[Frontiers in Bioscience 12, 4336-4351, May 1, 2007]

Tetraplex DNA and its interacting proteins

Michael Fry

Department of Biochemistry, Rappaport Faculty of Medicine, Technion, Israel Institute of Technology, Haifa 31096, Israel

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Biological significance of tetraplex DNA
3.1. Evidence for the existence of tetraplex DNA in vivo
3.2. Proposed biological roles of tetraplex DNA.
3.2.1. Functions in the regulation of gene transcription
3.2.1.1. Tetraplex formation in the DNA control region of c-myc
3.2.1.2. Tetraplex structures in DNA control regions of muscle-specific genes
3.2.2. Roles in telomere metabolism
3.2.3. Function in DNA recombination
3.2.4. Role in immunoglobulin class switch recombination
3.2.5. Involvement in nucleotide expansion
3.2.6. Role in the assembly of HIV nucleocapsid
4. Tetraplex DNA interacting proteins
4.1. Tetraplex DNA binding and stabilizing proteins
4.2. Proteins that promote the formation of tetraplex DNA
4.3. Tetraplex DNA unwinding helicases
4.4. Tetraplex DNA destabilizing proteins
4.5. Tetraplex DNA selective nucleases
4.6. Selection and design of tetraplex DNA interacting proteins
5. Perspectives
6. Acknowledgements
7. References

1. ABSTRACT

Mounting evidence indicates that certain nucleotide sequences impose non-canonical secondary structures on DNA. The resulting variable conformations are thought to bestow on the DNA informational content additional to that encoded by its linear arrangement of bases. DNA sequences that include clusters of contiguous guanine residues readily form *in vitro* diverse types of four-stranded structures collectively named tetraplex or quadruplex DNA. Data suggest that tetraplex DNA structures are likely to be formed *in vivo* and to have roles in key biological processes such as regulation of gene transcription, maintenance of telomeres, DNA recombination and the packaging of retroviral genome. A credible argument for the existence of quadruplex DNA *in vivo* is the prevalence of numerous viral and cellular proteins that interact physically and functionally with tetrahelical DNA. Some such proteins bind selectively and tightly to tetraplex DNA, others promote the formation of DNA tetrahelices or act to unwind them, and several nucleases cleave tetraplex DNA preferentially. The protein-mediated structural transformations of quadruplex DNA and its selective nucleolytic cleavage argue strongly for transient formation of tetrahelical DNA in the cell. This review surveys tetraplex structures of DNA and their interacting proteins and appraises the evidence for their biological roles.

2. INTRODUCTION

Evidence gathered in the last few decades revealed that DNA is able to adopt a variety of secondary structures other than the canonical B-DNA configuration. A large body of experimental data indicates that different DNA sequences form *in vitro*, and probably *in vivo*, non-B-DNA structures such as bent (1), triplex (2, 3) and Z-DNA (4, 5) formations. These secondary structures may potentially add a considerable informational capacity to DNA. Multiple lines of evidence suggest that secondary formations are likely to play roles in DNA transactions such as replication, transcription and recombination. Hence, the initial conception of DNA being exclusively the holder of linearly arranged genetic information is being revisited to add alternative DNA conformations as important informational elements.

Early observations made in the 1960s established that synthetic polyguanylic acid has a propensity to form aggregates under physiological-like *in vitro* conditions of temperature, pH and salt concentrations. This initially inexplicable finding was elucidated in the late 1980s in three seminal papers that demonstrated that guanine-rich nucleic acid sequences readily produce *in vitro* a variety of four-stranded formations termed tetraplex or quadruplex structures (6-8). At the core of these tetrahelices are cyclic guanine quartets in which four guanine residues that act as

