Published in final edited form as:

Mutat Res. 2013 ; 0: 118–131. doi:10.1016/j.mrfmmm.2012.11.005.

The Yin and Yang of Repair Mechanisms in DNA Structure-induced Genetic Instability

Karen M. Vasquez* and Guliang Wang

Division of Pharmacology and Toxicology, College of Pharmacy, The University of Texas at Austin – Dell Pediatric Research Institute, 1400 Barbara Jordan Blvd. R1800, Austin, TX 78723

Abstract

DNA can adopt a variety of secondary structures that deviate from the canonical Watson-Crick B-DNA form. More than 10 types of non-canonical or non-B DNA secondary structures have been characterized, and the sequences that have the capacity to adopt such structures are very abundant in the human genome. Non-B DNA structures have been implicated in many important biological processes and can serve as sources of genetic instability, implicating them in disease and evolution. Non-B DNA conformations interact with a wide variety of proteins involved in replication, transcription, DNA repair, and chromatin architectural regulation. In this review, we will focus on the interactions of DNA repair proteins with non-B DNA and their roles in genetic instability, as the proteins and DNA involved in such interactions may represent plausible targets for selective therapeutic intervention.

Repetitive sequences that do not code for proteins account for more than half of the total DNA in the human genome, which includes transposon-derived repeats, (e.g., Alu and LINE elements), pseudo-genes or duplications, and tandem simple repeats [1]. In contrast to its former nickname, “junk DNA”, repetitive sequences are now known to play important roles in regulating the genome, including putative involvement in shaping chromatin structure, regulating gene expression, and stimulating genomic rearrangement. Of particular interest, some simple repeats, which represent approximately 3% of the total human genome, have been found to form different types of non-canonical structures that differ from the Watson-Crick B-form (i.e., non-B DNA structures). Extensively studied microsatellite repeats, such as triplet repeats can contain tens, hundreds, or even thousands of trinucleotide repeat units [2, 3]. Expanded triplet repeats (such as CTG and CCG) have been implicated in more than 20 genetic diseases, including Fragile X syndrome, Huntington's disease, Friedreich's ataxia, and myotonic dystrophy [4, 5]. The role of microsatellites in disease is not limited to triplet repeats, as evidenced by Myotonic dystrophy type 2 (DM2), which is caused by the extreme expansion of CCTG tetranucleotide repeats from <30 repeats in normal individuals to up to 11,000 in some patients [6], and spinocerebellar ataxia type 10 patients, in which the number of pentanucleotide ATTCT repeats can be expanded from approximately 14 in normal individuals to up to 4,500 [7].

© 2012 Elsevier B.V. All rights reserved

*Address correspondence to: Karen M. Vasquez, PhD, Division of Pharmacology and Toxicology, College of Pharmacy, The University of Texas at Austin – Dell Pediatric Research Institute, 1400 Barbara Jordan Blvd. R1800, Austin, TX 78723. Tel.: +1 512 495 3040. Fax: +1 512 496 4946. karen.vasquez@austin.utexas.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest: There are no competing interests to declare.

In cells, the majority of genomic DNA exists in the B-conformation at any given time (Figure 1A), but the conformational state of DNA is dynamic and is generally negatively supercoiled corresponding to a high-energy state. The DNA is subjected to manipulation via replication, transcription, and DNA repair proteins. These processes unwind the DNA from the histones, generating more negative supercoiling stress and can open the double helix, giving rise to non-B structures at sequences (*e.g.*, repetitive sequences) conducive to their formation. To date, more than 10 different types of non-B DNA conformations have been identified and characterized [8]. Examples include, slipped structures formed at direct repeats, hairpin or cruciform structures formed at inverted repeats, intra-molecular triplex DNA structures (H-DNA) formed at polypurine-polypyrimidine elements with mirror repeat symmetry, four-stranded G-quartet structures formed at sequences comprising 4 guanine tracks of at least 3 continuous guanosine residues separated by 1–7 nucleotides, and left-handed Z-DNA structures formed at alternating pyrimidine-purine sequences (see Figure 1; for review, see [9]). Notably, under certain conditions, imperfect repeat sequences can also form non-B conformations. For example, short CG repeats (2–4 repeats) that are separated by 3-bp interruptions can adopt Z-DNA structures and form Z-Z junctions on plasmids *in vitro*, as assessed by two-dimensional gel electrophoresis and chemical and enzymatic probing [10]. Hairpin structure formation typically requires an inverted repeat symmetry, but can also form at CNG triplet repeats, which contain a mismatch in every two C-G [11, 12]. Thus, sequences with the potential to form non-B DNA structures (or non-B DNA-forming sequences) are very abundant in the human genome.

1. Non-B DNA structure and genetic instability

Neurological disorders related to triplet repeat expansion have been studied extensively (for reviews see [13, 14]); however, many types of non-B DNA conformations in addition to loop structures due to slippage and misalignment (*e.g.*, slipped DNA) of repetitive units or hairpins/cruciform formed at triplet repeat sequences have been implicated in genetic instability associated with a variety of human disorders [15][16–18], as we briefly outline below.

1.1. Intra-molecular triplex DNA (H-DNA) and genetic instability

H-DNA can form at polypurine-polypyrimidine sequences with mirror repeat symmetry, where half of the symmetry, when single-stranded, binds to the purine-rich strand of the duplex containing the other half of the symmetry *via* Hoogsteen hydrogen bonding through the major groove, forming a three-stranded DNA structure (Figure 1B) [19]. Genes containing long polypurine-polypyrimidine sequences show higher frequencies of alternative splicing and chromosomal translocations [20], and H-DNA-forming sequences frequently co-localize with chromosomal breakpoint hotspots and genomic rearrangements in disease-related human genes. For example, several polypurine-polypyrimidine tracks can be found near (within hundreds of bps) the major breakage hotspots in the *c-MYC* gene found in leukemias and lymphomas [21–28], and in the *BCL-2* gene major breakpoint region (Mbr) in follicular lymphomas. *In vitro* experiments demonstrated that these sequences were capable of forming non-B DNA structures, including H-DNA [29]. Furthermore, disruption of the H-DNA conformation by disruption of the mirror symmetry reduced the frequency of translocation events in the *BCL-2* Mbr [30]. The 21st intron of the human *PKD1* gene, which contains a 2.5 kb polypyrimidine tract with 23 mirror repeats [31] that can adopt H-DNA structures *in vitro* [32], has been implicated in the high mutation rate in this region in both germ line and somatic cells from autosomal dominant polycystic kidney disease patients (ADPKD) [33]. The 2.5-kb polypyrimidine sequence in a plasmid induced DNA double-strand breaks (DSBs) and resulted in large-scale deletions in *E. coli* [34]. In our previous work, we discovered that both an endogenous H-DNA-forming sequence from the human *c-MYC* promoter and model H-DNA-forming sequences, when

cloned into a mutation reporter shuttle plasmid, induced DSBs near the H-DNA locus and induced mutation frequencies ~20-fold above spontaneous background levels in mammalian COS-7 cells. Greater than 80% of the mutants exhibited large-scale deletions and/or rearrangements [35]. Moreover, when this H-DNA-forming sequence from the human *c-MYC* promoter was integrated into chromosomes of mutation-reporter mice, genetic instability was detected in ~8% of the offspring [36], indicating that H-DNA structures can be mutagenic in a chromosomal context *in vivo*.

1.2. Left-handed Z-DNA and genetic instability

Purine bases can adopt *syn* conformations in sequences of alternating purines and pyrimidines (*e.g.*, GT, GC), which can facilitate the formation of a left-handed helix with a zigzag-shape backbone (*i.e.*, Z-DNA, Figure 1C) that differs from the canonical Watson-Crick B-DNA helix [37, 38]. Among the alternating purine-pyrimidine sequences, GC repeats have the greatest capacity to form Z-DNA structures, followed by GT repeats, while AT repeats are more likely to form hairpin or cruciform structures rather than Z-DNA [39]. Although long GC repeats are rare in mammalian genomes, GT repeats are the most frequent simple repeating unit in the human genome, accounting for ~0.25% of the total DNA [40]. In addition to pure alternating purine-pyrimidine sequences, purine-pyrimidine repeats with interruptions have been shown to adopt Z-DNA *in vitro* under certain conditions [41, 42].

Algorithms designed to identify such sequences in the human genome have revealed co-localization of Z-DNA-forming sequences and genetic instability hotspots, particularly chromosome breakage and translocation sites in genes involved in human diseases such as cancer. For example, breakage hotspots in the *c-MYC* gene, *e.g.*, the P1 promoter and 3' downstream breakpoint region, contain multiple mixed GT and GC repeats capable of forming Z-DNA [43, 44]. Translocation breakpoints also map within hundreds of bps surrounding Z-DNA-forming sequences in lymphoid tumors [45]. Sequencing analysis of a 1.7 kb DNA breakage hotspot cluster region from a subgroup of B-cell precursor acute lymphoblastic leukemia (BCP-ALL) patients identified a region of Z-DNA-forming GT repeats to be involved in amplification and rearrangement events [46]. This study also implicated a chi-like heptamer, CCTCAGC, in the rearrangements, suggesting that both DNA secondary structures and recombinogenic mechanisms may contribute to BCP-ALL etiology [46].

Plasmids carrying Z-DNA-forming sequences are difficult to maintain in *E. coli*, suggesting that they are unstable in bacteria. For example, CG repeats, which can form stable Z-DNA structures, induce contractions of repeat units in bacteria, likely generated by misalignment during replication [4, 47–49]. Other simple repetitive sequences such as hairpin-forming triplet repeats have been implicated in replication-induced slippage events. When the mutagenic potential of a hairpin-forming AT(12) repeat was compared to that of a CG(12) repeat, which may form a hairpin or Z-DNA, it was found that the CG(12) repeat was more mutagenic than the AT(12) repeat [48], implicating Z-DNA, rather than a hairpin structure in CG repeat-induced mutation in bacteria. Moreover, in mammalian cells model Z-DNA-forming CG(14) repeats induced DSBs within a 300-bp region surrounding the Z-DNA-forming sequences and induced mutation frequencies ~20-fold above spontaneous background levels, largely in the form of deletion and/or rearrangement events [50]. These same Z-DNA-forming repeats resulted in even greater levels of genetic instability on chromosomes of transgenic mutation-reporter mice with ~7% of the population undergoing large deletion and/or rearrangement events in the vicinity of (within 300 bp) the non-B region [36]. On the contrary, this same sequence resulted exclusively in contractions or expansions of repeat units in bacteria [50]. Thus, there are at least two distinct mechanisms of mutagenesis stimulated by Z-DNA-forming sequences in various organisms.

1.3. Hairpins, cruciforms, and G4 DNA in genetic instability

Inverted repeat sequences (or palindromes) can adopt hairpin conformations *via* intra-strand base pairing, where half of the single-stranded symmetry folds back and forms intra-strand Watson-Crick base pairs with the other half on the same strand (Figure 1D) [51, 52]. Palindromes as short as 14 bp are sufficient for formation of a stable hairpins *in vivo* [53]. Long inverted repeats are capable of forming large and complex structures. For example, a long inverted repeat on the human X chromosome was reported to form a large double-cruciform structure involving multiple Holliday junctions which allowed for both intra-strand base pairing within the stems and interstrand base pairing of the spacer region bringing the hairpin stems together [54]. Inverted repeats are involved in generating deletions, gene amplification, and translocations, and can serve as substrates for recombination in bacteria [55], yeast [56, 57], and mammalian cells [54, 58–61]. The centers of the AT-rich long inverted repeats on human chromosome 11q23 and 22q11 co-localize with the breakpoints for t(11;22) translocations [62–64], and references therein]. Imperfect inverted repeat sequences, such as triplet repeats, can also adopt slippage loops or hairpin/cruciform structures (Figure 1E) with mismatches in the stems [13–15, 65], which can undergo expansion and/or deletion events *in vivo*, with a bias toward expansions in humans and deletions in bacteria, yeast, and mice [4, 5, 66, 67]. Further, long inverted repeats, and presumably the hairpin/cruciform structures formed, represent preferred integration sites for avian sarcoma virus and human immunodeficiency virus-1, at least *in vitro* [68]. The viruses tend to integrate within the stems of cruciform structures on plasmid DNA, adjacent to the loops in the cruciforms [68]. Long inverted repeats are also associated with chromosomal translocations in the human genome [69]. Taken together, these results provide strong evidence that cruciform structures formed at long inverted repeat sequences contribute to genetic instability and evolution.

G-quadruplexes or G4 DNA, four-stranded Hoogsteen-hydrogen bonded structures formed at G-rich sequences (Figure 1F), have been the subject of many studies related to the function of alternatively structured DNA (for reviews see [70–72]). G-quadruplex-forming sequences are found in telomeres, and are enriched in promoters and rDNA, and are even more abundant in mitochondrial DNA than in nuclear DNA in yeast [73, 74]. G4 DNA motifs were found to be more conserved than expected by chance [74], consistent with the suggested biological roles of G-quadruplexes in transcription, replication, and telomere maintenance [70]. G-quadruplex-forming sequences have also been identified in regions within 500 bp of mitotic and meiotic DSB sites [74, 75], and at hotspots of genetic instability in human disease [70]. For example, G-quadruplex structures can interfere with DNA replication leading to stalled and/or collapsed replication forks, which can result in DNA breakage and genetic instability [76–78]. It has also been demonstrated that when G4-forming sequences serve as the template for leading strand replication, G4 structure formation can result in deletions and complex rearrangement events in the absence of the G-quadruplex unwinding helicase Pif1 [78, 79]. The G-quadruplex-induced rearrangements are dependent on the homologous recombination proteins Rad51–Rad52 [79, 80]. Moreover, G4 DNA plays an important role in telomeric repeat stability [81]; on the one hand the single-stranded G-rich telomeric repeats are capable of forming G-quadruplex structures, which may protect the DNA ends from degradation, on the other hand, G-quadruplex structures must be resolved to allow for efficient replication of telomeres [81].

