

Dynamic roles for G4 DNA in the biology of eukaryotic cells

Nancy Maizels

Recent advances have made a persuasive case for the existence of G4 DNA in living cells, but what—if any—are its functions? Experiments have established how G4 DNA may contribute to the biology of eukaryotic cells, and genomic analysis has identified new ways in which the potential to form G4 DNA may influence gene regulation and genomic stability. This Perspective highlights those advances and identifies some key open questions.

G-quartets and G4 DNA

Synthetic G-rich oligonucleotides were shown to form four-stranded structures almost 20 years ago¹. These structures are variously referred to as G4 DNA (used here), G-quadruplexes and G-tetraplexes. G4 DNA is stabilized by G-quartets, planar arrays of four guanines in which each guanine pairs with two neighbors by Hoogsteen bonding² (Fig. 1a). In G4 DNA, the G-quartets define a very large core, comparable in size to a heme moiety, which is encircled by four phosphodiester backbones (Fig. 1b).

Both G-quartets and G4 DNA were discovered accidentally. When a concentrated solution of GMP unexpectedly formed a gel overnight, difraction analysis revealed the G-quartet². When G-rich oligonucleotides repeatedly refused to migrate with predicted mobility on a denaturing gel, chemical probing established that these molecules had spontaneously formed G-quartets¹. The readiness with which G-quartets and G4 DNA form *in vitro* suggested, of course, that these structures can also form *in vivo*, and in the last several years intracellular formation of G4 DNA has been documented^{3–5}.

G-rich nucleic acids have remarkable potential to form diverse conformations *in vitro*, in which G-quartets stabilize inter- or intramolecular interactions, between parallel or antiparallel DNA strands, with bases in *syn* or *anti* configurations^{6,7}. G4 DNA forms in sequences containing at least four guanine runs, each of which is at least 3 nt in length (G₃). Longer G-runs promote formation of more stable structures. G4 DNA forms spontaneously in physiological salts. It is stabilized by K⁺ at concentrations (10 mM) far below the levels typical of mammal cells (120 mM). As might be expected, intramolecular G4 DNA formation occurs in a concentration-independent manner, whereas high DNA concentrations accelerate intermolecular G4 DNA formation. G4 DNA is extremely stable and resists thermal denaturation, although it

is readily denatured upon treatment with alkali. The stability of this structure *in vitro* implies that mechanisms must exist to remove G4 DNA formed *in vivo*.

G4 DNA potential and the genome

Genomic analyses can identify sequences in living cells with G4 DNA formation potential (G4P) using algorithms based on analysis of structures formed by synthetic oligonucleotides *in vitro*. Algorithms typically search for the presence of at least four runs of at least three guanines within a window 40–100 nucleotides (nt) in length, sufficient for G4 DNA formation within synthetic oligonucleotides. It is important to emphasize that algorithms can establish only the potential to form this conformation, because most genomic DNA is maintained as a Watson-Crick duplex, in which G-C pairing prevents formation of G4 DNA. G4 DNA is a dynamic structure, and its formation depends on denaturation of the duplex, as occurs during replication (for example, see Fig. 2), transcription or recombination.

Three key repetitive functional domains are characterized by high G4P: the telomeres, the ribosomal DNA (rDNA) and, in mammals, the immunoglobulin heavy-chain switch regions. High G4P also characterizes repetitive G-rich microsatellites, including some of the most unstable⁸. In the human genome, more than 300,000 distinct sites have potential to form G4 DNA⁹. Several kinds of evidence support the functional importance of these sites. First, specific sequence motifs are evident at sites with high G4P, particularly within the loops separating G-runs¹⁰. Second, among single-copy genes, G4P correlates with gene function as defined by gene ontology terms¹¹. Notably, G4P is low in tumor-suppressor genes, which maintain genomic stability and are commonly haploinsufficient; but it is high in proto-oncogenes, which promote cell proliferation¹¹. G4 DNA formation can correlate with genomic instability, and the low G4P of tumor-suppressor genes may result in selection that protects genes in this class from instability. The high G4P of proto-oncogenes is more difficult to explain, but there is an intriguing possibility that it may reflect shared regulation, as discussed below.

