Taylor & Francis
Taylor & Francis Group

AUTHOR'S VIEWS

OPEN ACCESS   Check for updates

Distinct mechanisms of mutagenic processing of alternative DNA structures by repair proteins

Jennifer A. McKinney ⬤*, Guliang Wang ⬤*, and Karen M. Vasquez

Division of Pharmacology and Toxicology, College of Pharmacy, The University of Texas at Austin, Dell Pediatric Research Institute, Austin, TX, USA

ABSTRACT
Repetitive sequences can form a variety of alternative DNA structures (non-B DNA) that can modulate transcription, replication, and repair. However, non-B DNA-forming sequences can also stimulate mutagenesis, and are enriched at mutation hotspots in human cancer genomes. Interestingly, different types of non-B DNA stimulate mutagenesis via distinct repair processing mechanisms.

The etiology of many human diseases, including cancer, involves some form of genetic instability.[1] Distributions of genetic instability "hotspots" in human genomes are not random; however, the mechanisms involved are not fully understood. Repetitive DNA can adopt alternative structures (non-B DNA) that differ from the canonical Watson-Crick B-DNA helix, and are often enriched at hotspots of genetic instability, suggesting that the integrity of the DNA can be influenced by secondary structures. More than a dozen types of non-B DNA structures have been reported, including left-handed DNA (Z-DNA), intramolecular triplex DNA (H-DNA), cruciform/hairpins, G-quadruplexes (G4-DNA), R-loops, etc. [reviewed in[1]], many of which are reported to stimulate genetic instability, and have been implicated in the etiology of various diseases, including neurological disorders and cancer.

Aside from the obvious differences in physical conformation and sequence composition between various non-B DNA structures, each type of structure can lead to specific mutation events that are associated with different types of diseases. For example, triplet repeats (e.g. CAG/CTG, CGG/CCG) that can form loop-out hairpin structures with CG base-pairs and mismatches in the stems, can lead to repeat expansions in neurological disorders.[1] However, hairpin or cruciform structures formed at perfect inverted repeats can stimulate DNA double-strand breaks (DSBs), resulting in large-scale deletions and translocations in mammalian cells.[2] G4-DNA is a four-stranded structure that is stabilized by uninterrupted stacking of at least three G-quartets, where each is stabilized by Hoogsteen-hydrogen bonding of the four guanines in a coplanar G-quartet arrangement. The guanines can be provided from tracks of G runs on the same or different DNA strands.[1] Z-DNA structures result in a zig-zag appearance of the backbone at alternating purine and pyrimidine regions such as GC and GT repeats, or mixtures of GC and GT dinucleotides that maintain an alternating purine and pyrimidine

pattern.[1] H-DNA can form at polypurine-polypyrimidine elements with mirror repeat symmetry.[1] One of the strands from the tract can be separated from its complementary strand and bind to the purine strand in the underlying duplex via Hoogsteen-hydrogen bonding through the major groove to form a three-stranded DNA structure.[1] Interestingly, G4-DNA, H-DNA, and Z-DNA have all been shown to stimulate DSBs, leading to deletions, translocations, and recombination events.[1]

Although these structures can be mutagenic in various *in vitro* and *in vivo* model systems (including bacteria, yeast, mice, and human cells), their mutagenic potential can differ from one species to another. For example, while both H-DNA and Z-DNA can stimulate DSBs in mammalian cells, leading to deletions and additional downstream mutagenic events,[3,4] H-DNA-forming sequences have not been shown to be mutagenic in bacterial cells.[3] In contrast, Z-DNA-forming GC repeats, similar to other simple repeats, can cause repeat unit expansion or contraction in bacteria.[4] In addition, while CAG repeats predominantly result in expansions in human disease, the same repeats have been shown to cause more deletions than expansions in bacteria, yeast, and mice.[1,5]