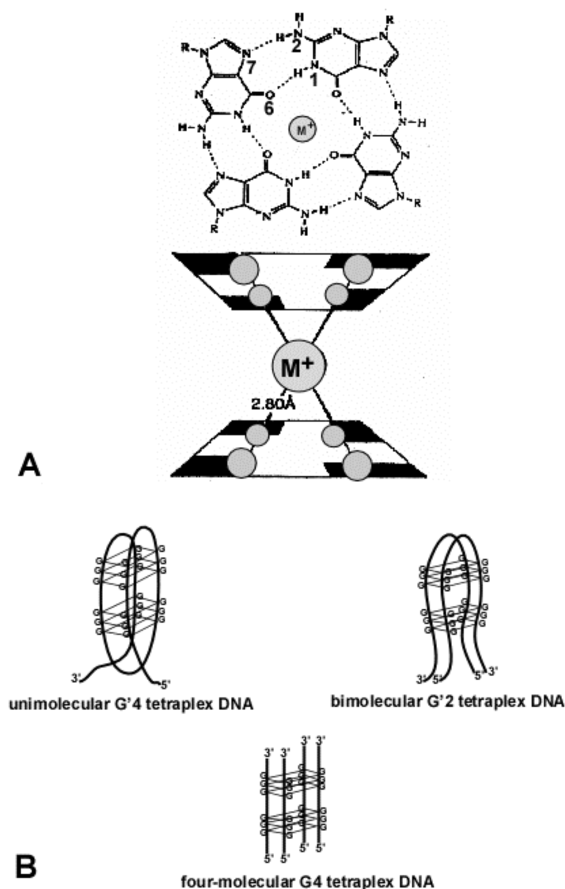

Figure 1. A. Top – scheme of a guanine quartet. Each of the four guanine residues serves as both a donor of two hydrogen bonds (groups NH1 and $NH_2$2) and an acceptor of two (groups O6 and N7). A monovalent metal ion, commonly K^+ or Na^+, is positioned below and above the center of the quartet. Bottom – schematic depiction of two stacked guanine quartets. The metal ion is located between the two quartets forming eight coordinative bonds with the carbonyl groups of the guanine residues. B. Scheme of the three major classes of quadruplex DNA. The illustrated quadruplexes are stabilized by quartets formed between two clusters of three contiguous guanine residues each. The G'4 unimolecular and G'2 bimolecular tetraplexes are shown as having antiparallel orientation and the G4 four-molecular tetraplex as a parallel-stranded tetrahelix. Notably, however, G'4 and G'2 quadruplexes can also adopt parallel-stranded polarity (15, reviewed in refs. 18, 19).

both acceptors and donors of Hoogsteen hydrogen bonds form a closed circle (Figure 1A). The presence of monovalent alkali ion, most commonly K^+ or Na^+, is essential for both the generation and the stabilization of the four-stranded structures. When intercalated between two stacked quartets, the metal ion forms coordinative bonds with the eight carbonyl groups of the guanine residues (Figure 1A). Tetraplex DNA formations can be generated by naturally occurring guanine-rich single-strands such as the overhanged telomeric DNA single-strand (9). Alternatively, quadruplex structures may be generated in guanine-rich sequences extruded from double-stranded DNA (10). A variety of tetraplex structures can be produced *in vitro* by nucleic acid sequences that include clusters of contiguous guanine residues. Four-stranded formations whose structure was confirmed by X-ray crystal diffraction and NMR analyses, (11-15), are classified into three major groups of unimolecular, bimolecular and four-molecular tetraplexes as schematically depicted in Figure 1B. In addition, different structures are further sub-grouped according to criteria such as the parallel or anti-parallel orientation of their DNA strands, their glycosidic torsion angels, the participation of non-guanine nucleotides in tetrad formation, type of the coordinating cation and the base sequence and geometry of spacer stretches that separate the guanine clusters. Detailed description of the characteristics of quadruplex DNA structural variants, which is beyond the scope of this article, is to be found in other reviews (5, 16-19).

Indirect evidence suggests that tetrahelical DNA exists *in vivo*. First, sequences derived from genomic DNA readily fold and combine *in vitro* into quadruplex structures under physiological-like conditions. Second, tetraplex forming guanine clusters are represented at statistically disproportionate pervasiveness in biologically important genomic regions such as telomeres, gene promoters and enhancers and in hyper-recombining sites (see 3.2 below). The proposed occurrence of tetraplex DNA *in vivo* also gains support by the prevalence of multiple prokaryotic and eukaryotic proteins that interact physically and functionally with tetraplex DNA. Some of the proteins bind avidly, and in some cases stabilize tetrahelical DNA. By also binding to the DNA, other proteins promote the formation of tetraplex structures. Conversely, DNA destabilizing proteins and some catalytic helicases act to preferentially unwind tetrahelical DNA. Last, nucleases were described that incise DNA or RNA selectively within or next to tetraplex domains. The prevalence and diverse activities of the tetraplex interacting proteins are highly suggestive of both the presence *in vivo* of tetrahelical regions in DNA and of their dynamic formation, dissolution and digestion. These proteins and their potential biological significance are the subject of this review article.