Regions with potential to form G4 DNA have been identified within the promoters of several proto-oncogenes, including *c-MYC* and *c-KIT*^{12,13}. Regions with potential to form G4 DNA have also been reported to characterize bacterial promoters¹⁴, although the algorithm used in that analysis was relatively relaxed and may have overestimated potential for structuring. It has been proposed that G4 DNA formed upon transcription could be a target for a small molecule that would prevent gene expression upon binding. A number of compounds that recognize G4 DNAs have been developed, some selective for specific structures and others with broader specificity. However, none has as yet proven therapeutically useful. As many genomic regions have potential to form G4 DNA, off-target effects might be a serious concern.

Nancy Maizels is in the Departments of Immunology and Biochemistry, University of Washington School of Medicine, 1959 N.E. Pacific Street, Seattle, Washington 98195-7650, USA.

e-mail: maizels@u.washington.edu

Published online 5 December 2006; doi:10.1038/nsmb1171

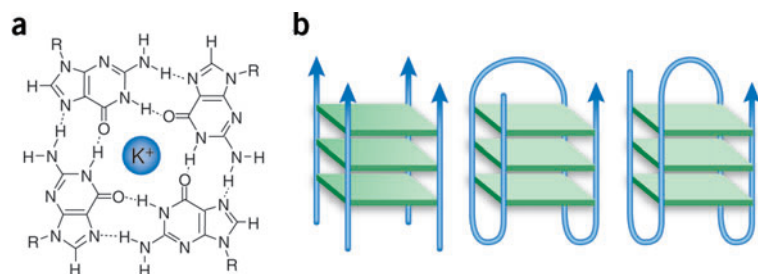


Figure 1 G-quartets and G4 DNA. (a) G-quartet, a planar ring of four guanines, with a monovalent cation (usually K^+ , as shown) in the central channel. (b) G4 DNA. Strands may be parallel (left), antiparallel (center) or mixed (right). The monovalent cation occupies the region between the planes defined by G-quartets. The actual potential for conformational diversity is enormous even within short synthetic oligonucleotides, encompassing stable parallel and antiparallel structures, *syn* and *anti* base configurations, interactions between nonadjacent G-runs and dimerization of separate modules^{6,7}.

G4P and regulation of gene expression

Factors associated with RNA processing, including hnRNP D, CBF, the UPI derivative of hnRNP A1, and nucleolin, bind G4 DNA through their conserved RNA-recognition motif and RNA-binding domain (RRM/RBD) regions^{15–18}. This raises the possibility that G4 conformations in RNA transcripts may enhance binding by factors that promote RNA processing and thus stimulate splicing or transport of specific transcripts. Shared regulation of some fraction of G-rich genes provides one plausible explanation for the correlation between high G4P and gene function¹¹ and for the presence of specific sequence motifs at sites with high G4P¹⁰.

The rDNA provides an extreme example of a repetitive gene family that is G rich and rapidly transcribed: electron microscopic analysis has shown that in actively dividing cells, the RNA polymerase complexes are very closely spaced, with distances as short as 67 nt separating polymerase complexes¹⁹. rDNA transcription and rRNA biogenesis occur in the nucleolus. Nucleolin, the major protein in the mesozoan nucleolus, contains RRM/RBD domains and binds G4 DNA with high affinity, leading to the suggestion that nucleolin may interact with G4 DNA formed on the nontemplate strand to prevent renaturation of the duplex and render the template strand available for multiple rounds of transcription¹⁶.