2. DNA repair-related proteins that associate with non-B DNA structures

The mutagenic potential of non-B DNA-forming sequences has been demonstrated in many species, implicating non-B DNA structures in disease etiology and genetic evolution [82]. It is thought that the non-B DNA structure itself, and not the sequence (in B-form DNA) is involved in this instability. Thus, structure formation and stability are key regulators of the

subsequent mutagenesis, as are the enzymes that process these structures. A variety of proteins can recognize and bind to non-B DNA structures to stabilize or resolve the conformation, recruit other factors to the structures, or directly process the non-B DNA regions. A family of winged-helix proteins with a general mechanism of Z-DNA recognition has been described suggesting the presence of a common “Z-DNA-binding motif” [83]. For example, the Z-alpha motif of the adenosine deaminase acting on RNA 1 (ADAR1) protein binds with high affinity to Z-DNA [84], and facilitates a B-DNA to Z-DNA structural transition even at the short TA(3) repeat, which is unable to adopt Z-DNA in the absence of ADAR1 [84]. Five 14-3-3 family isoforms (beta, gamma, epsilon, zeta, and sigma), which have been implicated in signal transduction and cell cycle control, can bind to cruciform structures [85–87]. Many architectural and regulatory proteins, such as histones H1 and H5, HMG proteins, and p53 preferentially bind to cruciform structures, or can facilitate the formation of cruciform structures upon binding (for review see [88]). In addition to cruciforms, some of these architectural proteins can bind other non-B DNA structures such as triplex DNA [89].

Non-B DNA structures may provide signals for structure-specific binding proteins involved in DNA damage recognition and repair mechanisms. Nucleotide excision repair (NER) proteins are thought to recognize DNA damage *via* chemical and physical alterations in the Watson-Crick B-DNA helix. It has been reported that the efficiency of lesion recognition and processing by NER is dependent on the extent of distortion induced in the DNA helix [90–92]; *i.e.*, lesions that induce greater distortions to the DNA helix are repaired more efficiently than less distorting lesions. Because non-B DNA deviates from the canonical B-DNA helix, these structures may be recognized as “damage” by the NER damage/distortion recognition proteins. This may also apply to other DNA repair proteins, as many non-B DNA-structure binding and/or processing proteins are involved in DNA repair, such as NER proteins, mismatch repair (MMR) proteins, helicases, and replication-related proteins, as discussed in detail below. Many other proteins are involved in processing the intermediates of non-B DNA-induced mutations, such as DSB repair proteins; here, we will focus on the enzymes that directly process non-B DNA structures, and not their intermediates, as the DSB repair pathways are beyond the scope of this review.

2.1. Nucleotide excision repair proteins

NER removes bulky DNA adducts by dual incision on either side of the damaged base(s) and removes the damaged nucleotide(s) together with the adjacent region. NER is a critical pathway for maintaining the integrity of the genome and is highly conserved in prokaryotes and eukaryotes [93, 94]. Because non-B DNA conformations result in helical distortions it is not surprising that some types of non-B structures can be recognized and processed by NER factors.

Inter-molecular triplexes, formed when a triplex-forming oligonucleotide (TFO) binds duplex DNA *via* Hoogsteen hydrogen bonding [95], have been shown to induce NER-dependent mutagenesis in mammalian cells [96]. In HeLa nuclear extracts, the NER recognition/verification proteins, xeroderma pigmentosum group A (XPA), and group C (XPC) were reported to be involved in processing inter-molecular triplexes [97]. Consistent with these results, *in vitro* binding studies have confirmed that the XPA-RPA complex [98] and the XPC-RAD23B complex [99] recognize covalent inter-molecular triplexes with high affinity and specificity, and can interact with each other on these structures [99]. The bacterial NER equivalent, the UvrABC endonuclease, has the capacity to process inter-molecular triplex structures *in vitro* [100], suggesting that triplex structures are substrates for NER across species.

Studies on the role of NER in hairpin/cruciform structure-induced genetic instability have largely focused on CNG triplet repeats. The *E. coli* NER protein UvrA can bind CAG repeats on a supercoiled plasmid that supports a hairpin conformation at a much higher affinity than to linear CAG repeats, and the deletion of CAG repeats was reduced in a UvrA-deficient *E. coli* strain [101]. Deficiencies in the endonuclease (UvrC) of the UvrABC complex or the UvrD helicase resulted in an enhancement of deletion events at CAG repeats [102], while deficiency of UvrB, which cleaves phosphodiester bonds downstream of the damaged sites, reduced the deletion events on the CAG tracts [102, 103]. In another study, the absence of UvrA in bacteria increased the deletion events at CAG repeats on plasmid DNA [103]. The authors speculated that UvrA could bind to the bubble and loop regions formed at the repeats [104, 105], and thereby prevent bypass synthesis and deletion of the repeats [103]. However, the stability of CAG repeats integrated in the bacterial genome was not affected by UvrA deficiency [102], while XPA depletion in human cells significantly reduced the transcription-induced deletion events on CAG repeats [106]. In *Drosophila*, CAG repeats exhibited length-dependent instability with a bias toward expansions [107]. Transcription through the repeats dramatically increased both contraction and expansion events, and deficiency of Mus201 (an ortholog of human XPG) suppressed repeat instability, suggesting an important role of transcription-coupled NER (TC-NER) in CAG repeat stability [107]. This was also the case in human cells where transcription through CAG repeats led to repeat contraction [106], and siRNA depletion of CSB, a protein required for TC-NER, and the NER endonucleases ERCC1 and XPG, suppressed CAG repeat contractions [108]. Together, these studies reveal complex effects of NER proteins on the stability of hairpin-forming triplet repeats. The role of NER in processing cruciform structures formed at perfect-inverted repeats is still not clear. Further work is required to better define the roles of NER proteins in triplet repeat instability in various organisms.

2.2. Mismatch repair proteins

Cells deficient in MMR are hypermutable and predisposed to tumorigenesis as evidenced by hereditary non-polyposis colon cancer in humans [109, 110]. MMR is responsible for the correction of base-pairing errors such as mismatches and insertion-deletion loops caused by spontaneous and induced base deamination, oxidation, methylation, reactive oxygen species, recombination intermediates, and replication errors. In this capacity, MMR proteins interact with components of other repair pathways including NER to prevent spontaneous mutation and microsatellite instability [111]. Single-stranded looped regions are found in many types of non-B DNA conformations, such as hairpins, slipped structures [82], and B-Z or Z-Z junctions [112, 113], which may represent substrates for MMR. For example, MMR has been shown to be required for removing a small cruciform structure (12 bp) in Chinese hamster ovary cells [114] and in yeast [53]; yet neither MSH2 nor MLH are required for repairing large loops (*e.g.*, 140 bp) in mammalian cells [57, 115, 116]. The MSH2–MSH3 complex binds to hairpin structures containing mismatches with higher affinities than hairpin structures formed at perfect inverted repeats [117]. So although MMR is efficient in removing small loops within hairpin or cruciform structures, it does not appear to process longer cruciforms that contain perfect base-pairing in the stem region. This is consistent with the well-known function of MMR in post-replicative repair of mismatches and small loops [118]. While MMR deficiency results in instability at microsatellite sequences in humans, the absence of the Msh2 in yeast does not significantly affect the large deletions (>12 repeat units) at CAG(50) repeats [119] or large expansions at CAG(25) (>13 repeat units) [120]. Interestingly, in mouse models of Huntington's disease or myotonic dystrophy (carrying long CNG repeats), deficiencies of Msh2 [121], Msh3 [122], or Pms2 [123] reduced repeat expansion events, or shifted the mutation spectra from expansions to deletions [124]. However, in human cells, MSH2 and MSH3 were responsible for the contraction events at CAG(95) repeats [106]. The reasons for these differences are not clear,

but suggest that MMR proteins are not the only proteins that can process hairpin structures formed at CNG repeats, thus subtle differences in intracellular environment may alter the outcomes of CNG repeat processing. The MSH2–MSH6 heterodimer can bind to G-rich repetitive sequences in activated immunoglobulin switch regions in B-cells, which can form G-quadruplexes and RNA/DNA hybrids during transcription [125, 126]. Further, MSH2 or MSH6 deficiency in B-cells can reduce switch recombination and heterogeneity of switch junctions [76–78]. MMR proteins have also been implicated in the recognition and processing of TFO-directed inter-molecular triplexes [127–129]. MSH2–MSH3 binds inter-molecular triplex structures with high affinity and specificity *in vitro* and *in vivo*, and MSH2–MSH3 can interact with XPA-RPA or XPC-RAD23B in recognizing these structures [129]. In *E. coli*, MutS and MutL are required for the large deletions induced by H-DNA-forming GAA repeats from the human *PKD1* gene [34]. Thus, MMR proteins clearly play a role in non-B DNA structure-induced genetic instability in a number of species.

2.3. Base excision repair proteins

Base excision repair (BER) removes lesions formed by alkylating agents and oxidative DNA damage, including non-bulky base modifications and abasic sites. Defects in BER have been associated with hereditary neurodegenerative diseases, immunodeficiency, cancer, and aging (For a review, see [130]).

BER proteins have been implicated in triplet repeat instability; for example, cells deficient in DNA polymerase beta and the BER cofactor, HMGB1, did not support CAG expansions [131]. DNA polymerase beta was found to be enriched at expanded CAG repeats in the striatum of mice representing a model for Huntington's disease (HD), where the tissue shows high somatic CAG instability, but not in the cerebellum where the CAG repeats are more stable [132]. On the contrary, the activities of Flap endonuclease-1 (Fen-1), Hmgb1, and Ligase 1 were significantly lower in the striatum of these mice [132, 133]. The authors speculated that a lower Fen-1/polymerase beta ratio in the striatum could lead to an accumulation of 5'-flap structures at CAG repeats generated by polymerase beta strand displacement during long patch BER of oxidative DNA lesions. If not repaired, the 5'-flap-containing CAG repeats can form hairpin structures and lead to expansion events [132]. Fen-1 can cleave flap structures and suppress hairpin-induced repeat expansion events [134], and the loss of Rad27 (Fen-1 homolog in yeast) in *S. cerevisiae* increased CNG repeat expansions and deletions several hundred-fold above background levels [135, 136]. In contrast, stable hairpins formed at longer repeats (*e.g.*, >20 CTG/CAG repeats) were resistant to the activity of Fen-1 [137, 138]. In fact, when the activity of Fen-1 was uncoupled from polymerase beta, it was found to promote expansions by facilitating ligation of the slipped repeats in the hairpins to the single-stranded DNA in the gap, resulting in expansions during the next cycle of replication [131].

8-oxoguanine glycosylase (OGG1) is a DNA glycosylase that is involved in the processing of 7,8-dihydro-8-oxoguanine (8-oxoG) by BER, a common lesion generated by exposure to reactive oxygen species. It was found that repair processing by OGG-1 can stimulate non-B DNA formation and result in repeat expansion in mice [139, 140], as described in the “toxic oxidation cycle” [141]. This model will be discussed in detail in section 3.3. “DNA damage and repair processes can alter DNA structure” below and in [142]. Interestingly, Ogg1 deficiency in yeast mitochondria resulted in an ~six-fold increase in GT repeat deletions as compared to the wild-type, and overexpression of Ogg1 completely suppressed GT repeat deletion events, suggesting that the activity of Ogg-1 could reduce repeat instability mitochondria [143]. In human cells, depletion of OGG-1 by siRNA to 30% of the endogenous levels did not affect the transcription-related CAG repeat instability [108]. Although the differences among these reports are not clear, it appears that Ogg-1 is involved in oxidative damage-related instabilities within repetitive sequences.

2.4. DNA helicases

DNA helicases are critical for unwinding duplex DNA, which is required for DNA metabolic processes including replication, transcription, and repair. Many DNA and RNA helicases have been characterized and it is estimated that as much as 1% of known open reading frames in the human genome code for predicted DNA and RNA helicases [144]. Helicases play a key role in genome maintenance, such that deficiencies in some helicases can result in premature aging and/or increased susceptibility to cancer [145]. For example, deficiencies in BLM, WRN, or RECQL4 are associated with human Bloom's, Werner's, and Rothmund-Thomson syndromes (RTS), respectively. Helicases play important roles in DNA repair and interact with many DNA repair proteins such as BRCA1 [146] and Topo III [145, 147, 148]. Some helicases are more efficient at resolving non-B DNA conformations than B-DNA duplexes [149], and see below], suggesting that they may protect the genome from mutagenic non-B DNA structures.

Non-B DNA structures can stall replicative DNA polymerases resulting in replication fork collapse and subsequent DNA breakage and deletions. Helicases are actively involved in preventing replication stalling, and can assist in replication fork restart, in part by resolving non-B DNA structures [144]. The RecQ helicase family, a superfamily 1 helicase group, has gained increasing attention in studies related to their activity on non-B DNA structures. *E. coli* and some lower eukaryotic species encode only one *RecQ* helicase (e.g., *RecQ* in *E. coli* and *Sgs1* in budding yeast and *S. cerevisiae*), while higher organisms contain multiple *RECQ* family members. RecQ helicases bind ssDNA or dsDNA with 3'-overhangs better than blunt substrates or those with 5'-overhangs, and unwind DNA in the 3' to 5' direction [150].