3. BIOLOGICAL SIGNIFICANCE OF TETRAPLEX DNA

3.1. Evidence for the existence of tetraplex DNA *in vivo*

Recently gathered data reveal high prevalence of potential tetraplex forming guanine clusters in the evolutionarily distant genomes of prokaryotes (20) and humans (21). Computer search for guanine clusters that have the potential to form tetraplex structures identified 226,157 such sites in the human genome (21). Most significantly, analysis of >61,000 open reading frames in the genomes of 18 prokaryotes detected statistically disproportionate incidence of tetraplex-forming sequences in gene promoters (20). Further, recent analysis of nearly 17,000 human genes indicated that gene function correlates with the potential to form tetrahelices. It was found that whereas some classes of genes such as tumor suppressors possess low tetraplex-forming potential others, such as proto-oncogenes have a very high likelihood of generating

Table 1. Tetraplex DNA motifs implicated in the regulation of gene transcription

Gene	Location of the tetraplex motif	Type of TETRAplex	Proposed effect on transcription	References
c-myc	-142 to -115 bp of the P1 promoter	Intramolecular	Suppression	32, 33,132
Insulin	-363 bp	Intra- and intermolecular	Activation or suppression, dependent upon bound proteins	165-167
Rb	-100 to +100 bp	Intra- and intermolecular	?	133-135
KRAS	-327 to -296	Intramolecular	Suppression	136, 137
MCK, *sMtCK*, α7 integrin	different sites along the promoter and enhancer regions	Intra- and intermolecular	Suppression	38, 39
VEGF	-85 to -50	Intramolecular	?	138
HIF-1α	-85 to -65	Intramolecular	?	139
c-kit	-160 to -140 and -109 to -87	Intramolecular	?	140, 141
Bcl-2	-58 to -19	Intramolecular (multiple forms)	?	142
β globin	-224 to -110	Intramolecular	?	168, 169

etrahelical domains (22). Despite the wide prevalence of genomic sequences that can potentially fold into tetraplex structures, a direct demonstration of their existence *in vivo* proved to be a rather difficult undertaking. Since most of the tetraplex generating sequences are relatively short and quadruplexes are likely to be transiently formed, the detection of tetrahelical regions in genomic DNA is analogous to a search for a needle in a haystack. Yet, use of quadruplex recognizing specific antibodies and dye molecules as well as direct electron microscopy provided rather convincing evidence for the presence of tetraplex DNA structures in the genomes of several organisms.

Despite their preferential binding *in vitro* to some types of tetraplex synthetic DNA, the specificity of the earliest reported anti tetraplex DNA antibodies from autoimmune mice was too low to unambiguously detect tetrahelical regions in genomic DNA (23, 24). More recently, however, anti telomeric tetraplex DNA antibodies that were selected from a human combinatorial library were reported to selectively bind parallel and antiparallel tetrahelical DNA with high affinity (K_d = 3-5 nM) (25). The fluorescently labeled antibodies stain telomeres in macronuclei of *Stylonychia lemnae* but do not mark DNA in micronuclei, suggesting that tetraplex domains exist in telomeres *in vivo*. Further, absence of immune fluorescence in replicating DNA suggests that the tetraplex structures are dissolved in the course of telomere replication (25). Use of these antibodies in subsequent studies indicated that the tetrahelical structures formed by telomeric DNA are generated upon telomere attachment to a sub-nuclear structure (26) and that the telomere end binding proteins TEBs α and β cooperate in controlling the formation of the tetrahelical DNA (27). In another line of investigation, the dye 3,6-bis(1-methyl-4-vinylpyridinium)carbazole diiodide (BMVC) was shown to possess high sensitivity and *in vitro* binding preference for the tetraplex form of the telomeric repeat sequence $d(TTAGGG)_4$ over double-stranded DNA. Also, quadruplex DNA-bound BMVC emits fluorescence at 575 nm whereas emission is at 545 nm for BMVC bound to double-stranded DNA (28). Detection of emission at 575 nm upon mixing BMVC with human chromosomal DNA suggests the presence of tetrahelical structure in the genomic DNA. Moreover BMVC fluorescence measured at the ends and other regions of metaphase chromosomes identifies tetraplex-bound BMVC fluorescence at telomere-proximal regions (28). Compelling evidence for the formation of tetraplex structures *in vivo* was obtained by their direct electron microscopic detection in transcribed plasmid DNA. Electron microscopy showed that loops of 150-500 bp are formed *in vitro* in transcribed pRX15F plasmids and in plasmids propagated in *E. coli* (29). These loops contain tetraplex structures in the guanine-rich non-template strand and DNA-RNA hybrid on the transcribed strand. The formation of quadruplex structure in the non-transcribed strand was verified by its cleavage by the highly specific tetraplex cutting GQN1 endonuclease (see 4.5 below) and by a selective tight binding of the tetraplex DNA protein nucleolin (29) (see also 4.1). In summary, the application of specific antibodies, use of quadruplex DNA binding dye and direct electron microscopic observations provide independent lines of evidence for the existence *in vivo* of tetraplex structures of DNA. It is reasonable to expect that these early observations will be confirmed as experimental approaches and reagents of higher resolution become available.