G4 DNA processing mechanisms

G4 DNA is very stable once formed, and cells have evolved mechanisms to process this structure. Because of its distinctive conformation (Fig. 1), G4 DNA is not a substrate for most enzymes active on single-stranded or duplex DNA, including even such avid nucleases as S1, snake-venom phosphodiesterase and DNase 1. As described below, some proteins active at the telomeres resolve G4 DNA by trapping the single-stranded telomeric repeat, thus shifting the equilibrium away from G4 DNA. Proteins shown to unwind G4 DNA include, most notably, helicases of the RecQ family, which are also necessary for maintenance of G-rich genomic regions^{20–27}. For example, G4 DNA is unwound by *Escherichia coli* RecQ, by *Saccharomyces cerevisiae* Sgs1p and by the human RecQ-family helicases BLM²⁰ and WRN²⁵. The signature RecQ conserved (RQC) domain, unique to RecQ-family helicases, binds G4 DNA with high affinity and may promote intracellular recognition of this conformation²⁸. In human cells lacking WRN helicase, telomeric sequence is lost owing to impaired replication of the telomeric G-rich strand²⁵, an adverse consequence of persistence of G4 DNA during replication (Fig. 2). WRN helicase is deficient in Werner syndrome, a human genetic disease characterized by premature aging as well as development of malignancies. This creates a tantalizing connection between persistence of G4 DNA and genomic instability related to aging.

Members of the DExH helicase family may also be important in G4 DNA unwinding. The nematode *dog-1* gene, which encodes a protein in this family, is essential to stability of very long G-runs²⁹; and the related murine gene *Rtel* is essential for telomere maintenance³⁰. Activities of those factors have not been analyzed *in vitro*, but the related human enzyme DHX36 has been identified as an ATP-dependent G4 DNA resolvase³¹.

G4 DNA is cleaved by the mammalian nuclease QQN1 (ref. 32) and *S. cerevisiae* Kem1p³³, although the native substrate of Kem1p may be RNA rather than DNA. The conserved nuclease Mre11p has been reported to cleave G4 DNA³⁴, but its activity on this substrate is quite weak. G4 DNA is a target for

binding and DNA synapsis by MutS α (the MSH2-MSH6 heterodimer), which also recognizes base mismatches and promotes their repair³⁵.

Many of the factors that actively unwind or cleave G4 DNA recognize this substrate with high affinity (in the nanomolar range) and considerable selectivity. Remarkably, considering the many structural analyses of proteins bound to duplex and single-stranded DNA, structural analysis has not yet provided a molecular picture of a protein bound to G4 DNA.

G4 DNA at telomeres

Telomeres in essentially all eukaryotes terminate with a 3' single-stranded G-rich overhang, or tail. Intermolecular G4 DNA may promote telomere-telomere interactions in ciliate macronuclei³; and intramolecular G4 DNA could also form within telomeric tails or stabilize t-loops, the lariat-like structures at ends of chromosomes (Fig. 3). Telomerase, the enzyme that maintains telomere length, uses the 3' overhang as a primer for addition of new DNA. Many telomerases cannot extend G4 DNA *in vitro*^{36,37}, so formation of G4 DNA may regulate extension by telomerase.

Ciliates first provided evidence that telomeres may adopt an unusual structure³⁸, and they currently provide the best-documented example of G4 DNA formation at telomeres *in vivo*. During specific stages of

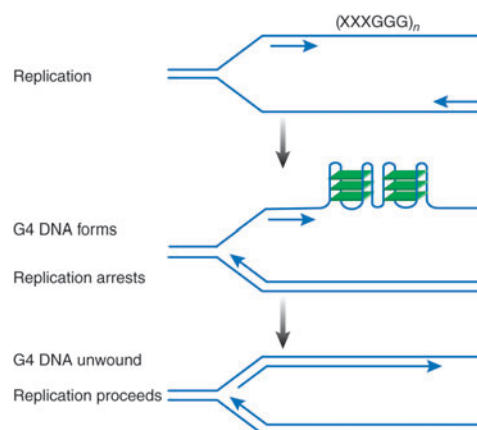


Figure 2 Formation of G4 DNA upon replication. Duplex DNA denatures to permit replication of a G-rich region (top); G4 DNA forms, causing replication to arrest (center); G4 DNA unwinding—for example, by a RecQ family helicase like BLM or WRN—allows replication to proceed (bottom). An antiparallel structure is diagrammed, for simplicity.

