BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Interconverting Conformations of Slipped-DNA Junctions Formed by Trinucleotide Repeats Affect Repair Outcome

Meghan M. Slean,[†,‡] Kaalak Reddy,[†,‡] Bin Wu,[§] Kerrie Nichol Edamura,[†,⊥] Mariana Kekis,[†,‡] Frank H. T. Nelissen,[§] Ruud L. E. G. Aspers,[§] Marco Tessari,[§] Orlando D. Schärer,[∥] Sybren S. Wijmenga,[§] and Christopher E. Pearson*[,†,‡]

[†]Program of Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 1L7, Canada

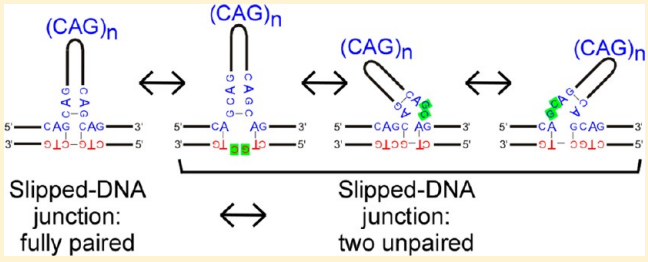
[‡]Department of Molecular Genetics, University of Toronto, Toronto, ON M5G 1L7, Canada

[§]Institute of Molecules and Materials, Department of Biophysical Chemistry, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

[∥]Department of Pharmacological Sciences and Department of Chemistry, Stony Brook University, Stony Brook, New York 11794-3400, United States

Supporting Information

ABSTRACT: Expansions of (CTG)·(CAG) repeated DNAs are the mutagenic cause of 14 neurological diseases, likely arising through the formation and processing of slipped-strand DNAs. These transient intermediates of repeat length mutations are formed by out-of-register mispairing of repeat units on complementary strands. The three-way slipped-DNA junction, at which the excess repeats slip out from the duplex, is a poorly understood feature common to these mutagenic intermediates. Here, we reveal that slipped junctions can assume a surprising number of interconverting conformations where the strand opposite the slip-out either is fully base paired or has one or two unpaired nucleotides. These unpaired nucleotides can also arise opposite either of the nonslipped junction arms. Junction conformation can affect binding by various structure-specific DNA repair proteins and can also alter correct nick-directed repair levels. Junctions that have the potential to contain unpaired nucleotides are repaired with a significantly higher efficiency than constrained fully paired junctions. Surprisingly, certain junction conformations are aberrantly repaired to expansion mutations: misdirection of repair to the nonnicked strand opposite the slip-out leads to integration of the excess slipped-out repeats rather than their excision. Thus, slipped-junction structure can determine whether repair attempts lead to correction or expansion mutations.



Genome-wide and gene-specific mutations can lead to evolutionary variation, cancer, and neurodegenerative and neuromuscular disease.[1,2] The discovery that genetic expansions of (CTG)·(CAG) repeated DNA sequences are the mutagenic cause of 14 progressive diseases, including Huntington's disease and myotonic dystrophy, stimulated renewed interest in the formation and repair of slipped-strand DNA structures, the supposed mutagenic intermediates of repeat tract mutations.[2−4] Expansion mutations can arise during postzygotic cell divisions[5] and continue in patients as they age, coinciding with worsening symptoms.[6] Patients exhibit intertissue repeat length differences as great as 5770 repeats, with larger expansions occurring in affected tissues such as cerebral cortex, muscle, and heart, indicating high levels of continued expansions coinciding with disease progression.[6] Understanding the mechanisms of repeat instability is crucial to arresting or reversing disease.[4]

All models proposed to explain repeat expansions involve DNA slippage at the repeats (Figure S1 of the Supporting Information; reviewed in refs 2−4). The formation and aberrant repair of slipped-strand DNAs are likely sources of repeat instability. Slipped-strand DNAs produced by misalignment of the repeats are thought to be transient mutagenic intermediates formed in mitotic cells at replication forks or in nonmitotic cells at sites of DNA damage or recombination (Figure S1 of the Supporting Information).

Structural features of slipped DNAs can affect repair outcome.[7−15] For example, slip-outs of CAG are repaired by human cell extracts with greater efficiency than CTG slip-outs,[14] and shorter slip-outs are repaired with greater efficiency than longer slip-outs.[15] Slipped DNA structure can determine which proteins are recruited for repair; for example, short but not long slip-outs require hMutSβ for repair.[15] An overlooked component of all slipped DNAs is the junction at which the slip-out extrudes from the complementary paired duplex. Models of slipped mono- and dinucleotide repeats have been

Received: October 5, 2012
Revised: December 8, 2012
Published: January 22, 2013

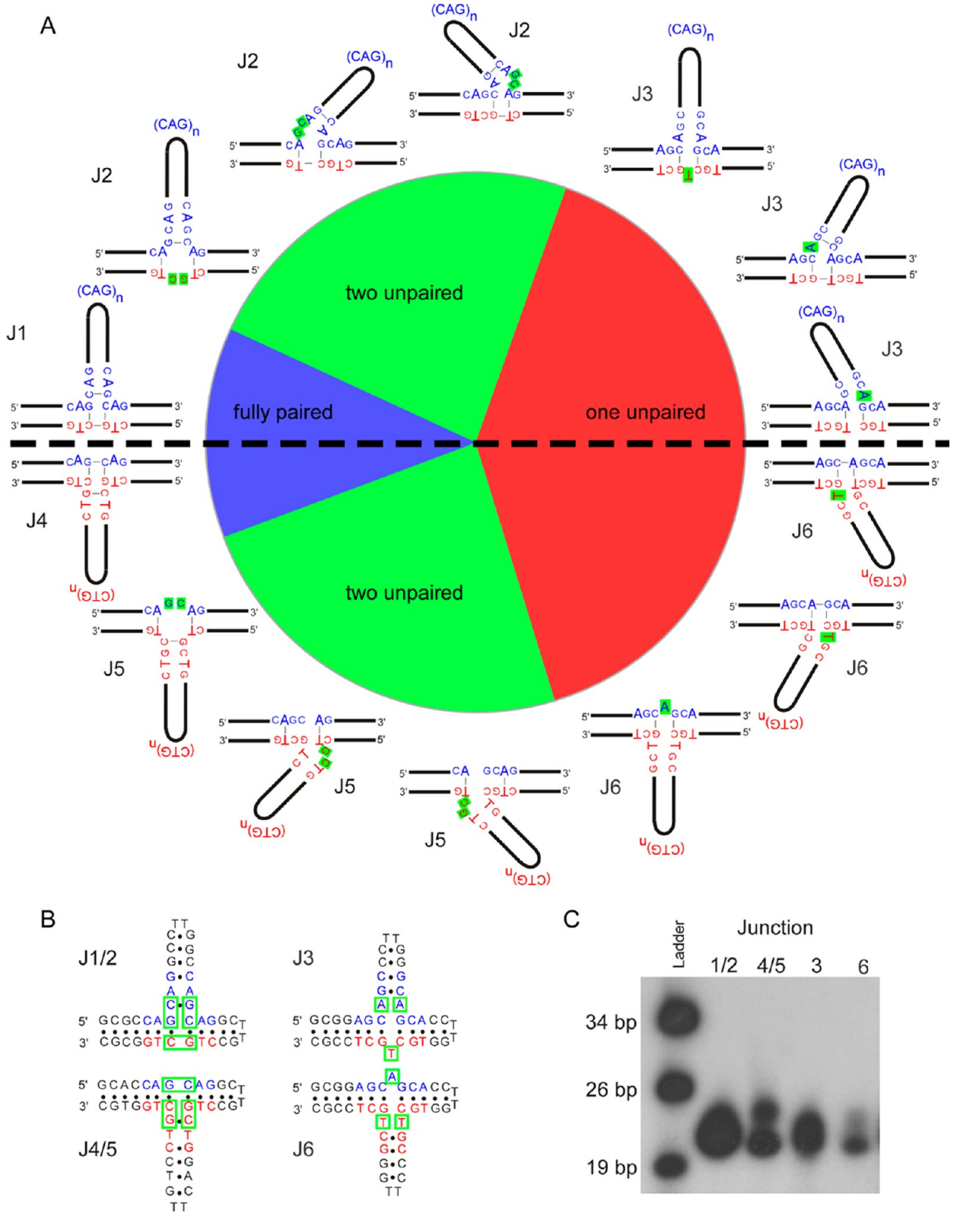


Figure 1. Slipped DNA junctions formed by (CTG)·(CAG) repeats. (A and B) Slipped DNAs are composed of three arms, two made of complementary repeat strands and the third being the CAG (blue) or CTG (red) repeat slip-out. Slip-outs assume intrastrand hairpins, and their structural characterizations have been extensive.[7−15] Possible junction conformations include fully paired strands at the slip-out or those with one or two unpaired bases at the junction, where the latter can potentially interconvert with the fully paired form. The angles displayed between the junction arms are not meant to reflect the actual angles observed in the three-way junctions. In panel A, unpaired nucleotides are highlighted in green; in panel B, nucleotides with the potential to be unpaired are shown in green boxes. (C) Electrophoretic migration of slipped-junction species was slower than expected. DNAs are a single species on denaturing gels (Figure S3 of the Supporting Information).

dx.doi.org/10.1021/bi301369b | *Biochemistry* 2013, 52, 773−785

limited to heteroduplexes that contain a repeat sequence in the absence of its complementary strand [for example, a $(CA)_4$ slip-out without a complementary $(TG)_n$ tract].[16−18] The structure of complementary three-way slipped junctions has not been characterized for trinucleotide repeats, and the influence of junction conformation upon repair outcome is unknown.