In BLM-deficient mice, cruciform-forming inverted repeats were found at deletion junctions, while deletions in BLM-proficient cells were associated with direct repeats, suggesting that the BLM helicase plays an important role in preventing cruciform-induced genetic instability [151]. WRN-deficient mice experienced an 8-fold increase in mutations at a long inverted repeat sequence integrated at a Line-1 element [152], demonstrating the importance of helicases in maintaining genome stability at non-B DNA structure-forming sequences. RecQ helicases bind Holliday junction (HJ) structures, which are conformationally similar to cruciform structures, with higher affinities than duplex DNA, and promote branch migration in an ATP-dependent manner [153, 154]. In fact, it has been reported that mammalian RECQ helicases, such as WRN, BLM [155–158], and RECQL1 [159] have the capacity to unwind HJ over thousands of bps.

The yeast *Sgs1* helicase, in a complex with Top3 and Rmi1, resolves aberrant recombination intermediates [160], disassembles branched DNA structures, and plays an integral role in meiotic recombination [161]. *Sgs1*-deficient yeast cells accumulate cruciform-shaped replication intermediates at damaged forks [162], and these unprocessed structures can be resolved by overexpression of the *E. coli* *RusA* or human GEN1 HJ resolvases [163], or the endogenous Mus81-Mms4 endonuclease [164].

Purified bacterial RecQ protein has been reported to melt triplex structures and facilitate primer extension by T7 DNA polymerase through polypurine-polypyrimidine tracts [165]. Purified BLM and WRN helicases can also unwind the third strand from TFO-induced inter-molecular triplex containing a 3' overhang [166]. An inter-molecular triplex containing a 3' tail might mimic the structural features of a stalled replication fork. WRN was shown to resolve H-DNA and Z-DNA structures *in vitro*, however, depletion of WRN protein in human cells resulted in increased mutation frequencies in H-DNA-forming, Z-DNA-forming, and B-DNA-forming sequences, suggesting that the increased genetic instability was independent of DNA structure [167]. Hairpin and G-quadruplex structures formed at

expanded CGG repeats can stall the eukaryotic replicative DNA polymerases alpha, delta, and epsilon; and the WRN helicase can interact with DNA polymerase delta to facilitate DNA synthesis through the CGG repeats [168].

The RQC domain, a conserved G-quadruplex DNA binding domain located at the C-terminus of the RecQ family helicases, is responsible for high affinity interactions of helicases with G-quadruplex structures [169]. Human telomeres contain long 3' single-stranded TTAGGG repeats and can adopt non-B DNA conformations including G-quadruplexes [170]. In fact, the first crystal structure of a G-quadruplex was obtained using human telomeric repeat sequences [171]. The 3'-end provides a substrate for RecQ helicase loading, and RecQ family helicases are important for maintaining G-rich sequences including the telomeres and immunoglobulin switch regions [145, 147, 169, 172].

In addition to the RecQ family helicases, the FANCI helicase can unwind non-B DNA conformations such as G-quadruplexes [173] and is involved in maintaining genomic stability. FANCI, also known as BRIP1 or BACH1, is a member of the superfamily 2 helicases. FANCI was originally identified through its interaction with BRCA1 and its role in DNA repair and breast cancer suppression [146]. Fanconi anemia patients deficient in FANCI are hypersensitive to DNA crosslinking agents [174], and consistent with this phenotype, FANCI has been implicated in crosslink repair via its interaction with MLH1 [175]. Different from the RECQ helicases, FANCI requires a 5' ssDNA tail for loading and it unwinds duplex DNA substrates in a 5' to 3' direction with respect to the helicase-loaded strand [176].

The *C. elegans Dog-1* (deletion of guanine-rich regions) gene, the *FANCI* homolog in worms, is required for suppressing deletion events that occur at G-rich regions capable of forming G-quadruplexes [177]. In *Dog-1*-deficient cells, mutation frequencies at G-rich repeats correlated with the tract lengths and G-richness, *i.e.*, the ability to form a G-quadruplexes, suggesting a critical role of *Dog-1* in suppressing G-quadruplex-induced genetic instability [178]. Moreover, *Dog-1* and *Him-6* (a *BLM* homolog) double mutants exhibited higher instability at G-quadruplex-forming regions than in *Dog-1* mutants [178], suggesting that there are multiple helicases involved in maintaining the stability of G-quadruplex-forming regions in worms. Similar to the functions of *Dog-1* in *C. elegans*, human FANCI can also resolve G-quadruplex structures *in vitro* in an ATP-dependent manner, and FANCI-deficient human cells have deletions at G-quadruplex-containing regions [173].

Another 5'–3' helicase that plays an important role in resolving non-B DNA is the Pif1 protein, a member of helicase superfamily 1, which is found in most eukaryotes [179]. Pif1 helicase is involved in the maintenance of mitochondrial DNA and genomic DNA, and has a particularly important role in telomeric DNA replication in eukaryotic species from yeast to human [179]. In yeast, Pif1 deficiency can result in replication stalling and DSBs at G-quadruplex regions [180]. Consistent with this finding, Pif1 was found to be enriched at G-quadruplex-forming sequences *in vivo* [78]. Human PIF1 also binds to a G-quadruplexes *in vitro* with higher affinity than to ssDNA or duplex DNA [181]. Pif1 can unwind G-quadruplex DNA and this reaction requires a 5' ssDNA tail [181], similar to the FANCI helicase.

DHX9 belongs to the helicase superfamily 2, and is also known as nuclear DNA helicase II (NDH II) or RNA helicase A (RHA) and is essential for efficient transcription and DNA repair. DHX9 unwinds DNA-DNA, RNA-RNA or DNA-RNA duplexes in a 3' to 5' direction, although it unwinds RNA hybrids more efficiently than the corresponding DNA-DNA substrates [182]. DHX9 has been shown to interact with WRN and co-localizes with

γ -H2AX at centrosomes in human cells [183]. Following DNA damage, the pre-assembled DHX9- γ -H2AX complex disassociated from the centrosome and was recruited to sites of damaged DNA to assist in their repair [183]. Interestingly, DHX9 helicase co-precipitates with inter-molecular triplex structures in mammalian cells [184], and can resolve these structures *in vitro* with greater efficiency than that of forked structures. The activity of DHX9 on triplex structures was found to be dependent on ATP hydrolysis and a 3'-free overhang on the third strand TFO [184]. DHX9 has the capacity to unwind intra-molecular H-DNA and Z-DNA structures on plasmids in human cells [167] and is enriched at these structure-forming sequences in cells (Jain *et al.*, unpublished results), suggesting that either a 3' tail is generated during the processing of non-B DNA or that a 3' tail is not necessary for resolving these structures by DHX9 *in vivo*. The generation of a 3' tail may lead to structure resolution independent of DHX9 because a nick or break would lead to the loss of negative supercoiling. Another possibility is that DHX9 is involved in processing the DSBs induced by non-B DNA as DHX9 has been shown to interact with non-homologous end-joining proteins [185]. In contrast to results from studies using WRN-depleted cells [167], DHX9 depletion resulted in an H-DNA-specific increase in mutagenesis in human cells, suggesting differences in structure-specific functions of helicases (Jain *et al.*, unpublished results). DHX9 helicase has also been shown to unwind RNA and DNA G-quadruplexes *in vitro* [182], implicating this helicase in the maintenance of genome stability at regions of alternatively structured DNA.

There are many other helicases that are capable of unwinding non-B DNA conformations, such as the SV40 large T antigen, which has been shown to unwind the third strand of triplex structures [186], and RTEL1 that acts on G4 DNA to assist in maintaining telomere integrity [187]. Another superfamily 1 helicase member, the ATP-dependent Srs2 DNA helicase can function with DNA polymerase delta *via* interactions with PCNA to unwind hairpin structures formed at CGG repeats (but not G4T4 repeats that form G-quadruplex structures), and facilitate replication through the repeats, thus preventing chromosome fragility in yeast [188]. Compared to wild-type cells, Srs2 deficient yeast cells showed an ~33-fold increase in CTG repeats expansions and contractions, ~4-fold more chromosomal breakage on yeast artificial chromosomes [189], and up to a 40-fold increase in expansion rates at CTG, CAG, and CGG repeats on episomal plasmids [190]. As summarized above, in most cases, helicase activity facilitates the unwinding of the mutagenic non-B DNA conformation and thus prevents DNA breakage and mutations at those sites. In addition to this direct effect, helicases are also involved in maintaining epigenetic stability during replication, which is thought to occur via coupling of replication with histone recycling. Sarkies *et al.* (2010) found that in REV-1-deficient cells, G4 DNA can inhibit processive replication, which uncouples DNA synthesis from histone recycling, resulting in loss of repressive chromatin marks [191]. Thus, in addition to utilizing the RECQ helicase to unwind non-B DNA structures and restart replication, cells can resume replication via REV1-dependent translesion synthesis (see below), and FANCD1 appears to coordinate these pathways at G-quadruplex motifs [191].

In contrast, helicases do not always function to maintain genomic stability. For example, deficiency of yeast Sgs1 did not affect hairpin-forming CAG repeat expansions [192], while depletion of Sgs1 reduced the expansion of H-DNA-forming GAA repeats [193]. This is very similar to the roles of DNA topoisomerases on non-B DNA-induced genetic instability; Topo I and II can relax the negative supercoiling on duplex DNA [194], and thereby reduce the formation of non-B DNA and subsequent genetic instability, while the breakage resulting from topoisomerase cleavage on non-B DNA structures [195, 196] can initiate genetic instability. In summary, the effects of helicases on non-B DNA metabolism appear to be structure-specific and are more complicated than simply resolving mutagenic structures to maintain genome stability.

2.5. DNA/RNA polymerases

Repetitive DNA sequences can present a challenge to DNA and RNA polymerases during replication or transcription, resulting in stalling of the polymerases, which can ultimately lead to genetic instability at these sites. For example, T7 RNA polymerase was partially blocked within and downstream of the H-DNA-forming sequence from the human *c-MYC* promoter region, and within a CG(14) Z-DNA-forming sequence on negatively supercoiled plasmids. It has been speculated that non-B DNA-induced RNA polymerase pausing may stimulate the recruitment and activation of transcription-coupled NER factors at the non-B DNA region [197, 198]. Replicative DNA polymerases delta and epsilon generated small insertions or deletions within CA repeats at frequencies that were 1000-fold higher than in non-repetitive sequences [199]. Replication of single-stranded GGC repeats by DNA polymerases alpha, beta, HIV reverse transcriptase, Taq DNA polymerase, or the Klenow fragment of DNA polymerase I caused expansions of repeat units likely due to misalignment of the repetitive nascent strands and the templates [200]. Some translesion synthesis (TLS) DNA polymerases can replace the replicative polymerases delta and epsilon at forks stalled by non-B DNA structures, and can synthesize through the non-B regions, thus suppressing the replication fork stalling and DNA breakage. For example, depletion of the TLS polymerases eta or kappa in HeLa cells engineered to contain multiple tandem copies of genomic non-B DNA-forming sequences resulted in more DSBs compared to wild-type cells, implicating these TLS polymerases in efficient DNA synthesis through these structure-forming sequences [201]. Rev1 is a member of the Y-family DNA polymerases and functions with polymerase zeta (a B-family DNA polymerase) in the error-prone processing of UV-induced DNA damage [202]. Yeast *rev1* mutants were found to have higher rates of both expansions and deletions of CTG repeats, while *rev1* deficiency had no impact on the stability of triplet repeats unable to adopt non-B DNA structures [203]. Interestingly, this effect on CTG triplet repeat stability did not require polymerase zeta, an essential Rev1 partner in translesion synthesis of UV-damaged templates, suggesting that Rev1 might interact with other proteins on non-B DNA structure-forming sequences [203].

In summary, it is important to note that the activities of repair-related proteins on non-B DNA structures may represent a double-edged sword in the context of DNA structure-induced genetic stability. Formation of a non-B DNA structure, *per se*, does not directly lead to breakage and instability, but can impair DNA metabolic processes, *e.g.*, blocking RNA or DNA polymerase progression. On the other hand, some secondary structures are required for efficient transcription, replication, and repair, demonstrating the “yin and yang” of repair and other DNA metabolic processes on DNA structure-induced genetic instability. Helicases can remove the non-B DNA-related structural obstacles to allow for efficient DNA metabolism and thereby reduce non-B DNA-induced genetic instability. Alternatively, the processing and removal of non-B DNA structures by repair proteins (or other enzymes) could lead to DNA nicks or DSBs, giving rise to an increased risk of genetic instability.

3. DNA damage/repair processes and non-B DNA structures

Common features of non-B DNA structures include single-stranded regions and alterations in the Watson-Crick double helix, such that non-B DNA may be hypersusceptible to DNA damage from endogenous and exogenous sources. Moreover, the contributions of repair protein binding to the formation or stabilization of non-B DNA structures *in vivo* are not clearly understood. Thus, the crosstalk between DNA repair proteins and non-B DNA may not be limited to the processing of the structured DNA by the repair proteins. Because DNA damage and repair processes, and the formation of non-B DNA structures can alter the landscape of the affected DNA in chromosomes, they can influence each other in several ways as we outline below.

3.1. Effect of non-B DNA on DNA damage patterns

DNA damage and mutation are not random processes, but are influenced by many factors, including sequence context, structural features of the DNA, and higher-order chromatin organization. For example, major groove interactions are altered within a triplex structure because the third strand occupies the major groove and forms Hoogsteen hydrogen bonds with the purine-rich strand of the underlying duplex. Moreover, the yields of radiation-induced DNA damage in triplex structures are different from those in duplex B-DNA structures [204], and triplex-duplex junctions are hyperreactive to a variety of agents, including DNA intercalators [205, 206]. Studies to determine the influence of DNA structure on radiation-induced strand breakage demonstrated that the local structure (e.g., a narrow minor groove that gives rise to a low accessibility of H4' and H5'2 and a low probability of H-atom abstraction by OH radicals) affected the likelihood of breakage events [207]. Single-stranded DNA regions are generally more accessible to damaging agents than duplex DNA because the bases are not protected inside the sugar-phosphate backbone. As mentioned above, exposure of single-stranded regions is a common feature of most non-B DNA conformations, and thus they may be preferred targets for DNA damage.

