfection display reduced random-integration frequency compared to that under 21% oxygen conditions [37], although the gene-targeting efficiency was little affected by the low-oxygen culture condition (our unpublished observations).

Loss of NHEJ in lower eukaryotes results in significantly reduced or no random integration events, and thus, as high as 100% gene-targeting efficiency can be achieved by inactivating NHEJ (for example, [39]). In human somatic cells, however, suppression of NHEJ does not result in decreased random-integration frequency, although the efficiency of gene targeting can be increased [40] (Figure 1A-D). These findings clearly indicate that NHEJ is not the sole mechanism of random integration in human somatic cells, and suggest the contribution of alternative end-joining to the residual random integration events by non-homologous recombination. Intriguingly, unlike vectors with no or shorter homology arms, integration frequency of targeting vectors with long homology arms was not affected by LIG4 deficiency [40] (Figure 1B, C, data not shown). It could be that in the absence of NHEJ, homology arms of the targeting vector served to prevent marker gene loss caused by large deletion (chew-back); however, as these homology arms contain a number of Alu elements, it is more likely that homology arms serve to trigger random integration in an NHEJ-independent fashion. Earlier studies using rodent cell lines, along with the fact that alternative end-joining favors micro-homologies, strongly support this idea [13,41,42].

Elimination of the HR protein Rad54 resulted in significantly reduced gene-targeting efficiency in human cells (Figure 1D), a finding consistent with previous reports using rodent and avian cell mutants [43,44]. Intriguingly, random-integration frequency was more than five-fold higher in RAD54-null cells than in their wild-type counterparts, implying that the observed reduction of gene-targeting efficiency in the absence of Rad54 is due, at least in part, to an unexpectedly increased random-integration frequency. Similar observations were made with mutant cell lines deficient in MUS81 and/or FANCB, genes implicated in HR [45,46] (Figure 1B). It is also important to note that the increased random-integration frequency associated with HR deficiency was suppressed by an additional loss of NHEJ, and this suppression was less pronounced when targeting vectors were used (Figure 1B,C). These
data further support the aforementioned idea that random-integration frequency is substantially influenced by homologous sequences present in the vector, and that these DNA sequences may serve to trigger NHEJ-independent, homology-based random integration. Thus, effective suppression of this mechanism will be a promising approach to reduce random integration events after targeting-vector transfection.

Despite the rapid progress on artificial nucleases (i.e., ZFN, TALEN, or CRISPR-based system) and their effective applications to targeted gene inactivation in various species [47], HR-mediated gene targeting (knock-in as well as knockout) without the use of artificial nucleases still provides an indispensable technique that must be further developed in the context of human-derived cells, as artificial nucleases are capable of causing DNA lesions that lead to deleterious off-target mutations [48-50]. It is expected that deciphering the molecular mechanism of random integration in terms of vector DNA sequence and precise DSB repair mechanisms will help improve human somatic cell gene targeting, for example, by developing a targeting vector that is most suitable for reducing random integrants.

References