Two general forms of slipped junctions can form between complementary $(CTG)_n$·$(CAG)_n$ repeats: those with slipped-out CAG repeats and those with slipped-out CTG repeats (Figure 1A). One arm of the three-way junction is composed of the excess repeats (the slip-out, CAG or CTG), while the other two arms are comprised of complementary paired repeat strands. There are several possible conformers for slipped junctions, differing by the base pairing occurring at the junction (Figure 1A,B). The three-way junction can contain two, one, or no unpaired bases in the strand opposite any one of the three arms (Figure 1A, nucleotides highlighted in green). Interconversion between the fully paired J1 conformation and the two-unpaired nucleotide J2 conformations could theoretically occur with limited exchanges of base-pairing partners immediately at the junction. Similarly, any of the J2 conformations may interconvert between each other. It is unlikely that either the fully paired J1 junction or the two unpaired J2 junctions could interconvert to the one-unpaired nucleotide conformation, J3, as this would require breaking all base pairs in the slip-out and at least a portion of the base pairs in one other arm of the junction. The J3 junction would be equivalent to a J1 junction with an out-of-register single-CTG repeat slipped-out opposite the CAG slip-out (Figure S2 of the Supporting Information). Similarly, interconversion between the fully paired J4 and the two-unpaired nucleotide J5 conformations could theoretically occur, as might interconversions among any of the single unpaired nucleotide J6 forms, but interconversion between J6 and either J4 or J5 forms is unlikely. It is unknown which of these possible conformers are present in slipped structures, if they interconvert, or if they are processed in the same manner.

These slipped-junction conformations provide a potential biological situation for unpaired bases at DNA junctions. Unpaired bases may biophysically alter slipped junctions, as observed for three-way DNA junctions formed by heterologous sequences,[19,20] and could modulate recognition by DNA-metabolizing proteins. While unpaired nucleotides at RNA junctions have known biological consequences,[21] the biological relevance of unpaired nucleotides at three-way DNA junctions of trinucleotide repeats is unknown.

Here we determined that models of slipped-DNA junctions can assume multiple interconverting isoforms that can be differentially recognized by structure-specific DNA-binding proteins, and we reveal that the junction conformation can affect DNA repair, determining whether slip-outs are correctly repaired or lead to expansion mutations.

EXPERIMENTAL PROCEDURES

Structure Formation. Oligonucleotides were ordered from Sigma-Genosys and gel-purified twice from denaturing (sequencing) gels (6%) to ensure they were free of $n - 1$ products. DNAs were uniquely ^{32}P-labeled and structures formed by alkaline denaturation and/or renaturation as described in detail.[11] Briefly, ^{32}P-labeled oligos were dried in a speed vacuum and then denatured in a solution of 500 mM NaOH (yielding a pH of 13) and 0.5 M NaCl at room temperature for 5 min. Samples were then neutralized by the addition of a 50-fold volume of 50 mM Tris-HCl (pH 8) and 5 mM EDTA, resulting in a final solution with 0.01 M NaOH and 0.03 M NaCl (pH 8). These conditions favor full renaturation. Samples were then incubated at 68 °C for 3 h, followed by two precipitations in ethanol. Longer incubations did not change the pattern of products formed. Caution was taken to avoid sample dehydration.

Purified Proteins. Recombinant, baculovirus-expressed human HMGB1 (amino acids 1−215) was purchased from ATGen (catalog no. HMG0801). Purified bacterial (*Escherichia coli*) MutS protein was purchased from Gene Check Inc. (catalog no. GC-001). Recombinant, His-tagged human MutSα and MutSβ complexes were expressed in insect cells and purified to near homogeneity as described previously.[15] Baculoviruses expressing hMSH2 and His-tagged hMSH3 and hMSH6 were kindly provided by G.-M. Li. XPG and XPF-ERCC1 were purified as previously described.[22]

Electrophoretic Mobility Shift Assays and Cleavage Assays. Purified HMGB1 was incubated with 100 fmol of [γ-^{32}P]ATP-labeled DNA in reaction buffer containing 50 mM NaCl, 25 mM Tris (pH 7.5), 1 mM EDTA, 1 mM DTT, 0.3% glycerol, and 10 ng/μL BSA in a final volume of 20 μL. Reaction mixtures were incubated on ice for 20 min. Samples were then analyzed on a 6% native polyacrylamide gel in Tris-borate EDTA buffer. Gels were subsequently dried and exposed to X-ray film.

Purified bacterial MutS, human MutSα/β complex, XPG, and XPF-ERCC1 were incubated with 100 fmol of [γ-^{32}P]ATP-labeled DNA (only 50 fmol for MutSα/β) in reaction buffer containing 10 mM HEPES (pH 7.5), 110 mM KCl, 1 mM EDTA, and 1 mM DTT with 2.5 ng/μL poly(dI-dC) in a final volume of 20 μL. Reaction mixtures were incubated on ice for 20 min. Samples were then analyzed on a 6% native polyacrylamide gel in Tris-borate EDTA buffer. Gels were subsequently dried and exposed to X-ray film.

Repair Substrates. A set of circular heteroduplex substrates were prepared using $(CTG)_n$·$(CAG)_n$ repeats (n = 0, 1, 2, or 17) contained within the human DM1 flanking sequence as previously described.[14] Briefly, circular heteroduplex DNA was produced by hybridizing circular single-stranded DNA with linearized double-stranded DNA with a different number of repeats, and then heteroduplex molecules were gel purified. The combination of zero repeats paired with 17 repeats creates a perfectly paired three-way junction; one repeat across from 17 repeats leaves one unpaired nucleotide at the junction, and two repeats across from 17 repeats can create a junction with zero to two unpaired nucleotides. Substrates are named on the basis of the junctions that they model (from Figure 1A): a substrate modeling junction J1 would be called p-J1.

Extracts. Whole cell extracts were prepared as described previously[14] from HeLa S3 cells (National Cell Culture Center, National Center for Research Resources, National Institutes of Health, Bethesda, MD) and LoVo cells (ATCC).

Repair Reactions. Repair reactions were conducted and efficiencies quantified as previously described.[14] After incubation of the circular repair substrates with cell extracts, the repeat-containing fragment was excised from the substrates using *Eco*RI and *Hin*dIII and run on 4% polyacrylamide gels before Southern blotting was conducted. Southern probes were generated using an *Eco*RI/*Hin*dIII fragment from DNA containing 17 repeats. Repair percentages were calculated as previously described[14,15] by determining the intensity of the homoduplex bands in each lane as a percentage of all the