Base extrusion is a general feature of B-DNA to Z-DNA junctions and Z-Z junctions [112, 113, 208, 209]. B-DNA to Z-DNA junctions are hyperreactive to DNA interactive agents such as 4-NQO [210], hydroxylamine, and osmium tetroxide [211]. Guanine bases within Z-DNA are in a *syn* position and are located over the sugar without protection, and are therefore more sensitive to dimethylsulfate and diethylsulfate modification [211, 212]. In Z-DNA the guanines are more sensitive, and the cytosines less sensitive, to ionizing radiolytic attack than in B-DNA [213].

A guanine in the loop of the hairpin formed at a (CAG)₁₀ repeat was shown to be hypersensitive to modification by peroxynitrite when compared to a guanine in a CAG(10) duplex, and hOGG1 removed 8-oxoG from the loop of hairpin structure at an efficiency that was 700-fold lower than on duplex DNA [214]. G-quadruplex motifs are targets for oxidative damage during hypoxia-induced transcription, resulting in the recruitment of BER enzymes and the formation of transient strand breaks [215]. Ndlbe & Schuster (2006) demonstrated that while charge migration from radical cation reactions with guanines occurred vertically along a DNA chain, it occurred horizontally within a G-quartet [216]. In addition, oxidizing radicals can become trapped within folded quadruplex structures and attack the external tetrads of the quadruplexes with higher efficiency than the center tetrads [217]. These examples emphasize the roles of DNA secondary structures on DNA damage patterns *in vitro*, suggesting that non-B DNA structure-forming sequences should be considered when analyzing DNA damage patterns in cells.

3.2. Effect of non-B DNA conformations on DNA repair

Genomic DNA is constantly under attack by both endogenous and exogenous DNA damaging agents. Because cell survival depends on genomic stability and the ability to repair DNA damage, exquisitely sensitive DNA repair mechanisms have evolved to process DNA damage. Defects in these repair systems can lead to severe disorders such as xeroderma pigmentosum, Cockayne syndrome, ataxia telangiectasia, Fanconi anemia, Bloom syndrome, and trichothiodystrophy, and an enhanced predisposition to cancer [218]. The consequences of genomic instability are also manifest in several hereditary human neurological disorders that involve the expansion of triplet repeat sequences [14].

The formation of non-B DNA structures introduces helical distortions into the DNA and alters Watson-Crick B-DNA helical parameters, which can influence DNA repair processes. It has been shown that bubble and loop regions in duplex DNA can serve as entry sites for

the *E. coli* NER complex UvrAB, prior to its translocation on the DNA [104, 105]. And the human NER nucleases, ERCC1-XPF and XPG are structure specific in that they incise DNA at pre-initiation bubble structures containing bulky adducts. As discussed above, DNA damage/distortion recognition proteins can bind non-B DNA, but the extent to which non-B DNA (in the absence of DNA damage *per se*) is subsequently processed by repair proteins is not well understood. The accessibility of repair proteins to damaged sites within non-B DNA regions may be altered compared to those in B-DNA. For example, the removal of N7-methylguanine by DNA glycosylase was shown to be less efficient in Z-DNA than in B-DNA [219, 220], and the repair efficiency of O6-methylguanine in a poly(dG-m5dC).poly(dG-m5dC) sequence in the ZDNA conformation was only 10% of that measured in B-DNA [221]. The hOGG1 protein showed reduced affinity to 8-oxoG when it was located within a hairpin structure, compared to B-DNA, and the repair efficiency was significantly lower [141]. In the striatal tissue of HD mice where CAG repeats are unstable, abasic sites located near the 5' ends of the CAG repeats were repaired with lower efficiency than other locations, such as the 3' ends of the repeats, or within non-repetitive sequences [133]. Further, the interactions of repair proteins with non-B DNA structures may result in altered activity and fidelity of repair. For example, the ATPase activity of the MSH2-MSH3 complex was inhibited, and the MSH2-MSH3-ADP/ATP affinity was altered upon binding to hairpin structures formed at CAG repeats [117, 222].

3.3. DNA damage and repair processes can alter DNA structure

Lesions in the DNA can change the energetics of conformational transitions, which may facilitate or disfavor the formation of non-B DNA structures. For example, psoralen photocrosslinking has been used to map DNA secondary structures (*e.g.*, hairpins and loops) of single-stranded genomes [223], due to its ability to covalently crosslink duplex regions of DNA. Replacement of a guanine with an abasic site in a B-DNA substrate destabilized the duplex, while an abasic site located within the loop of a CAG repeat hairpin structure did not substantially alter the thermodynamic stability of the looped region, and eventually increased the population of looped structures [224]. On the other hand, one can speculate that an interstrand crosslink could inhibit the formation of local secondary DNA structures in double-stranded genomes due to the requirement for strand separation in the generation of most non-B DNA conformations.

Non-B DNA structures exist in high-energy states, and their formation is favored under conditions that generate negative supercoiling and destabilize the B-DNA structure (*e.g.*, DNA replication, transcription, repair). While genomic DNA is largely in the B-form at any given time, certain regions with the capacity to adopt non-B DNA dynamically fluctuate between B-DNA and non-B conformations. In the human genome, DNA is organized and compacted into chromosomes, and is wrapped around histone cores. Interaction of the positively charged arginines on the histone core particle surface with the negatively charged DNA favors a B-DNA conformation, where the arginines match the shape of the minor groove of B-DNA [225]. The left-handed wrapping renders the negative supercoiling torsional stress anonymous, thus favoring a B-DNA conformation in chromatin. DNA repair is coupled to chromatin remodeling to unwrap the DNA from the nucleosome for repair processing, and thereby generates negative supercoiling that can destabilize the duplex structure. The resulting single-stranded DNA is susceptible to conformational transitions, such that DNA repair processing may facilitate the formation of non-B DNA. For example, more expansions occur at CAG repeats when BER processes lesions that are located near the 5' end of the repeats compared to the 3' end of the repeats [133]. We have described this process as “repair stimulated DNA structure formation” model [142].

4. Non-B DNA conformations affect chromosomal DNA stability at a distance

Recently, our group and others have found that non-B DNA conformations can induce mutations (including point mutations, small deletions, or insertions) hundreds of base pairs from the non-B DNA site, while the non-B DNA sequence itself may or may not contain mutations [167, 193]. The locations and spectra of the mutations suggested that they were not directly initiated at the non-B DNA region itself, although they were dependent on the non-B DNA-forming sequences. The mechanisms involved are not known, however, one plausible explanation is that the formation of a non-B DNA structure alters the chromatin structure and/or histone modifications in a large region of adjacent DNA, rendering it more susceptible to damage and mutation. In support of this notion, Ruan & Wang (2008) found that GAA repeats were refractory to nucleosome assembly *in vitro* when in a triplex structure relative to the same repeats in a duplex conformation [226]. Inter-molecular triplexes have been found to interrupt histone-DNA contacts flanking the triplex structure, and can serve as nucleosome barriers [227]. Triplexes have been reported to form on nucleosome-bound poly(dA).poly(dT) tracts, but they typically form at the ends of nucleosomal DNA fragments [228] with lower stability [229], accompanied by changes in nucleosome structure [230]. Thus, nucleosome formation and triplex formation are thought to be mutually inhibitory processes with the tendency for triplex exclusion from the nucleosome [231].

Despite the inherent CG-richness of the nucleosome core, the Z-DNA-forming CG(9) repeat was shown to be excluded from nucleosomes in *S. cerevisiae* [232], and methylation of CpG repeats capable of Z-DNA formation in the promoter region of the chicken adult beta-globin gene, prevented DNA from binding to the histone octamer *in vitro* [233]. In contrast, cytosine methylation of CpG repeats unable to adopt Z-DNA did not affect nucleosome positioning, suggesting that Z-DNA formation was required for nucleosome exclusion of methylated CpG repeats [234]. We have recently discovered that cytosine methylation can induce/stabilize Z-DNA structures at sequences that do not adopt Z-DNA conformations when unmethylated *in vitro*. And when methylated ZDNA-forming sequences were introduced into mammalian cells, they were excluded from nucleosome assembly *in vivo* (Wang *et al.*, unpublished data). Interestingly, a short unmethylated CGG repeat that represented an unfavorable site for histone octamer binding converted to a strong nucleosome positioning sequence upon methylation [235]. However, the expanded CGG repeat that was capable of forming a hairpin or G-quadruplex structure was resistant to nucleosome assembly regardless of the methylation status [235], suggesting a direct relationship between DNA secondary structure and nucleosome assembly.

A computer-based analysis of the association between nucleosomes and putative G-quadruplex-forming sequences revealed that they were coincident with nucleosome depleted regions in both human cells and in *C. elegans* beyond that expected by chance alone [236]. It is not known whether G-quadruplex formation displaces nucleosomes or whether these structure-forming sequences evolved to occupy nucleosome-free regions, but in either case, these findings suggest that these sequences serve a functional role in the genome.

The central AT-rich region of the Ars insulator sequence, a boundary element that can block enhancer-promoter communications, can adopt non-B DNA conformations as evidenced by its sensitivity to diethylpyrocarbonate modification [237]. Moreover, this central AT-rich region was found to be excluded from nucleosomes in DNA isolated from sea urchin embryos [237], implicating non-B DNA in altering nucleosome formation.

Alternatively, some triplet repeat sequences that can adopt hairpin structures have been found to be strong nucleosome positioning elements. Long CTG repeats (75 or 130 repeats) formed stable nucleosomes with calf thymus histone octamers, with nearly 10-fold greater propensity than the 5S RNA gene, a strong natural nucleosome-positioning element [238, 239]. CTG repeats at the 3' noncoding region of the *DMPK* gene at the Myotonic dystrophy (DM1) locus were found to be centrally located within nucleosomal regions [240]. Moreover, the sequences adjacent to the expanded CTG repeats (over a thousand repeats) from DM1 patients were less accessible for nuclease digestion, suggesting that they were protected in the nucleosome structure [241]. Thus, some non-B DNA conformations may represent strong nucleosome assembly elements and facilitate nucleosome remodeling at adjacent regions.

DNA conformation can also affect epigenetic modifications, such as cytosine methylation and histone acetylation. For example, cytosines within Z-DNA structures are known to be resistant to methyltransferases, resulting in substantially lower levels of methylation compared to the same sequences in B-form DNA [220, 242]. Using a bioinformatics approach, Bock *et al.* (2006) found that certain genomic characteristics such as DNA sequence and structure, and specific DNA repeats, were highly correlated with CpG island methylation on Chromosome 21 in human lymphocytes. Hence, cytosine methylation in CpG islands tends to occur at the regions that are predicted to form non-B DNA structures [243]. It was also reported that an H-DNA-forming AG repeat increased the level of histone acetylation in the smooth muscle myosin light chain kinase gene promoter [244]. It is not clear that the increased acetylation was caused by a direct interaction between the H-DNA structure and histones (or with histone acetyltransferases or histone deacetylases) or if it was an indirect affect of H-DNA-induced chromatin remodeling.

5. Summary

Non-B DNA conformations interact with a wide variety of DNA processing proteins in cells. The complex interplay of DNA structure and repair can affect each other in both positive and negative ways, demonstrating the “yin and yang” of DNA repair mechanisms in DNA structure-induced genetic instability. For example, the formation of alternatively structured DNA can alter the patterns of DNA damage and the efficiency of repair within that region, while DNA damage and repair can either facilitate or prohibit the formation of non-B DNA conformations. Further, some DNA processing proteins can suppress non-B DNA structure-induced genetic instability, while other proteins in the same pathway might enhance DNA structure-induced mutagenesis. A simplified model of the interactions between DNA damage, repair, and non-B DNA structure-induced mutagenesis is summarized in Figure 2. Based on our current understanding, it is reasonable to postulate that non-B DNA-forming regions may induce genetic instability in part by modulation of epigenetic factors, resulting in inefficient nucleosome assembly surrounding the non-B DNA regions, leaving the DNA more susceptible to damage and mutagenesis. Many questions remain unanswered, for example, what is the range (*i.e.*, how far upstream and/or downstream from a non-B DNA-forming sequence) of influence of non-B DNA on nucleosome assembly or genetic instability? How do non-B DNA-induced histone or DNA modifications influence genetic instability? What proteins recognize and process non-B DNA structures? What is the susceptibility of different non-B DNA-forming sequences to damage induced by various DNA damaging agents, and how efficiently are they repaired? Clearly, more studies are warranted to gain better mechanistic insights into the role of non-B DNA in chromosomal organization and genetic instability both locally and at a distance.

Acknowledgments

We thank Dr. Rick A. Finch for critical review of the manuscript. This work was supported by an NIH/NCI grant to K.M.V. (CA093729).