3.2. Proposed biological roles of tetraplex DNA

3.2.1. Functions in the regulation of gene transcription

A substantive body of data supports the claimed role of tetrahelical DNA structures in the regulation of gene transcription. As mentioned above, computer surveys demonstrated high prevalence of tetraplex-forming guanine clusters in promoter regions of numerous prokaryotic genes (20) and gene function was found to correlate with the relative abundance of such clusters in coding and intronic regions of human genes (22). Focusing on particular genes, many authors identified tetraplex generating guanine-rich sequences in regulatory regions of diverse genes and demonstrated that such nucleotide tracts readily form different types of tetrahelices *in vitro*. Further, evidence was provided in some cases for preferential interaction of specific transcription proteins with the tetraplex domains. Table 1 compiles data on the presence of tetraplex-forming sequences in gene regulatory regions and the proposed roles of the quadruplex structures in transcriptional control. In nearly all the cases the tetraplex generating sequences were identified in gene promoters or enhancers. The listed specific sub-types of the quadruplex structures were all formed *in vitro* and thus their relevance to the situation *in vivo* has yet to be demonstrated. Proposed negative or positive effects of a tetrahelical structure on transcription are also based in most cases on indirect or on *in vitro* gathered results. Nevertheless, the overall body of evidence is in line with the idea that by attracting or dispelling transcription factors, tetraplex DNA structures serve as regulatory elements in the control of gene expression. A complete survey of the data compiled in Table 1 which is

beyond the scope of this review article, can be found elsewhere (30). The discussion here will thus be limited to only two representative cases.

3.2.1.1. Tetraplex formation in the DNA control region of *c-myc*

An early report showed that a guanine-rich sequence derived from a *c-myc* major control region located -142 to -115 bp of the promoter P1, folds back to generate an intramolecular G'4 tetraplex structure (31). Since activation of *c-myc* renders the examined control region nuclease hypersensitive, it was speculated that by attracting the nuclease hypersensitive binding proteins hnRNPs A1 and K and CNBP, the tetraplex structure may act to promote transcription. While confirming the formation of tetraplex structure in the *c-myc* control region, more recent results indicate that it is likely to repress rather than stimulate transcription. Of the two intramolecular tetraplexes generated by the *c-myc* control sequence (32, 33) only one, a chair-form variant, was found to be biologically relevant (32). A G→A point mutation that obliterates the formation of this tetraplex structure increases by 3-fold the basal transcription of *c-myc* in HeLa S3 cells. Conversely, the cationic porphyrin TMPyP4 which stabilizes this quadruplex diminishes expression of *c-myc* (32). These data indicate, therefore, that the tetraplex domain in the *c-myc* control region acts to suppress transcription. Consequently, several additional ligands that bind and stabilize tetraplex telomeric DNA and inhibit telomerase activity, were found to also suppress the *in vitro* transcription of *c-myc* (34). Recently described finer structural analyses of the *c-myc* tetraplex structures and their interaction with TMPyP4 (35, 36) and the dynamic equilibrium between duplex and tetraplex structures of the *c-myc* sequence (37) may potentially contribute to a rational design of tetraplex stabilizing agents that inhibit *c-myc* expression and may act to suppress tumor initiation or progression.