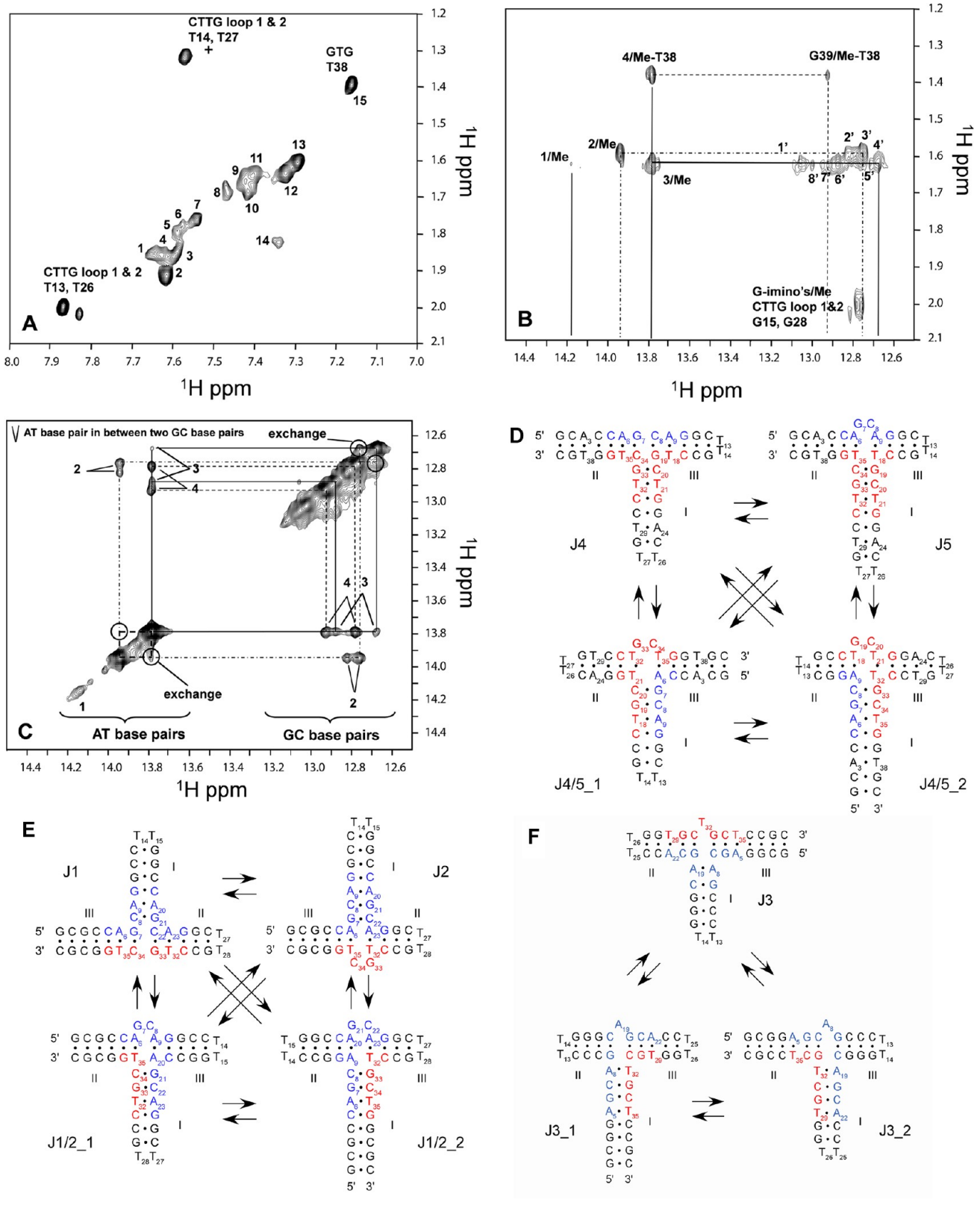


Figure 2. NMR data reveal that slipped junctions form multiple interconverting conformations. (A–C) NMR spectra for J4/5. (A) T Me–T H6 section of the 2D TOCSY spectrum in D_2O at 20 °C. (B) Imino to T Me and (C) imino to imino sections of the 600 MHz 2D NOESY spectrum of J4/5 recorded in H_2O at 10 °C with a mixing time of 300 ms. For an explanation of the J4/5 NMR spectra in panels A–C, see the text and a detailed description in section V of the Supporting Information. In panel A, the cross-peaks at characteristic chemical shift positions of the T's in the two CTTG tetraloops that cap two arms (panel D) are labeled with their residue numbers (T_{13}, T_{14}, T_{26}, and T_{27}). The other cross-peak resonances in panel A are numbered from 1 to 15; cross-peak 15 can be assigned to T_{38}. In panels B and C, the imino resonances from T's in A:T base pairs are numbered from 1 to 4. Except for number 1, which shows no cross-peaks to G imino resonances (panel C), all these imino resonances show two cross-peaks to G iminos and are thus from A:T base pairs flanked by two C:G base pairs. The presence of multiple conformations is indicated by, for

dx.doi.org/10.1021/bi301369b | *Biochemistry* 2013, 52, 773–785

Figure 2. continued

instance, the exchange cross-peaks in panel C (see, e.g., the encircled cross-peak between 2 and 3 labeled exchange in the bottom left of panel C near the diagonal; note that in NOESY spectra, the intensities of symmetry-related cross-peaks are not always equal), and the multiple imino−methyl contacts in panel B (see the text). Imino resonance 4 can be assigned to T_{38} because of the contact with the T_{38} Me seen in panel B (see the text). Imino resonance 3 includes more than two imino protons and/or multiple conformations as described in the text. (D−F) Schematics of conformations of J4/5, J1/2, and J3, respectively, that are consistent with their NMR spectra (NMR detailed in Figures S5−S7, Tables S3, and sections V−VII of the Supporting Information). In panel D, species J5, J4/5_1, and J4/5_2 are the J5 variants outlined in Figure 1A. In panel E, species J2, J1/2_1, and J1/2_2 are the J2 variants outlined in Figure 1A. In panel F, species J3, J3_1, and J3_2 are the J3 variants outlined in Figure 1A. The CAG repeats are colored blue and the CTG repeats red. The locations of pertinent nucleotides are indicated by their sequence order. The most likely stacking of the arms in the three-way junctions is shown using published arm nomenclature I, II, and III (Figure S4 of the Supporting Information).[61] Stacking preferences are based on published loop and pyrimidine rules.[20,25]

material in the lane, followed by subtraction of the corresponding band intensities in the starting material. Graphs shown are based on at least three experiments.

RESULTS

Oligonucleotide Models of Slipped-DNA Junctions. All of the possible junction conformers shown in Figure 1A can be modeled by four synthetic oligonucleotides (sequences in Table S1 of the Supporting Information): two for the slipped-out CAG and two for the slipped-out CTG junctions (Figure 1B; CAG repeats are colored blue and CTG repeats red). We prepared a series of 39−41-mer sequences, each expected to fold into three-way DNA junctions, in which the junction sequences varied but the junction arms and their hairpin tips and end sequences were similar (Figure 1B). In this manner, the structural and biological characterization could permit comparison of junction conformations. The following considerations guided oligo design. Sequence organization included two inverted repeat sequences, where the first would ensure the formation of an intrastrand hairpin constituting one of the two non-repeat-containing junction arms (Figure 1B, right arms). The second inverted repeat sequence constituted the "slip-out" region; it was adjoined to CTG or CAG units to facilitate its forming a distinct junction arm that minimized alternative mispairings with the opposing complementary CAG or CTG region. Interconversion between certain conformations could occur with small changes immediately at the junction. Oligonucleotide J1/2 could assume either the fully paired J1 conformation or any of the two-unpaired nucleotide J2 forms (shown in the top left section of Figure 1A). Oligonucleotide J3 could assume a junction conformation with one extra nucleotide (any of those shown in the top right section of Figure 1A). For CTG slip-outs, oligonucleotide J4/5 could assume either the fully paired J4 conformation or any of the two-unpaired nucleotide J5 forms (shown in the bottom left section of Figure 1A). Oligonucleotide J6 could assume a junction conformation with one extra nucleotide (any of those shown in the bottom right section of Figure 1A). Other interconversions are unlikely because of sequence constraints. The G and C content of the oligos was kept high, to optimize the self-association of the three-arm complex. Moreover, CTTG loops were included because they are highly stable and thus direct folding toward a three-arm conformation. AT base pairs were included in all three helical domains to provide spectroscopic markers for the nuclear magnetic resonance (NMR) studies.

These junction oligos were characterized structurally (by electrophoretic migration on polyacrylamide gels and NMR) and assessed for their interactions with DNA structure-specific binding proteins.

Slipped Junctions Form Distinct Isoforms As Determined by Electrophoretic Analysis. The self-association of strands was assessed by electrophoretic migration on native polyacrylamide gels, relative to a duplex size marker (Figure 1C). Each of the oligos resolved as a major species, which in all cases migrated slower than expected (>20 bp) relative to the migration of duplexes composed of the same number of nucleotides (Figure 1C). This suggested that each oligo had self-associated to form a non-B-DNA structure.[7] For certain oligos, additional slower-migrating species were evident, a phenomenon previously observed for three-way junctions containing extra nucleotides.[19] This was most notable for J4/5, in which two electrophoretic species were evident (Figure 1C). Additional electrophoretic species were previously observed for G-A mismatches in some but not all sequence contexts, a phenomenon confirmed by NMR analysis to be caused by structural interconversion.[23] Others have reported various structural forms of trinucleotide repeats and suggested the impact that such variations might have upon instability.[24] The slower migration of these extra species suggested that multiple structural conformations may be formed by each oligo; when denatured with formamide prior to native electrophoresis or when run on denaturing gels, each oligo ran as a single species (Figure S3 of the Supporting Information). Notably, gel purification of each of the J4/5 species resulted in a mixture of the two species, suggesting that they may be interconverting.

Interconversion of Fully Paired and Extra-Nucleotide Junction Isomers. Using NMR, we tested the possibility that a given sequence could form multiple junction conformations, as suggested by the distinct electrophoretic species. The same DNA sequences used above were used for the NMR studies, with an overview of the NMR experiments given in Table S2 of the Supporting Information. We first analyzed oligo J4/5, which yielded the greatest proportion of an alternate electrophoretic species (Figure 1C). For each of the junction-forming sequences, two of the stems of the predicted slipped junctions are each capped with stabilizing H2-type CTTG loops for which characteristic resonances are well established in NMR spectra.[25] For J4/5, the TOCSY spectrum shows T Me−T H6 correlations characteristic of two CTTG loops (Figure 2A and Table S3A and supplementary text section V of the Supporting Information), supporting the formation of three-way junctions with a CTG slip-out. In addition, the cross-peak at 7.18 ppm/1.38 ppm in Figure 2A corresponds to the typical position for a T in GTG element in a B-helix; the cross-peak can be assigned to T_{38} because only one GTG element is present in the J4/5 sequence. The single cross-peak for T_{38} also shows that the branch arm containing T_{38} in J4 and J5 is well formed with a single local conformation. The presence of multiple conformations of J4/5 is indicated by two key observations. First,