Literature Cited

- [1]. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001; 409:860–921. [PubMed: 11237011]
- [2]. Guida M, Marger RS, Papp AC, Snyder PJ, Sedra MS, Kissel JT, Mendell JR, Prior TW. A molecular protocol for diagnosing myotonic dystrophy. *Clin Chem*. 1995; 41:69–72. [PubMed: 7813083]
- [3]. Cheng S, Barcelo JM, Korneluk RG. Characterization of large CTG repeat expansions in myotonic dystrophy alleles using PCR. *Hum Mutat*. 1996; 7:304–310. [PubMed: 8723679]
- [4]. Wells RD. Molecular basis of genetic instability of triplet repeats. *J. Biol. Chem*. 1996; 271:2875–2878. [PubMed: 8621672]
- [5]. Bowater RP, Wells RD. The intrinsically unstable life of DNA triplet repeats associated with human hereditary disorders. *Prog. Nucleic Acid Res. Mol. Biol*. 2001; 66:159–202. [PubMed: 11051764]
- [6]. Dere R, Wells RD. DM2 CCTG*CAGG repeats are crossover hotspots that are more prone to expansions than the DM1 CTG*CAG repeats in *Escherichia coli*. *J. Mol. Biol*. 2006; 360:21–36. [PubMed: 16753177]
- [7]. Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T. Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat. Genet*. 2000; 26:191–194. [PubMed: 11017075]
- [8]. Choi J, Majima T. Conformational changes of non-B DNA. *Chem. Soc. Rev*. 2011; 40:5893–5909. [PubMed: 21901191]
- [9]. Sinden, RR. DNA structure and function. Academic Press; San Diego: 1994.
- [10]. McLean MJ, Lee JW, Wells RD. Characteristics of Z-DNA helices formed by imperfect (purine-pyrimidine) sequences in plasmids. *J. Biol. Chem*. 1988; 263:7378–7385. [PubMed: 3284886]
- [11]. Owen BA, Yang Z, Lai M, Gajec M, Badger JD 2nd, Hayes JJ, Edelmann W, Kucherlapati R, Wilson TM, McMurray CT. (CAG)(n)-hairpin DNA binds to Msh2–Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol*. 2005; 12:663–670. [PubMed: 16025128]
- [12]. Petruska J, Arnheim N, Goodman MF. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res*. 1996; 24:1992–1998. [PubMed: 8668527]
- [13]. Di Prospero NA, Fischbeck KH. Therapeutics development for triplet repeat expansion diseases. *Nat. Rev. Genet*. 2005; 6:756–765. [PubMed: 16205715]
- [14]. Caskey CT, Pizzuti A, Fu YH, Fenwick RG Jr. Nelson DL. Triplet repeat mutations in human disease. *Science*. 1992; 256:784–789. [PubMed: 1589758]
- [15]. Pearson CE, Zorbas H, Price GB, Zannis-Hadjopoulos M. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J. Cell. Biochem*. 1996; 63:1–22. [PubMed: 8891900]
- [16]. Bacolla A, Wojciechowska M, Kosmider B, Larson JE, Wells RD. The involvement of non-B DNA structures in gross chromosomal rearrangements. *DNA Repair (Amst)*. 2006; 5:1161–1170. [PubMed: 16807140]
- [17]. Wells RD. Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci*. 2007; 32:271–278. [PubMed: 17493823]
- [18]. Zhao J, Bacolla A, Wang G, Vasquez KM. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci*. 2010; 67:43–62. [PubMed: 19727556]
- [19]. Lyamichev VI, Mirkin SM, Frank-Kamenetskii MD. Structures of homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dyn*. 1986; 3:667–669. [PubMed: 3271043]

- [20]. Bacolla A, Collins JR, Gold B, Chuzhanova N, Yi M, Stephens RM, Stefanov S, Olsh A, Jakupciak JP, Dean M, Lempicki RA, Cooper DN, Wells RD. Long homopurine*homopyrimidine sequences are characteristic of genes expressed in brain and the pseudoautosomal region. *Nucleic Acids Res.* 2006; 34:2663–2675. [PubMed: 16714445]
- [21]. Wiener F, Ohno S, Babonits M, Sumegi J, Wirschubsky Z, Klein G, Mushinski JF, Potter M. Hemizygous interstitial deletion of chromosome 15 (band D) in three translocation-negative murine plasmacytomas. *Proc. Natl. Acad. Sci. U. S. A.* 1984; 81:1159–1163. [PubMed: 6422467]
- [22]. Akasaka T, Akasaka H, Ueda C, Yonetani N, Maesako Y, Shimizu A, Yamabe H, Fukuhara S, Uchiyama T, Ohno H. Molecular and clinical features of non-Burkitt's, diffuse large-cell lymphoma of B-cell type associated with the c-MYC/immunoglobulin heavy-chain fusion gene. *J. Clin. Oncol.* 2000; 18:510–518. [PubMed: 10653866]
- [23]. Kovalchuk AL, Muller JR, Janz S. Deletional remodeling of c-myc-deregulating chromosomal translocations. *Oncogene.* 1997; 15:2369–2377. [PubMed: 9393881]
- [24]. Joos S, Haluska FG, Falk MH, Henglein B, Hameister H, Croce CM, Bornkamm GW. Mapping chromosomal breakpoints of Burkitt's t(8;14) translocations far upstream of c-myc. *Cancer Res.* 1992; 52:6547–6552. [PubMed: 1330296]
- [25]. Haluska FG, Tsujimoto Y, Croce CM. The t(8;14) breakpoint of the EW 36 undifferentiated lymphoma cell line lies 5' of MYC in a region prone to involvement in endemic Burkitt's lymphomas. *Nucleic Acids Res.* 1988; 16:2077–2085. [PubMed: 3128772]
- [26]. Saglio G, Borrello M, Grazia, Guerrasio A, Sozzi G, Serra A, di Celle PF, Foa R, Ferrarini M, Roncella S, Pignatti C, Borgna, et al. Preferential clustering of chromosomal breakpoints in Burkitt's lymphomas and L3 type acute lymphoblastic leukemias with a t(8;14) translocation. *Genes Chromosomes Cancer.* 1993; 8:1–7. [PubMed: 7691153]
- [27]. Care A, Cianetti L, Giampaolo A, Sposi NM, Zappavigna V, Mavilio F, Alimena G, Amadori S, Mandelli F, Peschle C. Translocation of c-myc into the immunoglobulin heavy-chain locus in human acute B-cell leukemia. A molecular analysis. *EMBO J.* 1986; 5:905–911. [PubMed: 3013623]
- [28]. Wilda M, Busch K, Klose I, Keller T, Woessmann W, Kreuder J, Harbott J, Borkhardt A. Level of MYC overexpression in pediatric Burkitt's lymphoma is strongly dependent on genomic breakpoint location within the MYC locus. *Genes Chromosomes Cancer.* 2004; 41:178–182. [PubMed: 15287031]
- [29]. Kinniburgh AJ. A cis-acting transcription element of the c-myc gene can assume an H-DNA conformation. *Nucleic Acids Res.* 1989; 17:7771–7778. [PubMed: 2678005]
- [30]. Raghavan SC, Chastain P, Lee JS, Hegde BG, Houston S, Langen R, Hsieh CL, Haworth IS, Lieber MR. Evidence for a triplex DNA conformation at the bcl-2 major breakpoint region of the t(14;18) translocation. *J. Biol. Chem.* 2005; 280:22749–22760. [PubMed: 15840562]
- [31]. Van Raay TJ, Burn TC, Connors TD, Petry LR, Germino GG, Klinger KW, Landes GM. A 2.5 kb polypyrimidine tract in the PKD1 gene contains at least 23 H-DNA-forming sequences. *Microb. Comp. Genomics.* 1996; 1:317–327. [PubMed: 9689215]
- [32]. Blaszak RT, Potaman V, Sinden RR, Bissler JJ. DNA structural transitions within the PKD1 gene. *Nucleic Acids Res.* 1999; 27:2610–2617. [PubMed: 10373576]
- [33]. Watnick TJ, Piontek KB, Cordal TM, Weber H, Gandolph MA, Qian F, Lens XM, Neumann HP, Germino GG. An unusual pattern of mutation in the duplicated portion of PKD1 is revealed by use of a novel strategy for mutation detection. *Hum. Mol. Genet.* 1997; 6:1473–1481. [PubMed: 9285784]
- [34]. Bacolla A, Jaworski A, Larson JE, Jakupciak JP, Chuzhanova N, Abeyasinghe SS, O'Connell CD, Cooper DN, Wells RD. Breakpoints of gross deletions coincide with non-B DNA conformations. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:14162–14167. [PubMed: 15377784]
- [35]. Wang G, Vasquez KM. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:13448–13453. [PubMed: 15342911]
- [36]. Wang G, Carbajal S, Vijg J, DiGiovanni J, Vasquez KM. DNA structure-induced genomic instability in vivo. *J. Natl. Cancer Inst.* 2008; 100:1815–1817. [PubMed: 19066276]

- [37]. Malfoy B, Rousseau N, Vogt N, Viegas-Pequignot E, Dutrillaux B, Leng M. Nucleotide sequence of an heterochromatic segment recognized by the antibodies to Z-DNA in fixed metaphase chromosomes. *Nucleic Acids Res.* 1986; 14:3197–3214. [PubMed: 3010230]
- [38]. Johnston BH. Generation and detection of Z-DNA. *Methods Enzymol.* 1992; 211:127–158. [PubMed: 1406305]
- [39]. McLean MJ, Wells RD. The role of DNA sequence in the formation of Z-DNA versus cruciforms in plasmids. *J. Biol. Chem.* 1988; 263:7370–7377. [PubMed: 2835374]
- [40]. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001; 409:860–921. [PubMed: 11237011]
- [41]. Feigon J, Wang AH, van der Marel GA, van Boom JH, Rich A. Z-DNA forms without an alternating purine-pyrimidine sequence in solution. *Science.* 1985; 230:82–84. [PubMed: 4035359]
- [42]. Eichman BF, Schroth GP, Basham BE, Ho PS. The intrinsic structure and stability of out-of-alternation base pairs in Z-DNA. *Nucleic Acids Res.* 1999; 27:543–550. [PubMed: 9862978]
- [43]. Rimokh R, Rouault JP, Wahbi K, Gadoux M, Lafage M, Archimbaud E, Charrin C, Gentilhomme O, Germain D, Samarut J, et al. A chromosome 12 coding region is juxtaposed to the MYC protooncogene locus in a t(8;12)(q24;q22) translocation in a case of B-cell chronic lymphocytic leukemia. *Genes Chromosomes Cancer.* 1991; 3:24–36. [PubMed: 2069907]
- [44]. Wolfl S, Wittig B, Rich A. Identification of transcriptionally induced Z-DNA segments in the human c-myc gene. *Biochim. Biophys. Acta.* 1995; 1264:294–302. [PubMed: 8547317]
- [45]. Boehm T, Mengle-Gaw L, Kees UR, Spurr N, Lavenir I, Forster A, Rabbitts TH. Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. *EMBO J.* 1989; 8:2621–2631. [PubMed: 2531086]
- [46]. Sinclair PB, Parker H, An Q, Rand V, Ensor H, Harrison CJ, Strefford JC. Analysis of a breakpoint cluster reveals insight into the mechanism of intrachromosomal amplification in a lymphoid malignancy. *Hum. Mol. Genet.* 2011
- [47]. Casanovas JM, Ellison MJ, Rodriguez-Campos A, Azorin F. The obtention of simian virus 40 recombinants carrying d(CG.GC)n, d(CA.GT)n and d(CT.GA)n sequences. Stability of the inserted simple repeating sequences. *Eur. J. Biochem.* 1987; 167:489–492. [PubMed: 2820728]
- [48]. Freund AM, Bichara M, Fuchs RP. Z-DNA-forming sequences are spontaneous deletion hot spots. *Proc. Natl. Acad. Sci. U. S. A.* 1989; 86:7465–7469. [PubMed: 2552445]
- [49]. Sinden RR. Neurodegenerative diseases. Origins of instability. *Nature.* 2001; 411:757–758. [PubMed: 11459042]
- [50]. 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; 103:2677–2682. [PubMed: 16473937]
- [51]. Collins J. Instability of palindromic DNA in *Escherichia coli*. *Cold Spring Harb. Symp. Quant. Biol.* 1981; 45(Pt 1):409–416. [PubMed: 6271486]
- [52]. Collins J, Volckaert G, Nevers P. Precise and nearly-precise excision of the symmetrical inverted repeats of Tn5; common features of recA-independent deletion events in *Escherichia coli*. *Gene.* 1982; 19:139–146. [PubMed: 6292045]
- [53]. Nag DK, Petes TD. Seven-base-pair inverted repeats in DNA form stable hairpins in vivo in *Saccharomyces cerevisiae*. *Genetics.* 1991; 129:669–673. [PubMed: 1752412]
- [54]. Losch FO, Bredenbeck A, Hollstein VM, Walden P, Wrede P. Evidence for a large double-cruciform DNA structure on the X chromosome of human and chimpanzee. *Hum. Genet.* 2007; 122:337–343. [PubMed: 17638018]
- [55]. Cromie GA, Millar CB, Schmidt KH, Leach DR. Palindromes as substrates for multiple pathways of recombination in *Escherichia coli*. *Genetics.* 2000; 154:513–522. [PubMed: 10655207]
- [56]. Gordenin DA, Lobachev KS, Degtyareva NP, Malkova AL, Perkins E, Resnick MA. Inverted DNA repeats: a source of eukaryotic genomic instability. *Mol. Cell. Biol.* 1993; 13:5315–5322. [PubMed: 8395002]