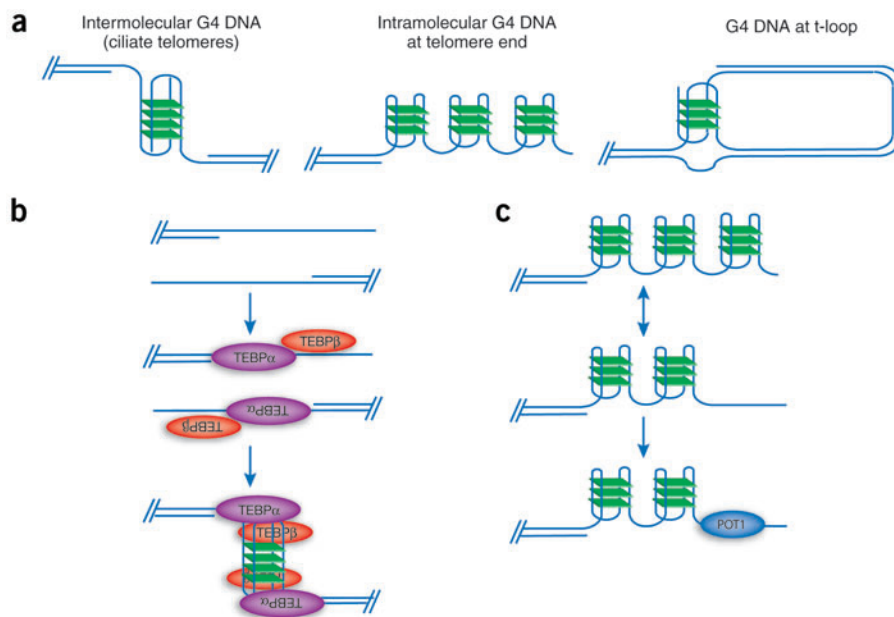


Figure 3 G4 DNA at telomeres. (a) Left, intermolecular, antiparallel G4 DNA joins two ciliate telomeres. Center, intramolecular, antiparallel G4 DNA within a vertebrate TTAGGG telomeric tail; the TTAGGG repeat can also form other structures. Right, G4 DNA could in principle stabilize the junction of the telomere tail and duplex DNA predicted to occur as part of a t-loop, the lariat structure at a chromosome end. (b) Formation of the G4 DNA–TEBP α/β complex at ciliate telomeres. Single-stranded regions (top) are bound by TEBP α , which allows TEBP β to load (center); this in turn promotes formation of intermolecular G4 DNA (bottom). (c) Resolution of telomeric G4 DNA by POT1. Rapid interconversion of G4 DNA (top) and single-stranded conformations (center) permit POT1 to load (bottom), making the telomere end accessible to telomerase.

interacts directly with the unstructured G-rich telomeric repeat, enabling TEBP β to load; and TEBP β binding promotes G4 DNA formation (Fig. 3b). Cell cycle-dependent phosphorylation of TEBP β promotes dissociation, so interaction with TEBP α/β provides a mechanism to regulate the presence of G4 DNA at telomere ends. Nonetheless, whether telomerase activity at a telomere is regulated by G4 structures is still not clear, particularly as ciliate telomerases can extend intermolecular G4 DNA *in vitro*³⁷.

In many eukaryotes, from fission yeast through humans, the conserved protection of telomeres-1 (POT1) protein is bound to telomere ends³⁹. POT1 is essential, and in its absence telomere ends initiate a DNA-damage response and aberrant recombination^{40,41}. Human telomerase cannot extend telomeric G4 DNA *in vitro*. But, like its homolog, TEBP α , POT1 contains OB folds that promote specific recognition of telomeric sequence⁴². By binding and trapping the single-stranded conformation (Fig. 3c), POT1 can shift the equilibrium away from G4 DNA and thereby promote extension by telomerase³⁶. POT1 also stimulates RecQ-family helicase activity on telomeric substrates, which would further promote a single-stranded conformation at telomeres *in vivo*⁴³.