3.2.1.2. Tetraplex structures in DNA control regions of muscle-specific genes

An early report indicated that the myogenic determination protein MyoD binds to a tetrahelical structure of guanine-rich enhancer sequence of the Muscle Creatine Kinase (*MCK*) gene more tightly than to its target d(CANNTG) E-box motif (38). This observation raised the possibility that tetraplex structures in regulatory sequences of muscle-specific genes may also control their transcription. A more recent study showed that promoter or enhancer sequences of the muscle-specific genes human sarcomeric Mitochondrial Creatine Kinase (*sMtCK*), mouse *MCK* and α7 integrin display disproportionately high incidence of guanine clusters (39). Furthermore, sequences derived from the regulatory regions of the three genes readily generates *in vitro* secondary structures that include hairpin formations and intra- and intermolecular tetraplexes (39). Binding of heterodimers of the basic helix-loop-helix proteins MyoD and E12 or E47 to the consensus E-box sequences in promoter or enhancer regions of muscle-specific genes activates the transcription of muscle-specific genes. By contrast, MyoD homodimers bind E-box less tightly and are less efficient activators of transcription. It was shown that homodimers of recombinant MyoD selectively form complexes with bimolecular tetraplex structures of muscle-specific regulatory sequences (40). Preferential binding of homodimeric MyoD to bimolecular tetraplex DNA structures over E-box DNA is reflected by ~20 to 40-fold lower dissociation constants, K_d, of the MyoD-tetraplex DNA complexes. Conversely, MyoD-E47 heterodimers form tighter complexes with E-box as indicated by their ~7 to 20-fold lower K_d values relative to the protein-bimolecular tetraplex DNA complexes (40). It was speculated that by binding tightly to MyoD homodimers, the tetraplex domains in promoter or enhancer regions of muscle-specific genes act to suppress their expression. Transcription is proposed to be initiated upon the formation of MyoD-E47 heterodimers, their dissociation from the quadruplex domains and preferential binding to the E-box motif.

3.2.2. Roles in telomere metabolism

Telomeres are specialized nucleoprotein complexes at the end of linear eukaryotic chromosomes. A guanine-rich single-stranded telomeric DNA overhang is arguably the most investigated of all tetrahelix forming sequences. The exact repeats of guanine clusters in the telomeric sequences $d(TTTTGGGG)_n$ in *Oxytricha nova* or *Stylonychia lemnae*, $d(TTGGGG)_n$ in *Tetrahymena thermophila* or $d(TTAGGG)_n$ in all vertebrates are ideally placed to promote the formation *in vitro* of hairpins that are stabilized by guanine·guanine base pairs and unimolecular and bimolecular tetraplex structures. Vertebrate telomere DNA is composed of hundreds to thousands $d(TTAGGG)_n{\cdot}d(CCCTAA)_n$ tandem repeats and has at its 3'-terminus a ~150 nt long single-stranded $d(TTAGGG)_n$ overhang. As detailed elsewhere in this article (see 4 below), the existence of multiple proteins in protozoa and eukaryotes that selectively bind, promote the formation of or destabilize tetraplex telomeric DNA is in line with the notion that these tetrahelical structures exist *in vivo*. Also, use of specific antibodies and tetraplex interacting dye provide more direct evidence for the presence of quadruplex telomere formations in nuclear DNA (25, 28). The major mode of telomere length maintenance is the extension of the protruding 3' guanine-rich strand by the ribonucleoprotein enzyme telomerase. An early study examined the capacity of the different folded forms of telomeric DNA to serve as primers for the *Oxytricha nova* telomerase *in vitro*. It was demonstrated that formation of a tetraplex structure in the guanine-rich telomeric primer blocks its extension by telomerase (41). This led to the suggestion that folding of the telomeric DNA into a tetraplex structure may act to impede elongation of the guanine strand by telomerase. Results of a considerable number of subsequent investigations support the idea that a stable tetrahelix in telomere ends serves to negatively control their extension by telomerase. First, it was reported that although the catalytic activity of human telomerase increases at elevated K^+ concentrations, its processivity is diminished. This observation was interpreted as an indication that K^+ ion dependent formation of tetraplex structure of the guanine-rich telomeric DNA strand impedes the progression of telomerase (42). In a more recent study it was demonstrated that K^+-stabilized