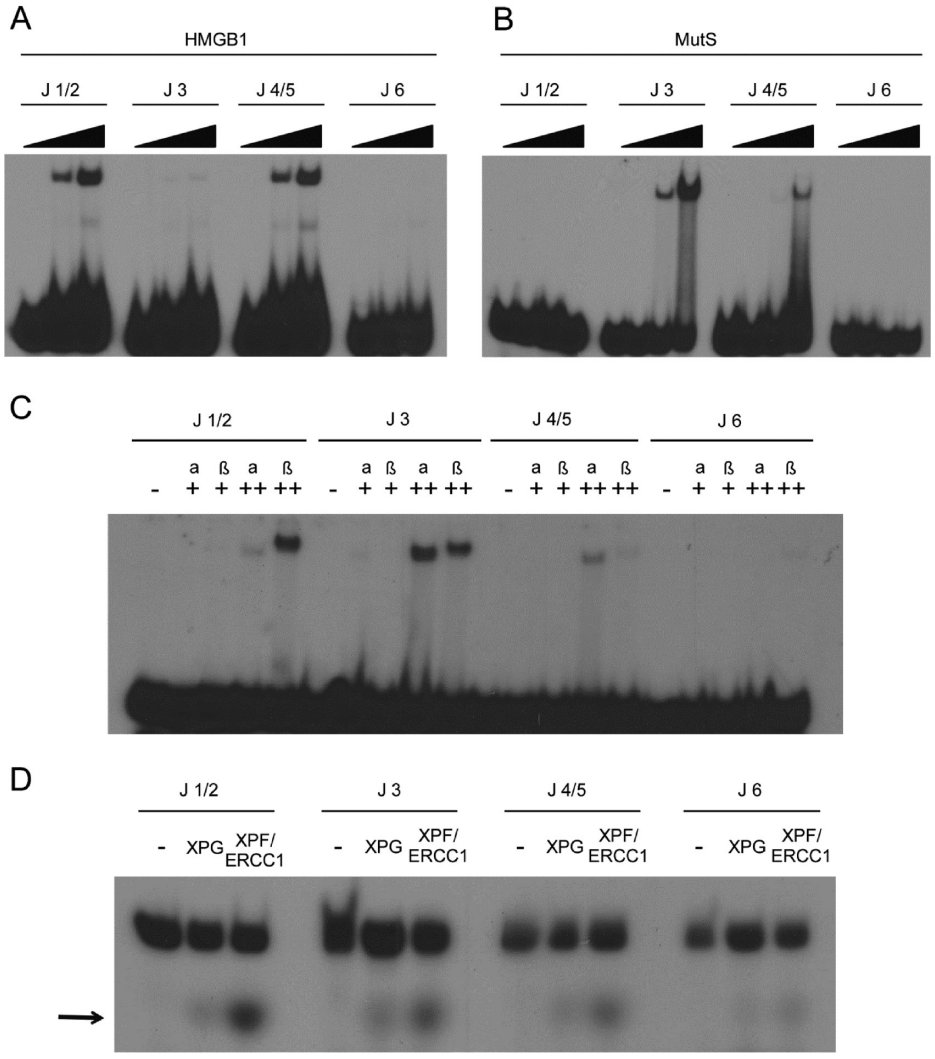


Figure 3. Electrophoretic mobility shift assays with [γ-^{32}P]ATP-end-labeled junction substrates J1/2, J3, J4/5, and J6. (A) Binding of purified human HMGB1 (0, 0.1, 1, and 10 pmol) to 100 fmol of each [γ-^{32}P]ATP-end-labeled junction substrate. (B) Binding of purified bacterial MutS (0, 0.1, 1, and 10 pmol) to 100 fmol of each [γ-^{32}P]ATP-end-labeled junction substrate. (C) Binding of purified human MutSα and MutSβ [0 (−), 250 (+), and 750 fmol (++)] to 50 fmol of each [γ-^{32}P]ATP-end-labeled junction substrate. (D) Cleavage of [γ-^{32}P]ATP end-labeled junction substrates J1/2, J3, J4/5, and J6 by human XPG and ERCC1-XPF nucleases. XPG and ERCC1-XPF proteins (100 fmol) were incubated with 100 fmol of each [γ-^{32}P]ATP-end-labeled junction substrate as indicated. EMSA and cleavage products were run on a 6% polyacrylamide gel that was dried and exposed to X-ray film. Arrows indicate the cleavage product for each junction substrate. Cleavage site mapping is shown in Figure S8 of the Supporting Information.

multiple conformations follow from the T Me−T H6 region of the TOCSY spectrum (Figure 2A, 1.6−1.9 ppm/7.2−7.6 ppm), where 14 resonances are seen for the five nonloop T's within CTG or GTC units (T_{18}, T_{21}, T_{29}, T_{32}, and T_{35}). Second, four imino A:T resonances are observed (Figure 2B,C and Table S3A of the Supporting Information). Resonances 2−4 are from A:T base pairs sandwiched between two GC base pairs. Resonance 4 can be assigned to T_{38} because it shows, together with one of its contacted G imino resonances, a cross-peak to a Me proton resonating at 1.38 ppm, a chemical shift corresponding to that of a T in a GTG sequence (Figure 2B). Most importantly, resonance 3 shows an exchange cross-peak with resonance 2 (Figure 2C; the encircled cross peaks just below and above diagonal in the lower left, labeled "exchange"), indicating multiple conformations. Moreover, imino resonance 3 consists of more than one imino resonance and involves multiple conformations, which is evident from the eight related Me contacts of the imino of the flanking C:G base pairs (Figure 2B). Also, more than 15 resonances are observed for the 13 cytosine residues in the sequence, indicating multiple conformations (Figure S5A,B of the Supporting Information). Hence, oligo J4/5 folds into multiple junction conformations that interconvert between each other (summarized in Figure 2D). The multiple NMR resonances show that the interconversion between conformations is in the slow-exchange regime; that is, the lifetime of the conformational states is in the millisecond range or longer. The presence of an exchange cross-peak in the two-dimensional (2D) NOESY indicates, given a mixing time of 300 ms, a lifetime of the conformational states not much longer than 0.5 s (Figure 2B,C). These lifetimes are similar to that of Hoogsteen base pairs in canonical duplex DNA.[26] Moreover, this observation is consistent with the time scale required for breaking one or two base pairs (e.g., the conversion from J4 to J5 involves breaking two base pairs at the junction). Our findings indicate that these junction con-

dx.doi.org/10.1021/bi301369b | *Biochemistry* 2013, 52, 773−785

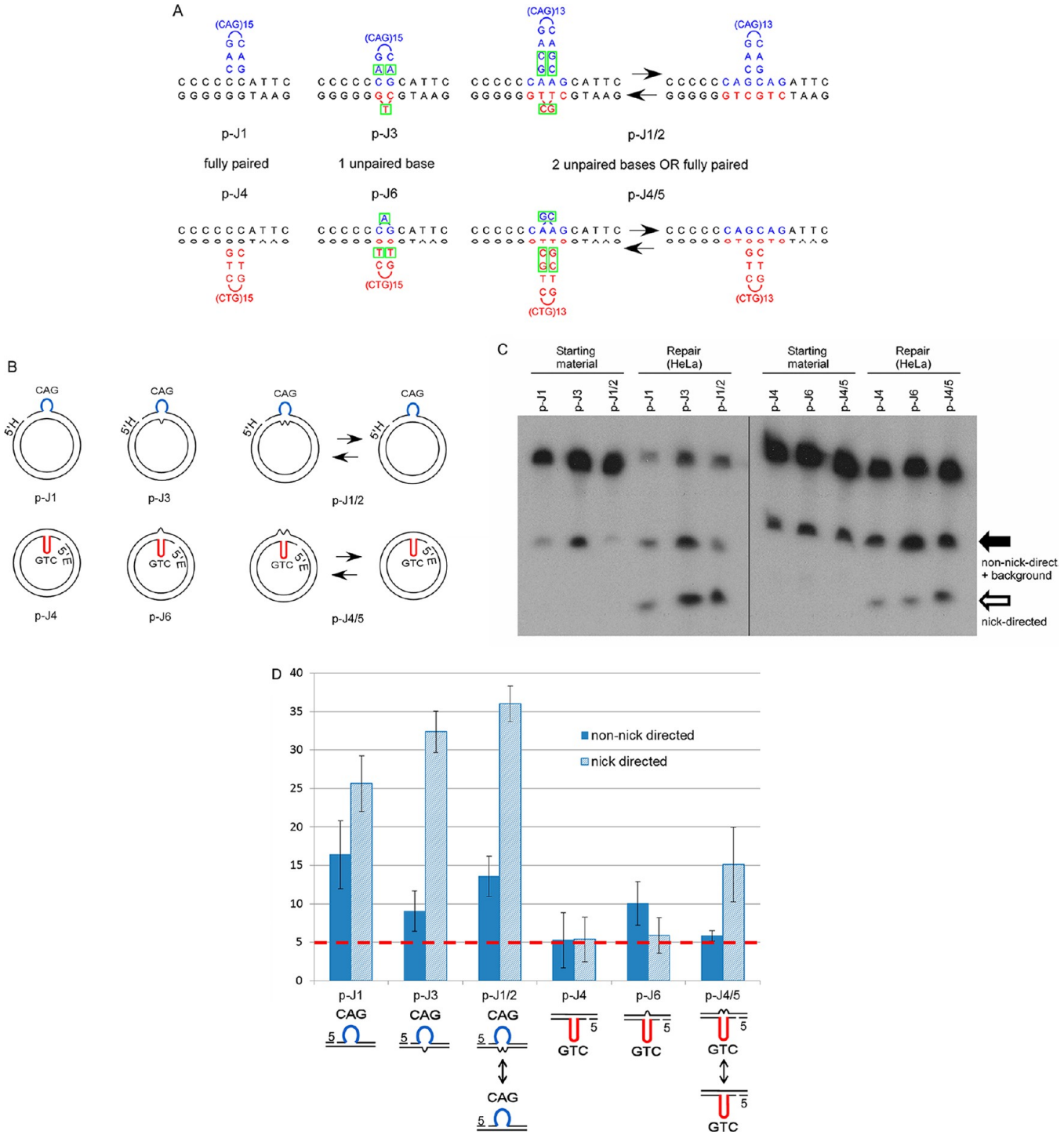


Figure 4. Repair of junction structures. (A) Junction sequences of the circular DNA repair substrates. Slip-outs are defined by both the nucleotides in the slip-out stem and the $(CNG)_n$ noted above the slip-out. Where applicable, only one possible unpaired nucleotide form is shown; other nucleotides with the potential to be unpaired are shown in green boxes. (B) DNA repair substrates modeling possible junction structures. Circular heteroduplexes with slipped $(CNG)_{17}$·$(CNG)_{0/1/2}$ repeats modeling intermediates of expansions with nicks in the slipped CAG or CTG strand. Substrates are named on the basis of the junctions (see Figure 1) that they model: p-J1 is a circular substrate modeling junction J1. (C) Southern blot analysis of repair of circular substrates using HeLa extracts. "Starting material" indicates unprocessed heteroduplexes, with the background band. The background band seen at varying levels in the starting material is the remaining double-stranded DNA from the heteroduplexing reaction mixture and cannot be completely eliminated; however, the amount of background is irrelevant as the starting material was subtracted from the repair efficiencies for quantification (as previously described[14,15]). (D) Circular substrate repair efficiencies (corrected for starting material). Nick-directed repair is shown as hatched bars, and non-nick-directed repair is shown as solid bars; error bars represent the standard deviation. Efficiency values are based on at least three replicates. The dashed red line indicates typical levels of non-nick-directed repair, as seen for a G-T mismatch.[36] Nick-directed repair: p-J1 vs p-J3, $p < 0.01$; p-J1 vs p-J1/2, $p < 0.01$; p-J3 vs p-J1/2, $p < 0.05$; p-J4/5 vs p-J4, $p < 0.05$; p-J4/5 vs p-J6, $p < 0.05$. Non-nick-directed repair: p-J1 vs p-J3, $p < 0.05$; p-J3 vs p-J1/2, $p < 0.05$; p-J6 vs p-J4/5, $p < 0.05$. p values determined by t test ($n \geq 3$).