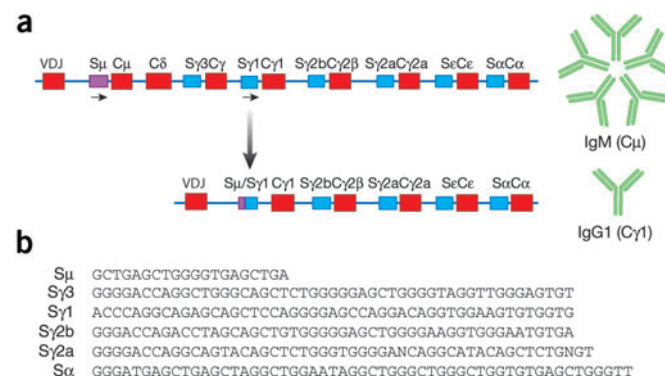
Conserved factors unrelated to POT1 also promote unstructuring of telomeric G4 DNA in mammalian cells. Two heterogeneous nuclear ribonucleoproteins, hnRNP A1 (and its derivative, UP1) and hnRNP D, bind with considerable specificity to the mammalian telomeric repeat and resolve G4 DNA conformations^{18,44–47}. hnRNP A1 has been shown to associate with telomeres *in vivo*, and its depletion inhibits telomere extension *in vitro*, suggesting that its function may be redundant to that of POT1 *in vivo*¹⁸. In contrast to POT1 and TEBP α/β , which interact with single-stranded DNA via OB folds, hnRNP A1 and hnRNP D bind via conserved RRM/RBD domains.

Results discussed above show that G4 DNA can form at telomeres, but they do not prove that this structure has an essential role in telomere biology. In principle, that question could be addressed experimentally by asking whether telomere maintenance is compromised if telomeres

the life cycle, these unicellular eukaryotes generate a macronucleus containing millions of gene-sized chromosomes, each with two telomeres. Staining cells of the ciliate *Stylonychia lemnae* with antibodies to G4 DNA localized antiparallel G4 DNA specifically to the macronucleus³, where it may promote chromosome-chromosome interactions characteristic of ciliate macronuclei (Fig. 3a, left). Antibody staining did not provide a quantitative measure of the fraction of telomeres at which G4 DNA was present, and antibody binding could, in principle, even promote G4 DNA formation. However, RNA interference-mediated silencing of either the α or β subunit of the heterodimeric telomere end-binding protein complex TEBP α/β was recently shown to cause loss of telomeric DNA and cell death, while at the same time abolishing staining by antibodies to G4 DNA⁵. This supports the idea that the antibody recognizes a specific structure that depends upon TEBP α/β for formation, stability or both.

TEBP α/β is a conserved factor containing oligonucleotide-oligosaccharide-binding (OB) folds that promote sequence-specific recognition of single-stranded telomeric repeats. *In vitro*, TEBP α

Figure 4 G-rich S regions are targets of immunoglobulin heavy-chain class switch recombination. (a) Class switch recombination at the murine immunoglobulin heavy-chain locus. The expressed heavy-chain locus is shown before (above) and after (below) switching from μ to γ 1. This changes the expressed antibody class from IgM to IgG1 (right). Transcription of the targeted S regions activates recombination (horizontal arrows denote transcription). VDJ, rearranged heavy chain variable region; S, switch region; C, constant region. (b) G-rich sequences at the murine S regions. S regions consist of degenerate repeats 2–10 kb in length. Consensus sequences for some S regions are shown.



cannot form G4 DNA. Telomeric sequences can be readily changed by mutation of the telomerase RNA template; but in a living cell, an entire constellation of proteins interacts at the telomere ends, and success of such an experimental strategy would require covariation of all factors that bind telomeric DNA with sequence specificity.

G-loops and recombinogenic G4 DNA

In mammalian B cells, regulated G4 DNA formation promotes recombination that modulates antibody structure and the immune response. At the expressed immunoglobulin heavy-chain gene of an activated B cell, class switch recombination occurs to join a new constant region to the variable region, promoting deletion of many kilobases (kb) of chromosomal DNA (Fig. 4a). Recombination junctions map within switch regions (S regions), G-rich, degenerate, repetitive DNA sequences 2–10 kb in length (Fig. 4b). The S regions targeted for recombination are transcribed to initiate recombination. Transcription of G-rich templates such as the S regions, either *in vitro* or intracellularly, causes characteristic large loops to form, which are hundreds of base pairs in length and readily imaged by EM^{4,48} (Fig. 5a,b). These loops, called ‘G-loops’, contain a cotranscriptional RNA-DNA hybrid on the C-rich template strand and G4 DNA interspersed with single-stranded regions on the G-rich strand^{4,48}. Class switch recombination is impaired by deficiencies in either subunit of MutS α , and this may reflect the ability of MutS α to interact with G4 DNA in G-loops (Fig. 5c) and to promote synapsis between G-loops³⁵.