- [57]. Nag DK, Kurst A. A 140-bp-long palindromic sequence induces double-strand breaks during meiosis in the yeast *Saccharomyces cerevisiae*. *Genetics*. 1997; 146:835–847. [PubMed: 9215890]
- [58]. Akgun E, Zahn J, Baumes S, Brown G, Liang F, Romanienko PJ, Lewis S, Jasin M. Palindrome resolution and recombination in the mammalian germ line. *Mol. Cell. Biol.* 1997; 17:5559–5570. [PubMed: 9271431]
- [59]. Tanaka H, Tapscott SJ, Trask BJ, Yao MC. Short inverted repeats initiate gene amplification through the formation of a large DNA palindrome in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 2002; 99:8772–8777. [PubMed: 12060719]
- [60]. Zhou ZH, Akgun E, Jasin M. Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proc. Natl. Acad. Sci. U. S. A.* 2001; 98:8326–8333. [PubMed: 11459971]
- [61]. Cunningham LA, Cote AG, Cam-Ozdemir C, Lewis SM. Rapid, stabilizing palindrome rearrangements in somatic cells by the center-break mechanism. *Mol. Cell. Biol.* 2003; 23:8740–8750. [PubMed: 14612414]
- [62]. Kurahashi H, Emanuel BS. Long AT-rich palindromes and the constitutional t(11;22) breakpoint. *Hum. Mol. Genet.* 2001; 10:2605–2617. [PubMed: 11726547]
- [63]. Edelmann L, Spiteri E, Koren K, Pulijaal V, Bialer MG, Shanske A, Goldberg R, Morrow BE. AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am. J. Hum. Genet.* 2001; 68:1–13. [PubMed: 11095996]
- [64]. Kurahashi H, Shaikh TH, Hu P, Roe BA, Emanuel BS, Budarf ML. Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22). *Hum. Mol. Genet.* 2000; 9:1665–1670. [PubMed: 10861293]
- [65]. Panigrahi GB, Lau R, Montgomery SE, Leonard MR, Pearson CE. Slipped (CTG)ⁿ(CAG)^m repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat. Struct. Mol. Biol.* 2005; 12:654–662. [PubMed: 16025129]
- [66]. Chastain PD, Sinden RR. CTG repeats associated with human genetic disease are inherently flexible. *J. Mol. Biol.* 1998; 275:405–411. [PubMed: 9466918]
- [67]. Lahue RS, Slater DL. DNA repair and trinucleotide repeat instability. *Front. Biosci.* 2003; 8:s653–665. [PubMed: 12700078]
- [68]. Katz RA, Gravuer K, Skalka AM. A preferred target DNA structure for retroviral integrase in vitro. *The Journal of biological chemistry*. 1998; 273:24190–24195. [PubMed: 9727042]
- [69]. Kurahashi H, Inagaki H, Ohye T, Kogo H, Kato T, Emanuel BS. Palindrome-mediated chromosomal translocations in humans. *DNA Repair (Amst)*. 2006; 5:1136–1145. [PubMed: 16829213]
- [70]. Wu Y, Brosh RM Jr. G-quadruplex nucleic acids and human disease. *The FEBS journal*. 2010; 277:3470–3488. [PubMed: 20670277]
- [71]. Han H, Hurley LH. G-quadruplex DNA: a potential target for anti-cancer drug design. *Trends Pharmacol. Sci.* 2000; 21:136–142. [PubMed: 10740289]
- [72]. Di Antonio M, Rodriguez R, Balasubramanian S. Experimental approaches to identify cellular G-quadruplex structures and functions. *Methods*. 2012
- [73]. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 2007; 35:406–413. [PubMed: 17169996]
- [74]. Capra JA, Paeschke K, Singh M, Zakian VA. G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* 2010; 6:e1000861. [PubMed: 20676380]
- [75]. Juranek SA, Paeschke K. Cell cycle regulation of G-quadruplex DNA structures at telomeres. *Curr. Pharm. Des.* 2012; 18:1867–1872. [PubMed: 22376109]
- [76]. Wang Q, Liu JQ, Chen Z, Zheng KW, Chen CY, Hao YH, Tan Z. G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase. *Nucleic Acids Res.* 2011; 39:6229–6237. [PubMed: 21441540]
- [77]. Han H, Hurley LH, Salazar M. A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acids Res.* 1999; 27:537–542. [PubMed: 9862977]

- [78]. Paeschke K, Capra JA, Zakian VA. DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell*. 2011; 145:678–691. [PubMed: 21620135]
- [79]. Lopes J, Piazza A, Bermejo R, Kriegsman B, Colosio A, Teulade-Fichou MP, Foiani M, Nicolas A. G-quadruplex-induced instability during leading-strand replication. *The EMBO journal*. 2011; 30:4033–4046. [PubMed: 21873979]
- [80]. Piazza A, Boule JB, Lopes J, Mingo K, Largy E, Teulade-Fichou MP, Nicolas A. Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 2010; 38:4337–4348. [PubMed: 20223771]
- [81]. Chaires JB. Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *The FEBS journal*. 2010; 277:1098–1106. [PubMed: 19951355]
- [82]. Wang G, Vasquez KM. Non-B DNA structure-induced genetic instability. *Mutat. Res*. 2006; 598:103–119. [PubMed: 16516932]
- [83]. Schwartz T, Behlke J, Lowenhaupt K, Heinemann U, Rich A. Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins. *Nat. Struct. Biol*. 2001; 8:761–765. [PubMed: 11524677]
- [84]. Herbert A, Schade M, Lowenhaupt K, Alfken J, Schwartz T, Shlyakhtenko LS, Lyubchenko YL, Rich A. The Zalpha domain from human ADAR1 binds to the ZDNA conformer of many different sequences. *Nucleic Acids Res*. 1998; 26:3486–3493. [PubMed: 9671809]
- [85]. Zannis-Hadjopoulos M, Yahyaoui W, Callejo M. 14-3-3 cruciform-binding proteins as regulators of eukaryotic DNA replication. *Trends Biochem. Sci*. 2008; 33:44–50. [PubMed: 18054234]
- [86]. Todd A, Cossons N, Aitken A, Price GB, Zannis-Hadjopoulos M. Human cruciform binding protein belongs to the 14-3-3 family. *Biochemistry (Mosc.)*. 1998; 37:14317–14325.
- [87]. Callejo M, Alvarez D, Price GB, Zannis-Hadjopoulos M. The 14-3-3 protein homologues from *Saccharomyces cerevisiae*, Bmh1p and Bmh2p, have cruciform DNA-binding activity and associate in vivo with ARS307. *The Journal of biological chemistry*. 2002; 277:38416–38423. [PubMed: 12167636]
- [88]. Brazda V, Laister RC, Jagelska EB, Arrowsmith C. Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol Biol*. 2011; 12:33. [PubMed: 21816114]
- [89]. Reddy MC, Christensen J, Vasquez KM. Interplay between human high mobility group protein 1 and replication protein A on psoralen-cross-linked DNA. *Biochemistry (Mosc.)*. 2005; 44:4188–4195.
- [90]. Mitchell DL, Nairn RS. The biology of the (6–4) photoproduct. *Photochem Photobiol*. 1989; 49:805–819. [PubMed: 2672059]
- [91]. Gunz D, Hess MT, Naegeli H. Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J Biol Chem*. 1996; 271:25089–25098. [PubMed: 8810263]
- [92]. Buschta-Hedayat N, Buterin T, Hess MT, Missura M, Naegeli H. Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proc Natl Acad Sci U S A*. 1999; 96:6090–6095. [PubMed: 10339546]
- [93]. Ogrunc M, Becker DF, Ragsdale SW, Sancar A. Nucleotide excision repair in the third kingdom. *J Bacteriol*. 1998; 180:5796–5798. [PubMed: 9791138]
- [94]. Wood RD. DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie*. 1999; 81:39–44. [PubMed: 10214908]
- [95]. Cooney M, Czernuszewicz G, Postel EH, Flint SJ, Hogan ME. Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro. *Science*. 1988; 241:456–459. [PubMed: 3293213]
- [96]. Wang G, Seidman MM, Glazer PM. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science*. 1996; 271:802–805. [PubMed: 8628995]
- [97]. Wang G, Chen Z, Zhang S, Wilson GL, Jing K. Detection and determination of oligonucleotide triplex formation-mediated transcription-coupled DNA repair in HeLa nuclear extracts. *Nucleic Acids Res*. 2001; 29:1801–1807. [PubMed: 11292853]

- [98]. Vasquez KM, Christensen J, Li L, Finch RA, Glazer PM. Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. *Proc. Natl. Acad. Sci. U. S. A.* 2002; 99:5848–5853. [PubMed: 11972036]
- [99]. Thoma BS, Wakasugi M, Christensen J, Reddy MC, Vasquez KM. Human XPC-hHR23B interacts with XPA-RPA in the recognition of triplex-directed psoralen DNA interstrand crosslinks. *Nucleic Acids Res.* 2005; 33:2993–3001. [PubMed: 15914671]
- [100]. Christensen LA, Wang H, Van Houten B, Vasquez KM. Efficient processing of TFO-directed psoralen DNA interstrand crosslinks by the UvrABC nuclease. *Nucleic Acids Res.* 2008; 36:7136–7145. [PubMed: 18996898]
- [101]. Oussatcheva EA, Hashem VI, Zou Y, Sinden RR, Potaman VN. Involvement of the nucleotide excision repair protein UvrA in instability of CAG*CTG repeat sequences in *Escherichia coli*. *J. Biol. Chem.* 2001; 276:30878–30884. [PubMed: 11413147]
- [102]. Szwarcrocka ST, Staczek P, Parniewski P. Chromosomal model for analysis of a long CTG/CAG tract stability in wild-type *Escherichia coli* and its nucleotide excision repair mutants. *Can. J. Microbiol.* 2007; 53:860–868. [PubMed: 17898841]
- [103]. Parniewski P, Bacolla A, Jaworski A, Wells RD. Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in *Escherichia coli*. *Nucleic Acids Res.* 1999; 27:616–623. [PubMed: 9862988]
- [104]. Ahn B, Grossman L. The binding of UvrAB proteins to bubble and loop regions in duplex DNA. *J. Biol. Chem.* 1996; 271:21462–21470. [PubMed: 8702929]
- [105]. Kirkpatrick DT, Petes TD. Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature.* 1997; 387:929–931. [PubMed: 9202128]
- [106]. Lin Y, Dion V, Wilson JH. Transcription promotes contraction of CAG repeat tracts in human cells. *Nat. Struct. Mol. Biol.* 2006; 13:179–180. [PubMed: 16388310]
- [107]. Jung J, Bonini N. CREB-binding protein modulates repeat instability in a *Drosophila* model for polyQ disease. *Science.* 2007; 315:1857–1859. [PubMed: 17332375]
- [108]. Lin Y, Wilson JH. Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. *Mol. Cell. Biol.* 2007; 27:6209–6217. [PubMed: 17591697]
- [109]. Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell.* 1993; 75:1027–1038. [PubMed: 8252616]
- [110]. Peltomaki P, Vasen HF. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology.* 1997; 113:1146–1158. [PubMed: 9322509]
- [111]. Fedier A, Fink D. Mutations in DNA mismatch repair genes: implications for DNA damage signaling and drug sensitivity (review). *Int. J. Oncol.* 2004; 24:1039–1047. [PubMed: 15010846]
- [112]. de Rosa M, de Sanctis D, Rosario AL, Archer M, Rich A, Athanasiadis A, Carrondo MA. Crystal structure of a junction between two Z-DNA helices. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:9088–9092. [PubMed: 20439751]
- [113]. Ha SC, Lowenhaupt K, Rich A, Kim YG, Kim KK. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature.* 2005; 437:1183–1186. [PubMed: 16237447]
- [114]. Taghian DG, Hough H, Nickoloff JA. Biased short tract repair of palindromic loop mismatches in mammalian cells. *Genetics.* 1998; 148:1257–1268. [PubMed: 9539440]
- [115]. Tran H, Degtyareva N, Gordenin D, Resnick MA. Altered replication and inverted repeats induce mismatch repair-independent recombination between highly diverged DNAs in yeast. *Mol. Cell. Biol.* 1997; 17:1027–1036. [PubMed: 9001255]
- [116]. Bill CA, Taghian DG, Duran WA, Nickoloff JA. Repair bias of large loop mismatches during recombination in mammalian cells depends on loop length and structure. *Mutat. Res.* 2001; 485:255–265. [PubMed: 11267836]

- [117]. Owen BA, Yang Z, Lai M, Gajek M, Badger JD, Hayes JJ, Edelmann W, Kucherlapati R, Wilson TM, McMurray CT. (CAG)_n-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol.* 2005; 12:663–670. [PubMed: 16025128]
- [118]. Modrich P, Lahue R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* 1996; 65:101–133. [PubMed: 8811176]
- [119]. Miret JJ, Pessoa-Brandao L, Lahue RS. Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 1997; 17:3382–3387. [PubMed: 9154837]
- [120]. Miret JJ, Pessoa-Brandao L, et al. Lahue RS. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 1998; 95:12438–12443. [PubMed: 9770504]
- [121]. Manley K, Shirley TL, Flaherty L, Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat. Genet.* 1999; 23:471–473. [PubMed: 10581038]
- [122]. van den Broek WJ, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJ, Wieringa B. Somatic expansion behaviour of the (CTG)_n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum. Mol. Genet.* 2002; 11:191–198. [PubMed: 11809728]
- [123]. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, Monckton DG. Pms2 is a genetic enhancer of trinucleotide CAG/CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Hum. Mol. Genet.* 2004; 13:1815–1825. [PubMed: 15198993]
- [124]. Savouret C, Brisson E, Essers J, Kanaar R, Pastink A, te Riele H, Junien C, Gourdon G. CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J.* 2003; 22:2264–2273. [PubMed: 12727892]
- [125]. Larson ED, Duquette ML, Cummings WJ, Streiff RJ, Maizels N. MutS α binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. *Curr. Biol.* 2005; 15:470–474. [PubMed: 15753043]
- [126]. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* 2004; 18:1618–1629. [PubMed: 15231739]
- [127]. Wu Q, Christensen LA, Legerski RJ, Vasquez KM. Mismatch repair participates in error-free processing of DNA interstrand crosslinks in human cells. *EMBO Rep.* 2005; 6:551–557. [PubMed: 15891767]
- [128]. Wu Q, Vasquez KM. Human MLH1 protein participates in genomic damage checkpoint signaling in response to DNA interstrand crosslinks, while MSH2 functions in DNA repair. *PLoS Genet.* 2008; 4:e1000189. [PubMed: 18787700]
- [129]. 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. [PubMed: 19468048]
- [130]. Robertson AB, Klungland A, Rognes T, Leiros I. DNA repair in mammalian cells: Base excision repair: the long and short of it, Cellular and molecular life sciences. *CMLS.* 2009; 66:981–993. [PubMed: 19153658]
- [131]. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, McMurray CT, Wilson SH. Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *The Journal of biological chemistry.* 2009; 284:28352–28366. [PubMed: 19674974]
- [132]. Goula AV, Berquist BR, Wilson DM 3rd, Wheeler VC, Trottier Y, Merienne K. Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet.* 2009; 5:e1000749. [PubMed: 19997493]
- [133]. Goula AV, Pearson CE, Della Maria J, Trottier Y, Tomkinson AE, Wilson DM 3rd, Merienne K. The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. *Biochemistry (Mosc.).* 2012; 51:3919–3932.
- [134]. Ruggiero BL, Topal MD. Triplet repeat expansion generated by DNA slippage is suppressed by human flap endonuclease 1. *J. Biol. Chem.* 2004; 279:23088–23097. [PubMed: 15037629]