dx.doi.org/10.1021/bi301369b | *Biochemistry* 2013, 52, 773−785

formations are in thermodynamic equilibrium rather than being trapped kinetically.

The J1/2 sequence assumed the predicted slipped-junction conformations of a CAG slip-out, which could interconvert between a two-unpaired nucleotide junction, J2 (G-C nucleotides in the CTG strand opposite the CAG slip-out, or within the CAG slip-out), and a fully paired junction, J1 (Figure 2E, NMR detailed in Figure S6A−C, Table S3B, and section VI of the Supporting Information). In addition, the NMR spectra indicate the presence of two other conformations (Figure 2E). Thus, both the fully paired and two-unpaired nucleotide slipped-junction variants can form and interconvert in J1/2 and J4/5.

The slipped-junction structure of J3 was proposed to have the potential to have a single unpaired nucleotide either opposite or within the CAG slip-out. The NMR spectra reveal the presence of multiple conformations (Figure 2F, NMR detailed in Figure S7, Table S3C, and section VII of the Supporting Information). The NMR spectra suggest a conformation with a bulged-out T_{32} in the CTG strand opposite the CAG slip-out and are consistent with the presence of an A_8:A_{19} mispair. The NMR spectra also indicate the presence of at least two other conformations, in which either A_8 or A_{19} of the A_8:A_{19} pair is bulged out while the remaining A base pairs with T_{32} (Figure 2F, bottom two schematics). The presence of the single-unpaired nucleotide slipped-junction form supports the concept that these can form at slip-outs of CTG or CAG repeats.

In summary, the NMR data show that all the sequences form three-way junctions having multiple conformations, where isoforms include the predicted base pairings or unpairings and interconversions.

Slipped-Junction Isoforms Are Differentially Recognized by DNA Repair Proteins. Structure-specific DNA binding proteins can sensitively detect perturbations in DNA conformation. To further confirm the variations between related slipped-DNA junctions and to demonstrate that not only the slip-out but also its three-way junction can affect protein recognition, we assessed whether a series of DNA sequence-independent structure-specific proteins could distinguish between junction isoforms.

HMGB1. The high-molecular weight protein HMGB1 binds DNA in a sequence-independent manner, preferring DNA fragments with preexisting bends such as those at cisplatin lesions, Holliday recombination intermediates, and microsatellite repeats.[27−29] Binding of HMGB1 to the slipped-junction models revealed distinct shifted products, and in some cases multiple shifts, likely the result of complexes with variable protein:DNA ratios and/or protein-induced bending or unwinding of the nucleic acid structure and/or the result of weak binding equilibrium, as previously observed[28,30] (Figure 3A). HMGB1 showed preferential binding for junction substrates J1/2 and J4/5, junctions that can interchange between fully paired and two-unpaired base conformations. There was minimal binding of HMGB1 to J3 at the highest protein concentration tested and no detectable interaction with J6 (Figure 3A). These results suggested that HMGB1 can differentially bind to the model junctions.

Mismatch Repair Proteins MutS, hMutSα, and hMutSβ. The bacterial MutS mismatch repair protein initiates repair via recognition and binding to DNA base−base mismatches and DNA heteroduplexes with extrahelical bases. Bacterial MutS selectively bound to substrates J3 and J4/5 as distinct shifts as well as smears at higher protein concentrations, likely the result of complexes with variable protein:DNA ratios and/or protein-induced bending or unwinding of the nucleic acid structure and/or the result of weak binding equilibrium, as previously observed[31] (Figure 3B). The MutS-bound junctions contain one or zero/two unpaired bases with a CAG or CTG slip-out, respectively. MutS did not bind to the J1/2 or J6 oligos. The differential interaction of MutS with the slipped-junction isoforms supported their being in conformationally distinct forms.

MSH2, a human homologue of the bacterial MutS protein, forms heterodimeric complexes with MSH6 and MSH3, constituting the hMutSα and hMutSβ complexes, respectively. Both MutSα and MutSβ can bind to base−base mismatches and insertion−deletion loops with as many as 5−24 excess nucleotides (ref 15 and references cited therein). Binding of human MutSα or MutSβ to the slipped-junction models revealed distinct shifted products (Figure 3C). Notably, hMutSα and hMutSβ selectively bound certain slipped-junction isomers, indicating a binding preference. hMutSα bound to J3 and only weakly to J4/5 and J1/2, but not to J6 (Figure 3C). hMutSβ selectively bound the CAG junction conformers J1/2 and J3, but not the CTG conformers J4/5 and J6 (Figure 3C). These results support a binding preference of hMutSβ for CAG slip-out junctions over CTG slip-out junctions, and a preference for specific junction conformations. Taken together, the differential recognition of the slipped-junction substrates by several DNA structure-specific repair proteins confirms their structural variations revealed by gel electrophoresis and NMR.

Nucleotide Excision Repair Proteins XPG and ERCC1-XPF. The structure-specific nucleotide excision repair nucleases XPG and ERCC1-XPF recognize DNA substrates that contain single-stranded DNA−double-stranded DNA junctions, such as bubble substrates and hairpins, to initiate repair through coordinated 5′ and 3′ cleavage of the lesion by ERCC1-XPF and XPG, respectively.[32] We tested the ability of human XPG and ERCC1-XPF proteins to cleave the slipped junctions (Figure 3D). We observed differential cleavage of each radiolabeled junction substrate by ERCC1-XPF, as evidenced by varying amounts of a faster-migrating product (Figure 3D). Cleavage sites of ERCC1-XPF were mapped on oligos J1/2, J3, and J4/5 using denaturing gels (Figure S8 of the Supporting Information). Interestingly, in each case, cleavage occurred on the slipped-out repeat arm, with varying locations relative to the junction point. The differential cleavage efficiency and varying scission locations confirm that there are structural variations between the model junctions.

Because several DNA structure-specific proteins with different biological functions show variation in binding preference for the slipped junctions, this supports the NMR findings that the junctions assume distinct conformations.

Slipped-Junction Isoforms Are Differentially Repaired by Human Cell Extracts. One might expect that because different DNA slip-outs are repaired with varied efficiencies, distinct slipped-junction isoforms may differentially influence repair outcome. To test this, circular DNA repair substrates were prepared with a slip-out of up to 17 CTG or CAG repeats, with either zero, one, or two unpaired nucleotides at the junction (Figure 4A,B). These were made by heteroduplexing single-stranded circular DNAs containing a segment of the DM1 locus with a related linearized double-stranded plasmid containing a greater number of repeats (causing a slip-out to form on one strand upon hybridization). Substrates were

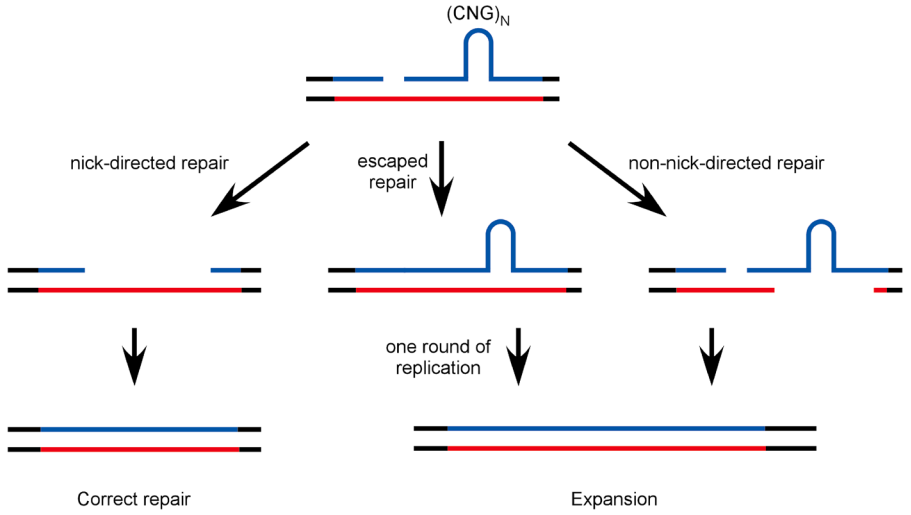


Figure 5. Nick-directed vs non-nick-directed repair. Correct repair is always nick-directed, involving excision of the excess repeats and gap filling using the continuous strand as a template; non-nick-directed repair leads to retention of the excess repeats by using the nicked slipped strand as the template for gap filling, leading to an expansion mutation product.