Cotranscriptional hybrids and genomic instability

G-loops form very efficiently in G-rich templates: loops are evident in more than 50% of substrates after transcription *in vitro* and in 15% of plasmids recovered after transcription in *E. coli*⁴. Deficiency in RecQ helicase and RNase H increases the fraction of recovered molecules containing G-loops⁴, showing both that the loops are produced intracellularly and that highly conserved mechanisms exist for elimination of alternative nucleic acid structures.

G-loops contain a cotranscriptional RNA-DNA hybrid^{4,48}. The persistence of such hybrids in living cells can result in genomic instability, and conserved factors associated with RNA processing (THO/TREX and ASF/SF2) normally function to prevent formation of such hybrids^{49,50}. It is not yet known whether cotranscriptional hybrid formation correlates closely with the potential to form G4 DNA or also occurs in regions in which guanines are interspersed rather than in runs. In the former case, regions of high G4P would be predicted to show instability in somatic cells and also on an evolutionary timescale.

Perspective: where do we go from here?

G4 DNA is a dynamic structure, making it a particular challenge to establish its biological functions. Nonetheless, the past few years have seen real progress in overcoming skepticism about whether G4 DNA forms at all in a living cell, establishing where in the genome it can form and identifying proteins that recognize this conformation and may put it to use or remove it. Genetics and genomics together can now build on this foundation to provide a systematic picture of the biological functions of G4 DNA. Modulating protein–G4 DNA recognition by mutation of domains or sites necessary for this interaction should reveal the importance of G4 DNA to regulatory pathways and genomic mechanisms for maintenance of G-rich regions. Genomic comparisons can establish whether potential for G4 DNA formation is a conserved property of specific genes, how this potential correlates with genomic stability or instability and how it may contribute to genomic evolution.

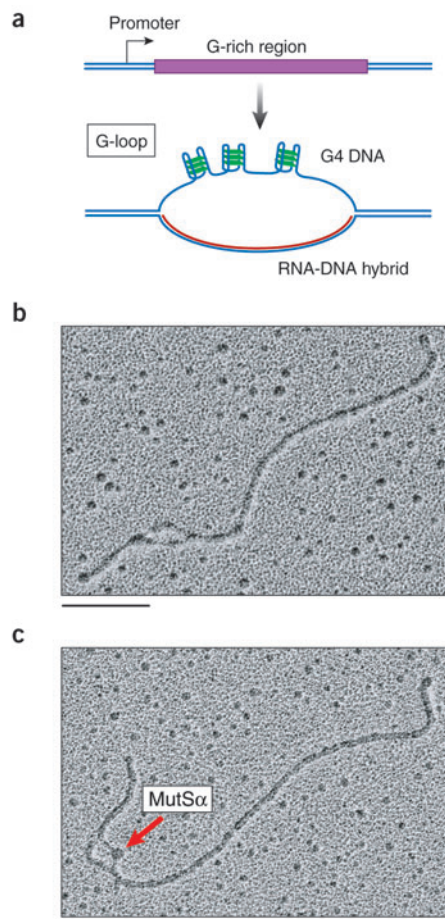


Figure 5 G-loop formed in a transcribed G-rich region is bound by MutS α . (a) Transcription of a G-rich region in duplex DNA (above) results in formation of a G-loop (below), which contains a cotranscriptional RNA-DNA hybrid on the template strand and G4 DNA interspersed with single-stranded regions on the G-rich strand. Red, RNA strand; green, G-rich region of DNA. (b) Transmission electron micrograph of G-loop in transcribed G-rich region³⁵. (c) Transmission electron micrograph of G4 DNA in G-loop bound by MutS α ³⁵ (arrow). Bars in b and c, 200 nm. Panels b and c reproduced courtesy of *Current Biology*.