- [135]. Schweitzer JK, Livingston DM. Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Hum. Mol. Genet.* 1998; 7:69–74. [PubMed: 9384605]
- [136]. Freudenreich CH, Kantrow SM, Zakian VA. Expansion and length-dependent fragility of CTG repeats in yeast. *Science.* 1998; 279:853–856. [PubMed: 9452383]
- [137]. Lee S, Park MS. Human FEN-1 can process the 5'-flap DNA of CTG/CAG triplet repeat derived from human genetic diseases by length and sequence dependent manner. *Exp. Mol. Med.* 2002; 34:313–317. [PubMed: 12515398]
- [138]. Spiro C, Pelletier R, Rolfsmeier ML, Dixon MJ, Lahue RS, Gupta G, Park MS, Chen X, Mariappan SV, McMurray CT. Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell.* 1999; 4:1079–1085. [PubMed: 10635332]
- [139]. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature.* 2007; 447:447–452. [PubMed: 17450122]
- [140]. Kovtun IV, Johnson KO, McMurray CT. Cockayne syndrome B protein antagonizes OGG1 in modulating CAG repeat length in vivo. *Aging.* 2011; 3:509–514. [PubMed: 21566259]
- [141]. Jarem DA, Wilson NR, Schermerhorn KM, Delaney S. Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. *DNA Repair.* 2011; 10:887–896. [PubMed: 21727036]
- [142]. Wang G, Vasquez KM. Models for chromosomal replication-independent non-B DNA structure-induced genetic instability. *Mol. Carcinog.* 2009; 48:286–298. [PubMed: 19123200]
- [143]. Vongsamphanh R, Wagner JR, Ramotar D. *Saccharomyces cerevisiae* Ogg1 prevents poly(GT) tract instability in the mitochondrial genome. *DNA Repair.* 2006; 5:235–242. [PubMed: 16293446]
- [144]. Chu WK, Hickson ID. RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer.* 2009; 9:644–654. [PubMed: 19657341]
- [145]. Bohr VA. Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. *Trends Biochem. Sci.* 2008; 33:609–620. [PubMed: 18926708]
- [146]. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA, Livingston DM. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell.* 2001; 105:149–160. [PubMed: 11301010]
- [147]. Rezazadeh S. RecQ helicases; at the crossroad of genome replication, repair, and recombination. *Mol. Biol. Rep.* 2012; 39:4527–4543. [PubMed: 21947842]
- [148]. Singh DK, Ghosh AK, Croteau DL, Bohr VA. RecQ helicases in DNA double strand break repair and telomere maintenance. *Mutat. Res.* 2011
- [149]. Vindigni A, Marino F, Gileadi O. Probing the structural basis of RecQ helicase function. *Biophys. Chem.* 2010; 149:67–77. [PubMed: 20392558]
- [150]. Zhang XD, Dou SX, Xie P, Hu JS, Wang PY, Xi XG. *Escherichia coli* RecQ is a rapid, efficient, and monomeric helicase. *The Journal of biological chemistry.* 2006; 281:12655–12663. [PubMed: 16507576]
- [151]. Tereshchenko IV, Chen Y, McDaniel LD, Schultz RA, Tischfield JA, Shao C. Small scale genetic alterations contribute to increased mutability at the X-linked Hprt locus in vivo in Blm hypomorphic mice. *DNA Repair.* 2010; 9:551–557. [PubMed: 20299287]
- [152]. Leder A, Lebel M, Zhou F, Fontaine K, Bishop A, Leder P. Genetic interaction between the unstable v-Ha-RAS transgene (Tg.AC) and the murine Werner syndrome gene: transgene instability and tumorigenesis. *Oncogene.* 2002; 21:6657–6668. [PubMed: 12242664]
- [153]. Karow JK, Constantinou A, Li JL, West SC, Hickson ID. The Bloom's syndrome gene product promotes branch migration of holliday junctions, *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97:6504–6508.
- [154]. Harmon FG, Kowalczykowski SC. RecQ helicase, in concert with RecA and SSB proteins, initiates and disrupts DNA recombination. *Genes Dev.* 1998; 12:1134–1144. [PubMed: 9553043]

- [155]. Mohaghegh P, Karow JK, Brosh RM Jr, Bohr VA, Hickson ID. The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* 2001; 29:2843–2849. [PubMed: 11433031]
- [156]. Nora GJ, Buncher NA, Opresko PL. Telomeric protein TRF2 protects Holliday junctions with telomeric arms from displacement by the Werner syndrome helicase. *Nucleic Acids Res.* 2010; 38:3984–3998. [PubMed: 20215438]
- [157]. von Kobbe C, Thoma NH, Czyzewski BK, Pavletich NP, Bohr VA. Werner syndrome protein contains three structure-specific DNA binding domains. *The Journal of biological chemistry.* 2003; 278:52997–53006. [PubMed: 14534320]
- [158]. Ozgenc A, Loeb LA. Current advances in unraveling the function of the Werner syndrome protein. *Mutat. Res.* 2005; 577:237–251. [PubMed: 15946710]
- [159]. LeRoy G, Carroll R, Kyin S, Seki M, Cole MD. Identification of RecQL1 as a Holliday junction processing enzyme in human cell lines. *Nucleic Acids Res.* 2005; 33:6251–6257. [PubMed: 16260474]
- [160]. Mankouri HW, Hickson ID. The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvasome'? *Trends Biochem. Sci.* 2007; 32:538–546. [PubMed: 17980605]
- [161]. De Muyt A, Jessop L, Kolar E, Sourirajan A, Chen J, Dayani Y, Lichten M. BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism. *Mol. Cell.* 2012; 46:43–53. [PubMed: 22500736]
- [162]. Liberi G, Maffioletti G, Lucca C, Chiolo I, Baryshnikova A, Cotta-Ramusino C, Lopes M, Pelliccioli A, Haber JE, Foiani M. Rad51-dependent DNA structures accumulate at damaged replication forks in sgs1 mutants defective in the yeast ortholog of BLM RecQ helicase. *Genes Dev.* 2005; 19:339–350. [PubMed: 15687257]
- [163]. Mankouri HW, Ashton TM, Hickson ID. Holliday junction-containing DNA structures persist in cells lacking Sgs1 or Top3 following exposure to DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:4944–4949. [PubMed: 21383164]
- [164]. Ashton TM, Mankouri HW, Heidenblut A, McHugh PJ, Hickson ID. Pathways for Holliday junction processing during homologous recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 2011; 31:1921–1933. [PubMed: 21343337]
- [165]. Dixon BP, Lu L, Chu A, Bissler JJ. RecQ and RecG helicases have distinct roles in maintaining the stability of polypurine-polypyrimidine sequences. *Mutat. Res.* 2008; 643:20–28. [PubMed: 18582477]
- [166]. Brosh RM Jr, Majumdar A, Desai S, Hickson ID, Bohr VA, Seidman MM. Unwinding of a DNA triple helix by the Werner and Bloom syndrome helicases. *J. Biol. Chem.* 2001; 276:3024–3030. [PubMed: 11110789]
- [167]. Bacolla A, Wang G, Jain A, Chuzhanova NA, Cer RZ, Collins JR, Cooper DN, Bohr VA, Vasquez KM. Non-B DNA-forming sequences and WRN deficiency independently increase the frequency of base substitution in human cells. *The Journal of biological chemistry.* 2011; 286:10017–10026. [PubMed: 21285356]
- [168]. Kamath-Loeb AS, Loeb LA, Johansson E, Burgers PM, Fry M. Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)_n trinucleotide repeat sequence. *J. Biol. Chem.* 2001; 276:16439–16446. [PubMed: 11279038]
- [169]. Huber MD, Duquette ML, Shiels JC, Maizels N. A conserved G4 DNA binding domain in RecQ family helicases. *J. Mol. Biol.* 2006; 358:1071–1080. [PubMed: 16530788]
- [170]. Crabbe L, Karlseder J. In the end, it's all structure. *Curr. Mol. Med.* 2005; 5:135–143. [PubMed: 15974866]
- [171]. Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature.* 2002; 417:876–880. [PubMed: 12050675]
- [172]. Opresko PL. Telomere ResQue and preservation--roles for the Werner syndrome protein and other RecQ helicases. *Mech. Ageing Dev.* 2008; 129:79–90. [PubMed: 18054793]
- [173]. London TB, Barber LJ, Mosedale G, Kelly GP, Balasubramanian S, Hickson ID, Boulton SJ, Hiom K. FANCD1 is a structure-specific DNA helicase associated with the maintenance of

- genomic G/C tracts. *The Journal of biological chemistry*. 2008; 283:36132–36139. [PubMed: 18978354]
- [174]. Shen X, Do H, Li Y, Chung WH, Tomasz M, de Winter JP, Xia B, Elledge SJ, Wang W, Li L. Recruitment of fanconi anemia and breast cancer proteins to DNA damage sites is differentially governed by replication. *Mol. Cell*. 2009; 35:716–723. [PubMed: 19748364]
- [175]. Cantor SB, Xie J. Assessing the link between BACH1/FANCI and MLH1 in DNA crosslink repair. *Environ. Mol. Mutagen*. 2010; 51:500–507. [PubMed: 20658644]
- [176]. Wu Y, Brosh RM Jr. FANCI helicase operates in the Fanconi Anemia DNA repair pathway and the response to replicational stress. *Curr. Mol. Med*. 2009; 9:470–482. [PubMed: 19519404]
- [177]. Cheung I, Schertzer M, Rose A, Lansdorp PM. Disruption of dog-1 in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA. *Nat. Genet*. 2002; 31:405–409. [PubMed: 12101400]
- [178]. Youds JL, O'Neil NJ, Rose AM. Homologous recombination is required for genome stability in the absence of DOG-1 in *Caenorhabditis elegans*. *Genetics*. 2006; 173:697–708. [PubMed: 16547095]
- [179]. Bochman ML, Sabouri N, Zakian VA. Unwinding the functions of the Pif1 family helicases. *DNA Repair*. 2010; 9:237–249. [PubMed: 20097624]
- [180]. Brooks TA, Kendrick S, Hurley L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *The FEBS journal*. 2010; 277:3459–3469. [PubMed: 20670278]
- [181]. Sanders CM. Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *The Biochemical journal*. 2010; 430:119–128. [PubMed: 20524933]
- [182]. Chakraborty P, Grosse F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair*. 2011; 10:654–665. [PubMed: 21561811]
- [183]. Zhang S, Hemmerich P, Grosse F. Werner syndrome helicase (WRN), nuclear DNA helicase II (NDH II) and histone gammaH2AX are localized to the centrosome. *Cell Biol. Int*. 2007; 31:1109–1121. [PubMed: 17498979]
- [184]. Jain A, Bacolla A, Chakraborty P, Grosse F, Vasquez KM. Human DHX9 helicase unwinds triple-helical DNA structures. *Biochemistry (Mosc.)*. 2010; 49:6992–6999.
- [185]. Mischo HE, Hemmerich P, Grosse F, Zhang S. Actinomycin D induces histone gamma-H2AX foci and complex formation of gamma-H2AX with Ku70 and nuclear DNA helicase II. *The Journal of biological chemistry*. 2005; 280:9586–9594. [PubMed: 15613478]
- [186]. Kopel V, Pozner A, Baran N, Manor H. Unwinding of the third strand of a DNA triple helix, a novel activity of the SV40 large T-antigen helicase. *Nucleic Acids Res*. 1996; 24:330–335. [PubMed: 8628658]
- [187]. Vannier JB, Pavicic-Kaltenbrunner V, Petalcorin MI, Ding H, Boulton SJ. RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell*. 2012; 149:795–806. [PubMed: 22579284]
- [188]. Anand RP, Shah KA, Niu H, Sung P, Mirkin SM, Freudenreich CH. Overcoming natural replication barriers: differential helicase requirements. *Nucleic Acids Res*. 2012; 40:1091–1105. [PubMed: 21984413]
- [189]. Kerrest A, Anand RP, Sundararajan R, Bermejo R, Liberi G, Dujon B, Freudenreich CH, Richard GF. SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. *Nat. Struct. Mol. Biol*. 2009; 16:159–167. [PubMed: 19136956]
- [190]. Bhattacharyya S, Lahue RS. *Saccharomyces cerevisiae* Srs2 DNA helicase selectively blocks expansions of trinucleotide repeats. *Mol. Cell. Biol*. 2004; 24:7324–7330. [PubMed: 15314145]
- [191]. Sarkies P, Murat P, Phillips LG, Patel KJ, Balasubramanian S, Sale JE. FANCI coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic Acids Res*. 2012; 40:1485–1498. [PubMed: 22021381]
- [192]. Bhattacharyya S, Lahue RS. *Saccharomyces cerevisiae* Srs2 DNA helicase selectively blocks expansions of trinucleotide repeats. *Mol. Cell. Biol*. 2004; 24:7324–7330. [PubMed: 15314145]
- [193]. Shishkin AA, Voineagu I, Matera R, Chong N, Chernet BT, Krasilnikova MM, Narayanan V, Lobachev KS, Mirkin SM. Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol. Cell*. 2009; 35:82–92. [PubMed: 19595718]