designed to have limited base pairing conformations at the slip-out junction: circular substrates p-J1 and p-J4 were designed to assume only the fully paired junctions, J1 and J4 (Figure 1A), as there were no repeats in the complementary strand; substrates p-J3 and p-J6 could assume any of the one-unpaired nucleotide junction forms; substrates p-J1/2 and p-J4/5 could form either the fully paired or any of the two-unpaired nucleotide forms, where these could potentially interconvert (see Figures 1A and 4A). It is likely that the slip-out would extrude in the middle of the complementary repeat tract.[11] Following the initial denaturation, the repeat tracts would form intrastrand hairpins; when the flanking nonrepetitive sequences formed base pairs, a pseudocruciform would form that would then reanneal into a single central slip-out. Additionally, it is likely that p-J1/2 and p-J4/5 would more readily assume the forms with two unpaired nucleotides, because three-way junctions with unpaired nucleotides are more stable than those that are fully paired,[19] but absolute proof of this would require more experimental support. All circular substrates contained a nick 5′ of the slip-out and modeled expansion intermediates (the nick was on the slipped-out strand). The limited conversion of the constrained circular slipped DNAs to other isoforms was experimentally confirmed.[11,13] Circular substrates p-J1/2, p-J3, p-J4/5, and p-J6 were structural equivalents of oligo junctions J1/2, J3, J4/5, and J6, respectively, used above for electrophoretic analysis, NMR, and protein−DNA interactions. The hairpins and three-way junctions formed by oligonucleotides containing CTG or CAG repeats have been experimentally confirmed to be similar to the junctions formed in plasmids.[7,13,14] Our findings parallel the oligo-to-plasmid studies of Holliday junctions and base−base mismatches.[23,33,34]

Repair results for the various circular substrates modeling slipped junctions by HeLa cell extracts are shown in Figure 4C and quantified in Figure 4D (p values are found in the figure legend). Typically, the correct repair of base−base mismatches, nonrepetitive insertion−deletion loops, and repeat slip-outs is nick-directed, where the lesion is excised from the nicked strand while the continuous strand is used as the template by DNA polymerases during repair.[14,15,35] For the expansion substrates used here, correct (nick-directed) repair involves excision of the excess repeats in the slip-out, yielding the shorter duplex repair products (hollow arrow in Figure 4C). Failed attempts at correct repair (escaped repair) can lead to expansion mutations through heteroduplex retention (Figure 5). Consistent with our previous observations,[14] the correct repair efficiency was considerably higher for plasmids with a CAG slip-out than those with a CTG slip-out [in Figure 4D, compare hatched bars for CAG slip-outs (p-J1, p-J3, and p-J1/2) with CTG slip-outs (p-J4, p-J6, and p-J4/5)]. Interestingly, we observed that for both CAG and CTG slip-outs, junction structure had a significant effect on correct repair. For CAG slip-out substrates, the correct repair efficiency differed among all three substrates with efficiency increasing as follows: p-J1 < p-J3 < p-J1/2. Similarly, the correct repair efficiency was altered for CTG slip-out substrates: p-J4 and p-J6 < p-J4/5. Thus, the least constrained junctions (p-J1/2 and p-J4/5) are repaired with a significantly higher efficiency than the most constrained fully paired junctions (p-J1 and p-J4). The differences in repair efficiencies for CAG versus CTG slip-outs likely reflect differences in the structures that they form (unpaired random coil vs intrastrand hairpin), as previously demonstrated.[11] If the decreased level of intrastrand pairing of CAG slip-outs enhances their repair versus that of CTG slip-outs, it is possible that decreasing the level of pairing at the junction (by going from fully paired to up to two unpaired nucleotides) is what permits the improved repair of certain junction structures. Thus, we have shown that both slip-out sequence and slip-out junction conformation can alter repair outcome and, consequently, determine the potential for a slip-out to escape repair and be mutagenic.

As expected, correct repair of these necessarily long slip-outs was independent of mismatch repair proteins (Figure S9 of the Supporting Information), yielding high levels of repair similar to what is seen for long slip-outs with unconstrained junction conformations.[14] Although the pattern of correct repair mediated by MMR-deficient LoVo cell extracts (for LoVo, p-J1 = p-J3 < p-J1/2; t test of p-J1 or p-J3 vs p-J1/2, $p < 0.05$) was not identical to that seen with MMR-proficient HeLa cell extracts (for HeLa, p-J1 < p-J3 < p-J1/2), we still observed considerably greater repair for p-J1/2 compared to the other junction substrates for LoVo; thus, MMR may not play a key role in differentiating between junction conformations. If MMR

dx.doi.org/10.1021/bi301369b | *Biochemistry* 2013, 52, 773−785

was the main determinant of junction discrimination, we would expect that p-J1 = p-J3 = p-J1/2; however, we cannot exclude the possibility that MMR is contributing to some degree in junction discrimination. Because junction conformation affects binding of MutSβ, it would be interesting to see if repair of short slip-outs (which requires mismatch repair proteins[15]) is affected by structural changes at the junction; however, it is not possible to constrain the junction to specific conformations with short slip-outs in our system.

Unexpectedly, certain junction conformations had a significantly increased level of misdirected repair to the non-nicked strand (see the solid arrow in Figure 4C; in Figure 4D, filled bars; the dashed red line indicate levels of non-nick-directed repair for a G-T mismatch). In contrast to correct nick-directed repair, this aberrant repair uses the nicked strand as the template for repair and yields fully paired products that incorporate the excess repeats from the slip-out, essentially causing an expansion mutation (Figure 5). Other studies using this system to analyze repair of base−base mismatches, random sequence heteroduplexes, and slip-outs do not report levels of non-nick-directed repair because they are much lower than the levels of correct repair in the same substrates.[36,37] For example, the G-T mismatch yields only 3−5% non-nick-directed repair compared to 35−40% nick-directed correct repair.[36] However, we observed significant differences in non-nick-directed repair efficiencies with various junctions, with the level of non-nick-directed repair being higher for some junctions than the level of correct repair of other junctions [in Figure 4D, compare p-J1 (filled bar) with p-J4 (hatched bar)]. This aberrant repair was most striking for the CAG slip-outs p-J1 (16.4%) and p-J1/2 (13.6%) (Figure 4D). Generally, the non-nick-directed repair efficiency was greater for CAG slip-outs than for CTG slip-outs, similar to the trend observed for correct repair efficiency. Altering the junction conformation had opposite effects for CAG slip-outs versus CTG slip-outs. For CAG slip-outs, the largest amounts of aberrant non-nick-directed repair arose from the fully paired (p-J1) and the zero/two-unpaired nucleotide junctions (p-J1/2), while the smallest amounts arose from the one-unpaired nucleotide junction (p-J3). In contrast, for CTG slip-outs, the one-unpaired nucleotide junction (p-J6) yielded the most aberrant non-nick-directed repair, while the other junctions showed negligible levels. Because the levels of both correct and aberrant repair differ between junctions, the junction conformation at a slip-out could determine whether repair is protective or mutagenic (Figure 5). While it is clear that junction conformation can determine the levels of both correct nick-directed and aberrant non-nick-directed repair, the mechanism by which this occurs has yet to be elucidated.

DISCUSSION

The structural features of DNA lesions are crucial to their proper recognition and processing during repair, recombination, or mutation. Slipped DNAs necessarily contain three-way DNA junctions and slip-outs that typically form hairpins. Hairpins of CAG repeats are distinct from CTG repeats, having unique hairpin tips,[38] and CAG slip-outs are more unpaired.[11,39] Recent studies have confirmed that CAG slip-outs within a three-way junction adopt more of an open loop than a hairpin.[40−42] Perturbed base pairing at junctions with CAG slip-outs has been reported,[40−42] with little detail about the junction conformation. Unpaired nucleotides at three-way DNA junctions of nonrepetitive sequences can dramatically affect the global structure of the junction and permit increased biophysical stability and flexibility.[20,25,43,44] It has been shown that for fully paired DNA three-way junctions, coaxial stacking of two helices is not possible without disrupting the base pairing, so fully paired junctions form only nonstacked conformations.[45−47] Introduction of a single-stranded region at the branch point provides additional flexibility, allowing for formation of alternative stacked conformations.[20,25,45] Moreover, the sequences surrounding the junction play an important role in structure formation and can be used to predict conformational preferences using the pyrimidine rule and the loop rule.[20] In most cases, one of the stacked conformers is predominant.[20,25] Recently, these rules were successfully employed to design DNA nanostructures,[48] where the rules were extended from junctions containing a quasi tetraloop to junctions with a quasi triloop at the branch point. Alteration of the three-way junction conformation (with unpaired bases) can be induced by various factors, including modifying the sequence around the junction, macromolecular crowding, hydration state, protein interaction, and small molecule interaction.[11,48,49] Here we reveal the structural heterogeneity at slipped junctions formed by complementary CTG/CAG strands, where unpaired nucleotides can arise in either strand opposite any arm of the junction.

Repair of various base−base mismatches and random sequence heteroduplexes is sensitive to DNA structure.[17,23,50−53] We previously reported that this is also true for repetitive slipped DNAs, where CAG slip-outs are repaired with greater efficiency than CTG slip-outs.[14] An extra layer of complexity has been added considering that slipped DNAs can assume multiple conformations at the junction, including having unpaired nucleotides. Because either a slip-out or unpaired nucleotides could be considered DNA damage, either one could predominate as the lesion demanding repair. If the wrong strand is used as the template for repair, excess slipped-out repeats can be retained leading to expansion mutations (Figure 5).