ACKNOWLEDGMENTS

I thank members of my laboratory for provocative discussions, and the US National Institutes of Health for supporting our research on G4 DNA (R01 GM65988) and switch recombination (R01 GM39799).

COMPETING INTERESTS STATEMENT

The author declares that she has no competing financial interests.

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366 (1988).
2. Gellert, M., Lipsett, M.N. & Davies, D.R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA* **48**, 2013–2018 (1962).
3. Schaffitzel, C. *et al.* In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Styloynchia lemnae* macronuclei. *Proc. Natl. Acad. Sci. USA* **98**, 8572–8577 (2001).
4. Duquette, M.L., Handa, P., Vincent, J.A., Taylor, A.F. & Maizels, N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618–1629 (2004).
5. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. & Lipps, H.J. Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*.

- Nat. Struct. Mol. Biol.* **12**, 847–854 (2005).
6. Lee, J.Y., Okumus, B., Kim, D.S. & Ha, T. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. USA* **102**, 18938–18943 (2005).
 7. Phan, A.T., Kuryavii, V. & Patel, D.J. DNA architecture: from G to Z. *Curr. Opin. Struct. Biol.* **16**, 288–298 (2006).
 8. Wong, Z., Wilson, V., Patel, I., Povey, S. & Jeffreys, A.J. Characterization of a panel of highly variable minisatellites cloned from human DNA. *Ann. Hum. Genet.* **51**, 269–288 (1987).
 9. Huppert, J.L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908–2916 (2005).
 10. Todd, A.K., Johnston, M. & Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **33**, 2901–2907 (2005).
 11. Eddy, J. & Maizels, N. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* **34**, 3887–3896 (2006).
 12. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. & Hurley, L.H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA* **99**, 11593–11598 (2002).
 13. Rankin, S. *et al.* Putative DNA quadruplex formation within the human c-kit oncogene. *J. Am. Chem. Soc.* **127**, 10584–10589 (2005).
 14. Rawal, P. *et al.* Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res.* **16**, 644–655 (2006).
 15. Dempsey, L.A., Sun, H., Hanakahi, L.A. & Maizels, N. G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D: a role for G-G pairing in immunoglobulin switch recombination. *J. Biol. Chem.* **274**, 1066–1071 (1999).
 16. Hanakahi, L.A., Sun, H. & Maizels, N. High affinity interactions of nucleolin with G-G-paired rDNA. *J. Biol. Chem.* **274**, 15908–15912 (1999).
 17. Khateb, S., Weisman-Shomer, P., Hershco, I., Loeb, L.A. & Fry, M. Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homologous conserved domains in three members of the hnRNP family. *Nucleic Acids Res.* **32**, 4145–4154 (2004).
 18. Zhang, Q.S., Manche, L., Xu, R.M. & Krainer, A.R. hnRNP A1 associates with telomere ends and stimulates telomerase activity. *RNA* **12**, 1116–1128 (2006).
 19. French, S.L., Osheim, Y.N., Cioci, F., Nomura, M. & Beyer, A.L. In exponentially growing *Saccharomyces cerevisiae* cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Mol. Cell. Biol.* **23**, 1558–1568 (2003).
 20. Sun, H., Karow, J.K., Hickson, I.D. & Maizels, N. The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.* **273**, 27587–27592 (1998).
 21. Sun, H., Bennett, R.J. & Maizels, N. The *S. cerevisiae* Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res.* **27**, 1978–1984 (1999).
 22. Fry, M. & Loeb, L.A. Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)n. *J. Biol. Chem.* **274**, 12797–12802 (1999).
 23. Wu, X. & Maizels, N. Substrate-specific inhibition of RecQ helicase. *Nucleic Acids Res.* **29**, 1765–1771 (2001).
 24. Mohaghegh, P., Karow, J.K., Brosh, R.M. Jr., Bohr, V.A. & Hickson, I.D. The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* **29**, 2843–2849 (2001).