The genetic instability events induced by non-B DNA are, in part, initiated by proteins from various DNA repair pathways that can recognize and process the structures in an effort to repair and remove the "damage" in the absence of actual DNA damage per se. Interestingly, the processing of non-B DNA appears to be specific to the structure and species. For example, we found that H-DNA-induced mutagenesis required functional nucleotide excision repair (NER) in yeast and human cells, yet H-DNA was not mutagenic in bacteria. In addition, there are replication-dependent and replication-independent mechanisms involved in processing H-DNA, where the replication protein, flap endonuclease 1 (FEN1), is involved in suppressing H-DNA-induced mutagenesis in a replication-dependent
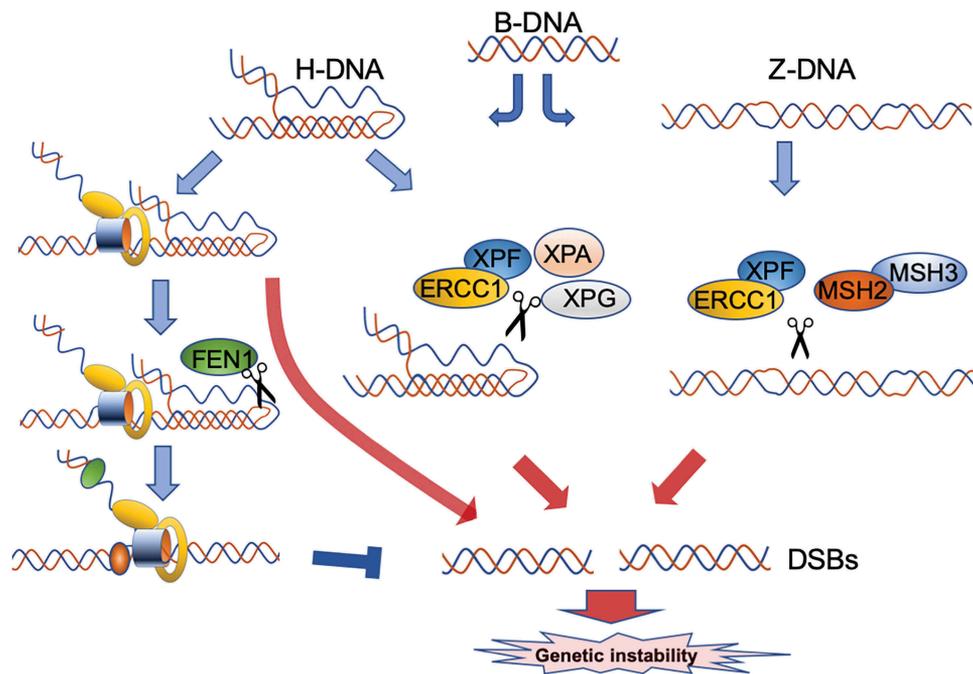
Figure 1. Structure-specific processing of non-B DNA conformations that impact genetic instability in mammalian cells. Under appropriate conditions certain repetitive sequences can adopt non-B DNA conformations such as intramolecular triplex DNA (H-DNA) and left-handed DNA (Z-DNA). (Left): DNA replication protein, flap endonuclease 1 (FEN1) attenuates H-DNA-induced mutagenesis by cleaving and removing the H-DNA structure during replication; (Middle): nucleotide excision repair (NER) proteins including xeroderma pigmentosum complementation groups A and G (XPA and XPG, respectively), and the complex consisting of excision repair cross-complemetation group 1 and xeroderma pigmentosum complementation group F (ERCC1-XPF) recognize and cleave H-DNA leading to DNA double-strand breaks (DSBs) and genetic instability, regardless of the replication status of the cell; (Right): the ERCC1-XPF and MutS homolog 2 and MutS homolog 3 (MSH2-MSH3) complexes, acting outside of their canonical roles in NER and mismatch repair (MMR), recognize and cleave Z-DNA structures resulting in DNA double-strand breaks (DSBs) and subsequent genetic instability. All proteins are depicted as ovals with their abbreviations listed.