Typically, it is thought that for slipped-DNAs, the nicked strand will be repaired and the continuous strand will serve as the template for repair. Loop-directed repair has been reported for large random sequence heteroduplexes, where loops are removed regardless of nick location.[51−53] We previously observed that loop-directed repair does not occur for slipped CTG/CAG repeats,[14] even following ligation of the nicks. The non-nick-directed repair that we observe is not the same as loop repair, as it involves incorporation of the excess repeats. In the expansion intermediates used here, the excess slipped-out repeats are on the nicked strand, meaning that nick-directed and loop-directed repair would both give a similar repair outcome: excision of the slip-out and resynthesis using the shorter continuous strand as a template. Despite this, we were able to observe repair of both strands, both nick-directed and non-nick-directed repair, depending on the junction conformation. While the aberrant repair was non-nick-directed, it is likely not nick-independent (there must be a nick located somewhere on the circular repair substrate) because we have previously observed in this system that repair will not occur without a nick.[14] Because the levels of aberrant non-nick-directed repair varied between slipped junctions that differed in only minor ways, we suggest that this aberrant repair depends upon junction structure. That such aberrant repair has not been reported for other nonrepeat heteroduplex DNA substrates[17,50−53] is consistent with the strong propensity of CAG/CTG repeats to mutate with an expansion bias, a

phenomenon very likely to be linked to the unique structures these repeats assume.[2,3,7,9]

Structural features of unpaired DNAs have previously been reported to affect repair outcome. The repair efficiency of base−base mismatches is sensitive to sequence context and increases with increasing G and C content in the neighboring sequence.[23,54,55] A G-A mismatch in a GC-rich context exhibited a regular intrahelical structure and a high repair efficiency, while the same mismatch in an AT-rich context assumed predominantly a looped-out conformation and was poorly repaired.[23] Similarly, base−base mismatches in the context of a mutation hot spot of the *K-ras* gene assume a dynamic equilibrium between an intrahelical G-A pair and a looped structure that is poorly repaired, permitting high levels of mutation.[34] Our observations of the effects of slipped-DNA junction structural dynamics on repair outcomes (protective vs mutagenic) further extend the effects of base−base mismatch structural variations.[23,34]

Several studies using slipped repeats did not detect substantial levels of repair to the non-nicked strand.[14,15,35] A potential source for the increased level of non-nick-directed repair observed herein is the manner in which the substrates were designed to restrain the circular substrates to more limited structural conformations (Figure 4A). The ability to form and interconvert between isoforms (zero, one, and two unpaired nucleotides opposite the slip-out) could permit sufficient competition for repair outcomes, thereby weakening the ability to detect products of non-nick-directed repair. Forcing the junction into more constrained conformations could be what permitted us to detect the different repair outcomes.

Differential protein binding to the junction isoforms (Figure 3) may explain some of the observed repair differences. Alteration of DNA pairing in heteroduplexes may be induced upon interaction with DNA repair proteins,[31,56,57] an action thought to affect subsequent nuclease cleavage.[31,58,59] Changes in DNA pairing dependent upon binding could affect both correct and aberrant repair. As we demonstrate herein, a decreased level of pairing seems to improve correct repair (e.g., CAG slip-outs that may be random coils are repaired better than CTG hairpin slip-outs). Additionally, if an endonuclease introduced a nick into the non-nicked strand because of the conformation after protein binding, this could cause an increase in the level of aberrant repair. Identifying proteins that may be involved in the repair of slipped trinucleotide repeats is an active area of investigation.

We have demonstrated the presence of multiple conformations for trinucleotide repeat junctions in a short DNA oligo model. These same aspects would show up in larger plasmid model systems of triplet repeats; however, because of their size, tertiary interactions may occur that could influence (stabilize, for instance) certain types of multiple conformations or could enhance allowed conformational space. In vivo, interactions with different proteins may further influence conformational space (narrow it or expand it). The remarkable observation is that a correlation exists between repair in plasmids and the degree of multiplicity in conformational space. This correlation is highly suggestive of a physically causal connection, but definitive proof would require direct observation of multiple conformations in the plasmid or even in vivo, neither of which is technically feasible. Nevertheless, we provide good evidence to assume that multiple conformations at the junction site in oligos would also play a role in larger plasmid systems and perhaps in vivo, but with the caveats stated above. Other elegant model systems of repeat structures suggest structural dynamics but share the same limitations.[60]

In summary, the slipped-junction conformation can determine whether expansion intermediates will be correctly repaired (excising the excess repeats), escape repair (retain the excess repeats as a heteroduplex), or be processed to expansions (incorporating the excess repeats). Thus, the potential for lesions in both DNA strands at a slipped junction can alter repair outcome, leading to correction or mutation. This improves our understanding of how repair attempts can lead to mutagenesis.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Tables S1−S3 and Figures S1−S9. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Program of Genetics and Genome Biology, The Hospital for Sick Children, TMDT 15-312, 101 College St., Toronto, ON M5G 1L7, Canada. Telephone: (416) 813-8256. Fax: (416) 813-4931. E-mail: cepearson.sickkids@gmail.com.

Present Address

⊥Sanofi Pasteur, Steeles Avenue West, North York, ON M2R 3T4, Canada.

Author Contributions

M.M.S. prepared human MMR proteins and repair substrates and performed all repair assays. K.R. performed binding and cleavage assays. M.M.S. and K.R. ran electrophoretic analyses. K.N.E. and M.K. prepared oligos. B.W., F.H.T.N., R.L.E.G.A., M.T., and S.S.W. performed all NMR analysis. O.D.S. prepared NER proteins. C.E.P. and S.S.W. designed and supervised the project. All authors wrote the manuscript.

Funding

This work was supported by the Canadian Institutes of Health Research (CIHR grants MOP-57813 & MOP-97896) (C.E.P.), the National Institutes of Health (Grant GM080454 to O.D.S.), and the Chemistry section of the Dutch Science Organization (S.S.W.). Both M.M.S. and K.R. were supported by studentships from The Hospital for Sick Children Research Training Competition and the CIHR Collaborative Graduate Training Program in Molecular Medicine. M.M.S. was also supported by an Ontario Graduate Scholarship.

Notes

The authors declare no competing interests.

■ REFERENCES

(1) Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R. S., Zborowska, E., Kinzler, K. W., Vogelstein, B., et al. (1995) Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. *Science 268*, 1336−1338.

(2) Pearson, C. E., Nichol Edamura, K., and Cleary, J. D. (2005) Repeat instability: Mechanisms of dynamic mutations. *Nat. Rev. Genet. 6*, 729−742.

(3) Mirkin, S. M. (2007) Expandable DNA repeats and human disease. *Nature 447*, 932−940.

(4) Lopez Castel, A., Cleary, J. D., and Pearson, C. E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol. 11*, 165−170.

(5) Seriola, A., Spits, C., Simard, J. P., Hilven, P., Haentjens, P., Pearson, C. E., and Sermon, K. (2011) Huntington's and myotonic dystrophy hESCs: Down-regulated trinucleotide repeat instability and