 25. Crabbe, L., Verdun, R.E., Haggblom, C.I. & Karlseder, J. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* **306**, 1951–1953 (2004).
 26. Chang, S. *et al.* Essential role of limiting telomeres in the pathogenesis of Werner syndrome. *Nat. Genet.* **36**, 877–882 (2004).
 27. Mandell, J.G., Goodrich, K.J., Bahler, J. & Cech, T.R. Expression of a RecQ helicase homolog affects progression through crisis in fission yeast lacking telomerase. *J. Biol. Chem.* **280**, 5249–5257 (2005).
 28. Huber, M.D., Duquette, M.L., Shiels, J.C. & Maizels, N. A conserved G4 DNA binding domain in RecQ family helicases. *J. Mol. Biol.* **358**, 1071–1080 (2006).
 29. Cheung, I., Schertzer, M., Rose, A. & Lansdorp, P.M. Disruption of dog-1 in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA. *Nat. Genet.* **31**, 405–409 (2002).
 30. Ding, H. *et al.* Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* **117**, 873–886 (2004).
 31. Vaughn, J.P. *et al.* The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *J. Biol. Chem.* **280**, 38117–38120 (2005).
 32. Sun, H., Yabuki, A. & Maizels, N. A human nuclease specific for G4 DNA. *Proc. Natl. Acad. Sci. USA* **98**, 12444–12449 (2001).
 33. Liu, Z. & Gilbert, W. The yeast *KEM1* gene encodes a nuclease specific for G4 tetraplex DNA: implication of in vivo functions for this novel DNA structure. *Cell* **77**, 1083–1092 (1994).
 34. Ghosal, G. & Muniyappa, K. *Saccharomyces cerevisiae* Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for replication of telomeric DNA. *Nucleic Acids Res.* **33**, 4692–4703 (2005).
 35. Larson, E.D., Duquette, M.L., Cummings, W.J., Streiff, R.J. & Maizels, N. MutSalpha binds to and promotes synthesis of transcriptionally activated immunoglobulin switch regions. *Curr. Biol.* **15**, 470–474 (2005).
 36. Zaug, A.J., Podell, E.R. & Cech, T.R. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. USA* **102**, 10864–10869 (2005).
 37. Oganessian, L., Moon, I.K., Bryan, T.M. & Jarstfer, M.B. Extension of G-quadruplex DNA by ciliate telomerase. *EMBO J.* **25**, 1148–1159 (2006).
 38. Lipps, H.J., Gruissem, W. & Prescott, D.M. Higher order structure in macronuclear chromatin of the hypotrichous ciliate *Oxytricha nova*. *Proc. Natl. Acad. Sci. USA* **79**, 2495–2499 (1982).
 39. Baumann, P. & Cech, T.R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171–1175 (2001).
 40. Wu, L. *et al.* Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell* **126**, 49–62 (2006).
 41. Hockemeyer, D., Daniels, J.P., Takai, H. & de Lange, T. Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres. *Cell* **126**, 63–77 (2006).
 42. Lei, M., Podell, E.R., Baumann, P. & Cech, T.R. DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* **426**, 198–203 (2003).
 43. Opresko, P.L. *et al.* POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. *J. Biol. Chem.* **280**, 32069–32080 (2005).
 44. Ishikawa, F., Matunis, M.J., Dreyfuss, G. & Cech, T.R. Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. *Mol. Cell. Biol.* **13**, 4301–4310 (1993).
 45. LaBranche, H. *et al.* Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat. Genet.* **19**, 199–202 (1998).
 46. Eversole, A. & Maizels, N. In vitro properties of the conserved mammalian protein hnRNP D suggest a role in telomere maintenance. *Mol. Cell. Biol.* **20**, 5425–5432 (2000).
 47. Enokizono, Y. *et al.* Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D. *J. Biol. Chem.* **280**, 18862–18870 (2005).
 48. Duquette, M.L., Pham, P., Goodman, M.F. & Maizels, N. AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. *Oncogene* **24**, 5791–5798 (2005).
 49. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711–721 (2003).
 50. Li, X. & Manley, J.L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122**, 365–378 (2005).