unimolecular quadruplex telomeric DNA failed to be extended by recombinant telomerases from *Tetrahymena* and *Euplotes* and was not bound by the protein component of these enzymes. By contrast, the two telomerases did bind to and extended a parallel-stranded intermolecular quadruplex structure of the telomeric sequence (43). In a more direct line of investigation it was demonstrated by several laboratories that stabilization of tetraplex structures of telomeric DNA by a variety of drugs inhibits the telomerase catalyzed extension of telomeres. In the main, drugs that bind to and stabilize tetrahelical structures of the guanine-rich telomeric sequence also block the action *in vitro* of telomerase and act *in vivo* to inhibit telomere elongation and proliferation of cancer cells. Since the scope of this article cannot allow a full survey the plethora of works describing diverse tetraplex-interacting drugs and their telomerase inhibitory effects, the reader is referred to several excellent reviews on the subject (44-48). All in all, this body of evidence gives credence to the proposed role of tetraplex telomeric DNA structures as negative regulators of telomerase activity. In addition to its inhibitory effect on telomerase progression, unimolecular tetraplex structure of the guanine-rich telomeric strand was reported to diminish the efficiency of assembly of telomeric complexes that include the proteins TRF2 and pot1, suggesting that guanine tetrads may hinder the formation of some telomeric nucleoprotein complexes (49). Last, additionally to their proposed roles in the regulation of telomere extension and organization, it was also suggested that intermolecular telomeric tetraplex structures participate in the pairing of homologous chromosomes (see 3.3.3).

3.2.3. Function in DNA recombination

The potential role of intermolecular tetrahelical complexes in DNA recombination was recognized concomitantly with the discovery of the self-association under physiological-like conditions of single-stranded DNA containing short guanine clusters to form Hoogsteen-bonded four molecular parallel tetrahelices (6). It was proposed that the self-recognition of guanine cluster in DNA may draw together and fasten four accurately aligned homologue chromatids during meiosis. In parallel, the observation that the telomeric guanine-rich DNA overhangs readily form bimolecular tetraplexes raised the possibility that these structures mediate the pairing of homologous chromatids at their telomere ends (7). That quadruplex DNA may promote recombination was recently directly demonstrated by the finding that fold-back tetraplexes promote RecA-independent homologous recombination in plasmids (50). Indirect indications suggest that quadruplex structures also play a role in the more intricate machinery of chromosome pairing in eukaryotes. To ascertain proper alignment of eukaryotic homologous chromosomes which is required for crossovers formation and accurate segregation, the chromosomes are held together during the pachytene stage of meiotic prophase by proteinaceous Synaptonemal Complexes (SC) (reviewed in 51). The DNA binding protein Hop1 which is a component of the *Saccharomyces cerevisiae* SC (52), has been shown to bind to and to promote the generation of tetraplex DNA (53) and to enhance the pairing of double-stranded DNA helices through the formation of Hoogsteen-bonded quadruplex structures (51).

3.2.4. Role in immunoglobulin class switch recombination

Immunoglobulin class switch recombination, a private and unique type of DNA recombination, is responsible for altering the mode of antigen clearance without changing the antigen specificity of the immunoglobulin molecule. In this process, an expressed rearranged heavy chain variable region (VDJ) is joined to a new 3' positioned constant (C) region, while the intervening DNA sequence is eliminated. Switch recombination is dependent upon the presence upstream to each C region of guanine-rich sequences termed switch (S) regions. These sequences exhibit a proclivity to form G4 tetraplex structures which are implicated in recombination (6). LR1, a heterodimer of nucleolin and hnRNP D, is a B cell specific protein that binds in a sequence-specific fashion to the S regions (54, 55). The activity of this protein correlates with the ability of cells to conduct recombination of switch DNA substrates. LR1 and its isolated nucleolin and hnRNP D subunits were reported to have a significantly higher affinity for G4 DNA (K_d = 0.25 to 0.5 nM) than for duplex DNA (K_d = 1.8 nM) (56, 57). Another protein, Activation-Induced cytidine Deaminase (AID) that activates both switch recombination and somatic hypermutation was also shown to bind specifically to tetraplex DNA structures that form co-transcriptionally (58). Last, switch recombination also depends on the activity of the mismatch repair heterodimer MutSα. This protein was reported to bind G4 DNA at a high affinity (≤ nM) and as electron microscopy reveals, it associates with transcribed S regions and promotes their synapsis (59). Put together, the predilection of the S region guanine-rich sequences to assemble into intermolecular tetraplex structures and the preferential binding of the switch proteins LR1 and its subunits, AID and MutSα to the tetrahelical DNA are in line with a role of intermolecular G4 structures in immunoglobulin gene diversification.