mismatch repair machinery expression upon differentiation. *Hum. Mol. Genet. 20*, 176−185.
(6) Lopez Castel, A., Nakamori, M., Tome, S., Chitayat, D., Gourdon, G., Thornton, C. A., and Pearson, C. E. (2011) Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum. Mol. Genet. 20*, 1−15.
(7) Pearson, C. E., and Sinden, R. R. (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry 35*, 5041−5053.
(8) Pearson, C. E., Ewel, A., Acharya, S., Fishel, R. A., and Sinden, R. R. (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum. Mol. Genet. 6*, 1117−1123.
(9) Pearson, C. E., Wang, Y. H., Griffith, J. D., and Sinden, R. R. (1998) Structural analysis of slipped-strand DNA (S-DNA) formed in $(CTG)_n \cdot (CAG)_n$ repeats from the myotonic dystrophy locus. *Nucleic Acids Res. 26*, 816−823.
(10) Pearson, C. E., Eichler, E. E., Lorenzetti, D., Kramer, S. F., Zoghbi, H. Y., Nelson, D. L., and Sinden, R. R. (1998) Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. *Biochemistry 37*, 2701−2708.
(11) Pearson, C. E., Tam, M., Wang, Y. H., Montgomery, S. E., Dar, A. C., Cleary, J. D., and Nichol, K. (2002) Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: Slipped-out repeats and slip-out junctions. *Nucleic Acids Res. 30*, 4534−4547.
(12) Marcadier, J. L., and Pearson, C. E. (2003) Fidelity of primate cell repair of a double-strand break within a (CTG)·(CAG) tract. Effect of slipped DNA structures. *J. Biol. Chem. 278*, 33848−33856.
(13) Tam, M., Erin Montgomery, S., Kekis, M., Stollar, B. D., Price, G. B., and Pearson, C. E. (2003) Slipped (CTG)·(CAG) repeats of the myotonic dystrophy locus: Surface probing with anti-DNA antibodies. *J. Mol. Biol. 332*, 585−600.
(14) Panigrahi, G. B., Lau, R., Montgomery, S. E., Leonard, M. R., and Pearson, C. E. (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat. Struct. Mol. Biol. 12*, 654−662.
(15) Panigrahi, G. B., Slean, M. M., Simard, J. P., Gileadi, O., and Pearson, C. E. (2010) Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. *Proc. Natl. Acad. Sci. U.S.A. 107*, 12593−12598.
(16) Fishel, R., Ewel, A., Lee, S., Lescoe, M. K., and Griffith, J. (1994) Binding of mismatched microsatellite DNA sequences by the human MSH2 protein. *Science 266*, 1403−1405.
(17) Umar, A., Boyer, J. C., and Kunkel, T. A. (1994) DNA loop repair by human cell extracts. *Science 266*, 814−816.
(18) Lang, W. H., Coats, J. E., Majka, J., Hura, G. L., Lin, Y., Rasnik, I., and McMurray, C. T. (2011) Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops. *Proc. Natl. Acad. Sci. U.S.A. 108*, 837−844.
(19) Leontis, N. B., Kwok, W., and Newman, J. S. (1991) Stability and structure of three-way DNA junctions containing unpaired nucleotides. *Nucleic Acids Res. 19*, 759−766.
(20) van Buuren, B. N., Overmars, F. J., Ippel, J. H., Altona, C., and Wijmenga, S. S. (2000) Solution structure of a DNA three-way junction containing two unpaired thymidine bases. Identification of sequence features that decide conformer selection. *J. Mol. Biol. 304*, 371−383.
(21) Stern, S., Powers, T., Changchien, L. M., and Noller, H. F. (1989) RNA-protein interactions in 30S ribosomal subunits: Folding and function of 16S rRNA. *Science 244*, 783−790.
(22) Staresincic, L., Fagbemi, A. F., Enzlin, J. H., Gourdin, A. M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S. G., Vermeulen, W., and Scharer, O. D. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J. 28*, 1111−1120.
(23) Fazakerley, G. V., Quignard, E., Woisard, A., Guschlbauer, W., van der Marel, G. A., van Boom, J. H., Jones, M., and Radman, M. (1986) Structures of mismatched base pairs in DNA and their recognition by the *Escherichia coli* mismatch repair system. *EMBO J. 5*, 3697−3703.
(24) Volker, J., Klump, H. H., and Breslauer, K. J. (2008) DNA energy landscapes via calorimetric detection of microstate ensembles of metastable macrostates and triplet repeat diseases. *Proc. Natl. Acad. Sci. U.S.A. 105*, 18326−18330.
(25) Wu, B., Girard, F., van Buuren, B., Schleucher, J., Tessari, M., and Wijmenga, S. (2004) Global structure of a DNA three-way junction by solution NMR: Towards prediction of 3H fold. *Nucleic Acids Res. 32*, 3228−3239.
(26) Nikolova, E. N., Kim, E., Wise, A. A., O'Brien, P. J., Andricioaei, I., and Al-Hashimi, H. M. (2011) Transient Hoogsteen base pairs in canonical duplex DNA. *Nature 470*, 498−502.
(27) Pil, P. M., and Lippard, S. J. (1992) Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science 256*, 234−237.
(28) Bianchi, M. E., Beltrame, M., and Paonessa, G. (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. *Science 243*, 1056−1059.
(29) Gaillard, C., and Strauss, F. (1994) Association of poly(CA)·poly(TG) DNA fragments into four-stranded complexes bound by HMG1 and 2. *Science 264*, 433−436.
(30) Stemmer, C., Schwander, A., Bauw, G., Fojan, P., and Grasser, K. D. (2002) Protein kinase CK2 differentially phosphorylates maize chromosomal high mobility group B (HMGB) proteins modulating their stability and DNA interactions. *J. Biol. Chem. 277*, 1092−1098.
(31) Surtees, J. A., and Alani, E. (2006) Mismatch repair factor MSH2-MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination. *J. Mol. Biol. 360*, 523−536.
(32) Fagbemi, A. F., Orelli, B., and Scharer, O. D. (2011) Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair 10*, 722−729.
(33) Lilley, D. M. (2008) Analysis of branched nucleic acid structure using comparative gel electrophoresis. *Q. Rev. Biophys. 41*, 1−39.
(34) Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., and Fazakerley, G. V. (1991) Solution structure of an oncogenic DNA duplex containing a G·A mismatch. *Biochemistry 30*, 5449−5458.
(35) Hou, C., Chan, N. L., Gu, L., and Li, G. M. (2009) Incision-dependent and error-free repair of $(CAG)_n/(CTG)_n$ hairpins in human cell extracts. *Nat. Struct. Mol. Biol. 16*, 869−875.
(36) Holmes, J., Jr., Clark, S., and Modrich, P. (1990) Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. *Proc. Natl. Acad. Sci. U.S.A. 87*, 5837−5841.
(37) Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) Heteroduplex repair in extracts of human HeLa cells. *J. Biol. Chem. 266*, 3744−3751.
(38) Hartenstine, M. J., Goodman, M. F., and Petruska, J. (2000) Base stacking and even/odd behavior of hairpin loops in DNA triplet repeat slippage and expansion with DNA polymerase. *J. Biol. Chem. 275*, 18382−18390.
(39) Petruska, J., Hartenstine, M. J., and Goodman, M. F. (1998) Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *J. Biol. Chem. 273*, 5204−5210.
(40) Degtyareva, N. N., Reddish, M. J., Sengupta, B., and Petty, J. T. (2009) Structural studies of a trinucleotide repeat sequence using 2-aminopurine. *Biochemistry 48*, 2340−2346.
(41) Degtyareva, N. N., Barber, C. A., Sengupta, B., and Petty, J. T. (2010) Context dependence of trinucleotide repeat structures. *Biochemistry 49*, 3024−3030.
(42) Degtyareva, N. N., Barber, C. A., Reddish, M. J., and Petty, J. T. (2011) Sequence length dictates repeated CAG folding in three-way junctions. *Biochemistry 50*, 458−465.
(43) Shlyakhtenko, L. S., Appella, E., Harrington, R. E., Kutyavin, I., and Lyubchenko, Y. L. (1994) Structure of three-way DNA junctions. 2. Effects of extra bases and mismatches. *J. Biomol. Struct. Dyn. 12*, 131−143.

(44) Overmars, F. J., Pikkemaat, J. A., van den Elst, H., van Boom, J. H., and Altona, C. (1996) NMR studies of DNA three-way junctions containing two unpaired thymidine bases: The influence of the sequence at the junction on the stability of the stacking conformers. *J. Mol. Biol. 255*, 702−713.
(45) Lilley, D. M. (2000) Structures of helical junctions in nucleic acids. *Q. Rev. Biophys. 33*, 109−159.
(46) Duckett, D. R., and Lilley, D. M. (1990) The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking. *EMBO J. 9*, 1659−1664.
(47) Sabir, T., Toulmin, A., Ma, L., Jones, A. C., McGlynn, P., Schroder, G. F., and Magennis, S. W. (2012) Branchpoint expansion in a fully complementary three-way DNA junction. *J. Am. Chem. Soc. 134*, 6280−6285.
(48) Seemann, I. T., Singh, V., Azarkh, M., Drescher, M., and Hartig, J. S. (2011) Small-molecule-triggered manipulation of DNA three-way junctions. *J. Am. Chem. Soc. 133*, 4706−4709.
(49) Muhuri, S., Mimura, K., Miyoshi, D., and Sugimoto, N. (2009) Stabilization of three-way junctions of DNA under molecular crowding conditions. *J. Am. Chem. Soc. 131*, 9268−9280.
(50) Kramer, B., Kramer, W., and Fritz, H. J. (1984) Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. *Cell 38*, 879−887.
(51) McCulloch, S. D., Gu, L., and Li, G. M. (2003) Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *J. Biol. Chem. 278*, 3891−3896.
(52) Corrette-Bennett, S. E., Mohlman, N. L., Rosado, Z., Miret, J. J., Hess, P. M., Parker, B. O., and Lahue, R. S. (2001) Efficient repair of large DNA loops in *Saccharomyces cerevisiae*. *Nucleic Acids Res. 29*, 4134−4143.
(53) Littman, S. J., Fang, W. H., and Modrich, P. (1999) Repair of large insertion/deletion heterologies in human nuclear extracts is directed by a 5′ single-strand break and is independent of the mismatch repair system. *J. Biol. Chem. 274*, 7474−7481.
(54) Jones, M., Wagner, R., and Radman, M. (1987) Repair of a mismatch is influenced by the base composition of the surrounding nucleotide sequence. *Genetics 115*, 605−610.
(55) Sanchez, A. M., Volk, D. E., Gorenstein, D. G., and Lloyd, R. S. (2003) Initiation of repair of A/G mismatches is modulated by sequence context. *DNA Repair 2*, 863−878.
(56) Drotschmann, K., Yang, W., Brownewell, F. E., Kool, E. T., and Kunkel, T. A. (2001) Asymmetric recognition of DNA local distortion. Structure-based functional studies of eukaryotic Msh2-Msh6. *J. Biol. Chem. 276*, 46225−46229.
(57) Dowen, J. M., Putnam, C. D., and Kolodner, R. D. (2010) Functional studies and homology modeling of Msh2-Msh3 predict that mispair recognition involves DNA bending and strand separation. *Mol. Cell. Biol. 30*, 3321−3328.
(58) Lyndaker, A. M., and Alani, E. (2009) A tale of tails: Insights into the coordination of 3′ end processing during homologous recombination. *BioEssays 31*, 315−321.
(59) Gupta, S., Gellert, M., and Yang, W. (2011) Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat. Struct. Mol. Biol. 19*, 72−78.
(60) Volker, J., Gindikin, V., Klump, H. H., Plum, G. E., and Breslauer, K. J. (2012) Energy landscapes of dynamic ensembles of rolling triplet repeat bulge loops: Implications for DNA expansion associated with disease states. *J. Am. Chem. Soc. 134*, 6033−6044.
(61) Altona, C. (1996) Classification of nucleic acid junctions. *J. Mol. Biol. 263*, 568−581.