- [194]. Wang JC. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* 2002; 3:430–440. [PubMed: 12042765]
- [195]. Froelich-Ammon SJ, Gale KC, Osheroff N. Site-specific cleavage of a DNA hairpin by topoisomerase II. DNA secondary structure as a determinant of enzyme recognition/cleavage. *J. Biol. Chem.* 1994; 269:7719–7725. [PubMed: 8125998]
- [196]. Pognan F, Paoletti C. Does cruciform DNA provide a recognition signal for DNA-topoisomerase II? *Biochimie.* 1992; 74:1019–1023. [PubMed: 1335757]
- [197]. Belotserkovskii BP, di Silva E, Tornaletti S, Wang G, Vasquez KM, Hanawalt PC. A triplex-forming sequence from the human c-Myc promoter interferes with DNA transcription. *J. Biol. Chem.* 2007; 282:32433–32441. [PubMed: 17785457]
- [198]. Ditlevson JV, Tornaletti S, Belotserkovskii BP, Teijeiro V, Wang G, Vasquez KM, Hanawalt PC. Inhibitory effect of a short Z-DNA forming sequence on transcription elongation by T7 RNA polymerase. *Nucleic Acids Res.* 2008; 36:3163–3170. [PubMed: 18400779]
- [199]. Abdulovic AL, Hile SE, Kunkel TA, Eckert KA. The in vitro fidelity of yeast DNA polymerase delta and polymerase epsilon holoenzymes during dinucleotide microsatellite DNA synthesis. *DNA Repair.* 2011; 10:497–505. [PubMed: 21429821]
- [200]. Ji J, Clegg NJ, Peterson KR, Jackson AL, Laird CD, Loeb LA. In vitro expansion of GGC:GCC repeats: identification of the preferred strand of expansion. *Nucleic Acids Res.* 1996; 24:2835–2840. [PubMed: 8759019]
- [201]. Betous R, Rey L, Wang G, Pillaire MJ, Puget N, Selves J, Biard DS, Shin-ya K, Vasquez KM, Cazaux C, Hoffmann JS. Role of TLS DNA polymerases eta and kappa in processing naturally occurring structured DNA in human cells. *Mol. Carcinog.* 2009; 48:369–378. [PubMed: 19117014]
- [202]. Lawrence CW. Cellular roles of DNA polymerase zeta and Rev1 protein. *DNA Repair.* 2002; 1:425–435. [PubMed: 12509231]
- [203]. Collins NS, Bhattacharyya S, Lahue RS. Rev1 enhances CAG:CTG repeat stability in *Saccharomyces cerevisiae*. *DNA Repair.* 2007; 6:38–44. [PubMed: 16979389]
- [204]. Barone F, Begusova M, La Nave E, Matzeu M, Mazzei F, Sy D. Radiation damage to triplex DNA induced by gamma-rays: a footprinting study and Monte Carlo simulation. *Int. J. Radiat. Biol.* 2000; 76:731–740. [PubMed: 10902726]
- [205]. Sun JS, Lavery R, Chomilier J, Zakrzewska K, Montenay-Garestier T, Helene C. Theoretical study of ethidium intercalation in triple-stranded DNA and at triplex-duplex junctions. *J. Biomol. Struct. Dyn.* 1991; 9:425–436. [PubMed: 1815636]
- [206]. Collier DA, Mergny JL, Thuong NT, Helene C. Site-specific intercalation at the triplex-duplex junction induces a conformational change which is detectable by hypersensitivity to diethylpyrocarbonate. *Nucleic Acids Res.* 1991; 19:4219–4224. [PubMed: 1870975]
- [207]. Sy D, Savoye C, Begusova M, Michalik V, Charlier M, Spothem-Maurizot M. Sequence-dependent variations of DNA structure modulate radiation-induced strand breakage. *Int. J. Radiat. Biol.* 1997; 72:147–155. [PubMed: 9269307]
- [208]. Kim D, Reddy S, Kim DY, Rich A, Lee S, Kim KK, Kim YG. Base extrusion is found at helical junctions between right- and left-handed forms of DNA and RNA. *Nucleic Acids Res.* 2009; 37:4353–4359. [PubMed: 19465399]
- [209]. Yang XL, Wang AH. Structural analysis of Z-Z DNA junctions with A:A and T:T mismatched base pairs by NMR. *Biochemistry (Mosc.)*. 1997; 36:4258–4267.
- [210]. Rodolfo C, Lanza A, Tornaletti S, Fronza G, Pedrini AM. The ultimate carcinogen of 4-nitroquinoline 1-oxide does not react with Z-DNA and hyperreacts with B-Z junctions. *Nucleic Acids Res.* 1994; 22:314–320. [PubMed: 8127667]
- [211]. Johnston BH, Rich A. Chemical probes of DNA conformation: detection of Z-DNA at nucleotide resolution. *Cell.* 1985; 42:713–724. [PubMed: 4053183]
- [212]. Zimmerman SB. The three-dimensional structure of DNA. *Annu. Rev. Biochem.* 1982; 51:395–427. [PubMed: 7051961]
- [213]. Tartier L, Michalik V, Spothem-Maurizot M, Rahmouni AR, Sabattier R, Charlier M. Radiolytic signature of Z-DNA. *Nucleic Acids Res.* 1994; 22:5565–5570. [PubMed: 7838708]

- [214]. Jarem DA, Wilson NR, Delaney S. Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry (Mosc.)*. 2009; 48:6655–6663.
- [215]. Clark DW, Phang T, Edwards MG, Geraci MW, Gillespie MN. Promoter G-quadruplex sequences are targets for base oxidation and strand cleavage during hypoxia-induced transcription. *Free Radic. Biol. Med.* 2012; 53:51–59. [PubMed: 22583700]
- [216]. Ndllebe T, Schuster GB. Long-distance radical cation transport in DNA: Horizontal charge hopping in a dimeric quadruplex. *Org. Biomol. Chem.* 2006; 4:4015–4021. [PubMed: 17047883]
- [217]. Delaney S, Barton JK. Charge transport in DNA duplex/quadruplex conjugates. *Biochemistry (Mosc.)*. 2003; 42:14159–14165.
- [218]. Friedberg, EC.; Walker, GC.; Siede, W.; Wood, RD.; Schultz, RA.; Ellenberger, T. DNA repair and mutagenesis. 2nd ed.. ASM Press; Washington, D.C.: 2006.
- [219]. Lagravere C, Malfoy B, Leng M, Laval J. Ring-opened alkylated guanine is not repaired in Z-DNA. *Nature*. 1984; 310:798–800. [PubMed: 6382026]
- [220]. Boiteux S, Costa de Oliveira R, Laval J. The Escherichia coli O6-methylguanine-DNA methyltransferase does not repair promutagenic O6-methylguanine residues when present in Z-DNA. *J. Biol. Chem.* 1985; 260:8711–8715. [PubMed: 3894348]
- [221]. Boiteux S, Laval F. Repair of O6-methylguanine, by mammalian cell extracts, in alkylated DNA and poly(dG-m5dC). (poly dG-m5dC) in B and Z forms. *Carcinogenesis*. 1985; 6:805–807. [PubMed: 4006067]
- [222]. Ahn B, Bohr VA. DNA secondary structure of the released strand stimulates WRN helicase action on forked duplexes without coordinate action of WRN exonuclease, *Biochem. Biophys. Res. Commun.* 2011; 411:684–689.
- [223]. Shen CK, Hearst JE. Psoralen-crosslinked secondary structure map of single-stranded virus DNA. *Proc. Natl. Acad. Sci. U. S. A.* 1976; 73:2649–2653. [PubMed: 1066675]
- [224]. Volker J, Plum GE, Klump HH, Breslauer KJ. DNA repair and DNA triplet repeat expansion: the impact of abasic lesions on triplet repeat DNA energetics. *J. Am. Chem. Soc.* 2009; 131:9354–9360. [PubMed: 19566100]
- [225]. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997; 389:251–260. [PubMed: 9305837]
- [226]. Ruan H, Wang YH. Friedreich's ataxia GAA.TTC duplex and GAA.GAA.TTC triplex structures exclude nucleosome assembly. *J. Mol. Biol.* 2008; 383:292–300. [PubMed: 18775729]
- [227]. Westin L, Blomquist P, Milligan JF, Wrange O. Triple helix DNA alters nucleosomal histone-DNA interactions and acts as a nucleosome barrier. *Nucleic Acids Res.* 1995; 23:2184–2191. [PubMed: 7610046]
- [228]. Brown PM, Madden CA, Fox KR. Triple-helix formation at different positions on nucleosomal DNA. *Biochemistry (Mosc.)*. 1998; 37:16139–16151.
- [229]. Brown PM, Fox KR. DNA triple-helix formation on nucleosome-bound poly(dA).poly(dT) tracts. *The Biochemical journal*. 1998; 333(Pt 2):259–267. [PubMed: 9657964]
- [230]. Brown PM, Fox KR. DNA triple-helix formation on nucleosome core particles. Effect of length of the oligopurine tract. *European journal of biochemistry / FEBS*. 1999; 261:301–310. [PubMed: 10103063]
- [231]. Espinas ML, Jimenez-Garcia E, Martinez-Balbas A, Azorin F. Formation of triple-stranded DNA at d(GA.TC)n sequences prevents nucleosome assembly and is hindered by nucleosomes. *J. Biol. Chem.* 1996; 271:31807–31812. [PubMed: 8943221]
- [232]. Wong B, Chen S, Kwon JA, Rich A. Characterization of Z-DNA as a nucleosome-boundary element in yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 2007; 104:2229–2234. [PubMed: 17284586]
- [233]. Davey C, Pennings S, Allan J. CpG methylation remodels chromatin structure in vitro. *J. Mol. Biol.* 1997; 267:276–288. [PubMed: 9096225]
- [234]. Davey C, Allan J. Nucleosome positioning signals and potential H-DNA within the DNA sequence of the imprinting control region of the mouse *Igf2r* gene. *Biochim. Biophys. Acta*. 2003; 1630:103–116. [PubMed: 14654240]

- [235]. Godde JS, Kass SU, Hirst MC, Wolffe AP. Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter. *The Journal of biological chemistry*. 1996; 271:24325–24328. [PubMed: 8798682]
- [236]. Wong HM, Huppert JL. Stable G-quadruplexes are found outside nucleosome-bound regions. *Mol. Biosyst.* 2009; 5:1713–1719. [PubMed: 19585004]
- [237]. Takagi H, Inai Y, Watanabe S, Tatemoto S, Yajima M, Akasaka K, Yamamoto T, Sakamoto N. Nucleosome exclusion from the interspecies-conserved central AT-rich region of the *Ars* insulator. *J. Biochem. (Tokyo)*. 2012; 151:75–87. [PubMed: 21930654]
- [238]. Wang YH, Griffith J. Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. *Genomics*. 1995; 25:570–573. [PubMed: 7789994]
- [239]. Wang YH, Amirhaeri S, Kang S, Wells RD, Griffith JD. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science*. 1994; 265:669–671. [PubMed: 8036515]
- [240]. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, Klesert TR, Lobanenko VV, Tapscott SJ. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* 2001; 28:335–343. [PubMed: 11479593]
- [241]. Otten AD, Tapscott SJ. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. Natl. Acad. Sci. U. S. A.* 1995; 92:5465–5469. [PubMed: 7777532]
- [242]. Vardimon L, Rich A. In Z-DNA the sequence G-C-G-C is neither methylated by Hha I methyltransferase nor cleaved by Hha I restriction endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 1984; 81:3268–3272. [PubMed: 6328508]
- [243]. Bock C, Paulsen M, Tierling S, Mikeska T, Lengauer T, Walter J. CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet.* 2006; 2:e26. [PubMed: 16520826]
- [244]. Han YJ, de Lanerolle P. Naturally extended CT . AG repeats increase H-DNA structures and promoter activity in the smooth muscle myosin light chain kinase gene. *Mol. Cell. Biol.* 2008; 28:863–872. [PubMed: 17991897]

Figure 1. Schematic diagram of DNA conformations

(A) Canonical B-DNA; (B) intra-molecular triplex , H-DNA; (C) left-handed Z-DNA; (D) cruciform; (E) hairpin (left) and slippage loop (right); (F) G-quadruplex.

Figure 2. DNA damage/repair and non-B DNA conformation-induced genetic instability

A repetitive sequence in the B-form is unwound from the histone core during DNA metabolism, *e.g.*, DNA replication, transcription, or repair (step 1), and an altered DNA secondary structure is formed at the repetitive sequence (here shown as a Z-DNA structure). The non-B conformation can alter nucleosome reassembly, and may result in chromatin remodeling in nearby regions (step 2). The altered DNA structure can affect DNA damage and repair activities (step 3), and interfere with DNA replication (step 4). DNA repair proteins can recognize, bind to, and process the non-B DNA conformation, resulting in either melting of the DNA conformation (step 5) or to the formation DSBs eventually leading to deletion and/or translocation events (step 6).