manner. However, the NER-associated protein complex consisting of excision repair cross-complemetation group 1 and xeroderma pigmentosum complementation group F (ERCC1-XPF) and xeroderma pigmentosum complementation group G (XPG) nucleases can stimulate H-DNA-induced mutations in a replication-independent manner.[6] In contrast, we found that Z-DNA-induced genetic instability can occur via a different mechanism than that of H-DNA. The mutagenic processing of Z-DNA does not require functional NER, but rather it involves cooperation between the NER nuclease complex, ERCC1-XPF, and the mismatch repair (MMR) complex, MutS homolog 2 and MutS homolog 3 (MSH2-MSH3). Other components of these pathways are not required for the mutagenic processing of Z-DNA, suggesting that ERCC1-XPF and MSH2-MSH3 are acting outside of their canonical roles in NER and MMR.[7] The model proposed in Figure 1 summarizes the structure-specific processing of H-DNA and Z-DNA. Similar non-canonical physical and functional interactions between the ERCC1-XPF and MSH2-MSH3 complexes have been previously demonstrated in yeast and mammalian cells[8] in processing recombination intermediates[9] and DNA interstrand crosslinks.[10]

The mechanisms proposed for the mutagenic processing of non-B DNA structures provide a possible explanation for the enrichment of non-B DNA-forming sequences at translocation hotspots in human cancer genomes.[2,6] While we and others have identified a number of proteins/pathways that are involved in the processing of various non-B DNA structures, it is likely that other components are involved as well, and further studies are warranted to fully understand these

mechanisms. A thorough understanding of the mechanisms involved in non-B DNA-induced genetic instability specific to each structure, as well as the steps involved in alternative DNA structure formation, stability, and recognition will provide essential insight into the etiology of human disease, and allow for therapeutic advances to better treat and/or prevent genetic instability-associated disorders.

Disclosure of potential conflicts of interest

The authors report no conflict of interest.

ORCID

Jennifer A. McKinney ⬦ http://orcid.org/0000-0002-9312-4883
Guliang Wang ⬦ http://orcid.org/0000-0001-6871-7540

References

1. Wang G, Vasquez KM. Effects of replication and transcription on DNA structure-related genetic instability. Genes (Basel). 2017;8 (1):17. doi:10.3390/genes8010017.
2. Lu S, Wang G, Bacolla A, Zhao J, Spitser S, Vasquez KM. Short inverted repeats are hotspots for genetic instability: relevance to cancer genomes. Cell Rep. 2015;10(10):1674–1680. doi: 10.1016/j.celrep.2015.02.039.

3. Wang G, Vasquez KM. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(37):13448–13453. doi:10.1073/pnas.0405116101.

4. Wang G, Christensen LA, Vasquez KM. Z-DNA-forming sequences generate large-scale deletions in mammalian cells. Proc Natl Acad Sci U S A. 2006;103(8):2677–2682. doi:10.1073/pnas.0511084103.

5. Jankowski C, Nasar F, Nag DK. Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. Proc Natl Acad Sci U S A. 2000;97(5):2134–2139. doi:10.1073/pnas.040460297.

6. Zhao J, Wang G, Del Mundo IM, McKinney JA, Lu X, Bacolla A, Boulware SB, Zhang C, Zhang H, Ren P, et al. Distinct mechanisms of nuclease-directed DNA-structure-induced genetic instability in cancer genomes. Cell Rep. 2018;22(5):1200–1210. doi:10.1016/j.celrep.2018.01.014.

7. McKinney JA, Wang G, Mukherjee A, Christensen L, Subramanian SHS, Zhao J, Vasquez KM. Distinct DNA repair pathways cause genomic instability at alternative DNA structures. Nat Commun. 2020;11(1):236. doi:10.1038/s41467-019-13878-9.

8. Lan L, Hayashi T, Rabeya RM, Nakajima S, Kanno SI, Takao M, Matsunaga T, Yoshino M, Ichikawa M, Te Riele H, et al. Functional and physical interactions between ERCC1 and MSH2 complexes for resistance to cis-diamminedichloroplatinum(II) in mammalian cells. DNA Repair (Amst). 2004;3(2):135–143. doi:10.1016/j.dnarep.2003.10.005.

9. Saparbaev M, Prakash L, Prakash S. Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in Saccharomyces cerevisiae. Genetics. 1996;142:727–736.

10. Zhao J, Jain A, Iyer RR, Modrich PL, Vasquez KM. Mismatch repair and nucleotide excision repair proteins cooperate in the recognition of DNA interstrand crosslinks. Nucleic Acids Res. 2009;37:4420–4429. doi:10.1093/nar/gkp399.