3.2.5. Involvement in nucleotide expansion

Side-by-side with its proposed multiple physiological roles, tetraplex structures in DNA may have detrimental effects as illustrated by their suggested causative role in some nucleotide expansion disorders. Following the discovery in 1991 that fragile X syndrome (FXS) is caused by dynamic expansion of a d(CGG) trinucleotide repeat sequence in the 5'-UTR region of the *FMR1* gene (60-62) and that Spinal and Bulbar Muscular Atrophy (SBMA) results from a d(CAG) expansion in the androgen receptor gene (63, 64), it was established that many repeats at different loci in the human genome are subject to dynamic expansion. Such expansion mutations result in a diverse class of about 20 different neurological, neuromuscular and neurodegenerative disorders known as nucleotide or repeat expansion disorders. Numerous studies demonstrated an inclination of many of the expanded repeats to form secondary structures, both at the DNA and RNA levels. Of note is the propensity of the FXS d(CGG) repeat sequence to fold into hairpin structures (65, 66) that readily pair to generate stable intermolecular tetraplexes (14, 67, 68).

Quadruplex structure of d$(CGG)_n$ hinder the progression *in vitro* of DNA polymerases (68, 69) and its unwinding by the human Werner syndrome helicase (WRN) alleviates the blocking of the polymerase (69). Taken together, these findings suggest that by obstructing the replication of the repeat sequence, the tetrahelical formation affects dissociation of the polymerase from the DNA template, which may subsequently result in slippage and expansion of the repeat sequence. In analogy, tetraplex structures of DNA may also be a causative factor in the nucleotide expansion disorder progressive myoclonus epilepsy 1 (EPM1). This disease is associated with expansion of the dodecamer (CCCCGCCCCGCG)·d(CGCGGGGCGGGG). Whereas the cytosine-rich sequence has been shown to fold into hairpin (70) or into unstable intermolecular i-motif protonated complex (71), the guanine-rich strand forms intramolecular tetraplex structure at neutral pH and an intermolecular tetrahelix in acid pH (71).

3.2.6. Role in the assembly of HIV nucleocapsid

Formation of tetrahelical RNA was described within a short time after tetraplex structures were identified in DNA. The retrovirus HIV-1 packages two single-stranded homologous RNA genomes within a gag protein-RNA complex. It was reported that an RNA fragment representing the 5' end of the gag gene (positions 732-858) dimerizes spontaneously into a monovalent cation-stabilized RNA-RNA tetraplex structure (72). This observation established that retroviral RNA - and RNA in general - can self-recognize to mediate genome organization and recombination. Further, it was speculated that the HIV-1 nucleocapsid protein acts as a non-catalytic RNA chaperone that lowers the activation barrier for tetraplex formation (72). Subsequent investigation demonstrated that the DNA reverse transcripts of HIV-1 also form several types of tetrahelical structures, with a G'2 antiparallel dimer of two hairpins being the most stable (73). It was proposed that such tetraplex constructs are likely to be transiently formed in the course of replication or repair of the viral guanine-rich DNA flaps. A subsequent report showed that the HIV-1 nucleocapsid protein binds with high affinity to the viral G'2 tetraplex DNA structure suggesting a role for this protein-DNA interaction in the compact assembly of the nucleocapsid (74). Interestingly, a more recent study revealed that binding of the HIV-1 nucleocapsid protein affects destabilization of a unimolecular tetraplex structure of the thrombin binding sequence 5'-d(GGTTGGTGTGGTTGG)-3' (75). Identification of HIV-1 quadruplex DNA structure raised the possibility that it may serve as target for anti AIDS drugs. A 16-nt guanine-rich sequence d(GGGGTGGGAGGAGGGT)-3' was reported to act as inhibitor of the HIV-1 reverse transcriptase and especially of its integrase, the retroviral enzyme that catalyzes integration of the retro-transcribed HIV-1 DNA into the host cell nuclear DNA (76). Detailed NMR analysis showed the tetraplex to be a novel "interlocked" bimolecular structure (77). This finding opens the way for rational design of tetraplex inhibitors of the enzyme that will have higher potency by being positioned within a channel/canyon that are, respectively, formed within a dimer and between the two dimers of the integrase tetramer (77, 78).

4. TETRAPLEX DNA INTERACTING PROTEINS

Following the initial description of the formation of tetraplex structures of guanine-rich DNA (6-8) many laboratories identified proteins in diverse organisms that exhibit assorted selective physical and functional interactions with tetrahelical DNA. The existence of cellular proteins that preferentially interact with tetraplex DNA structures provides a strong argument for the existence of quadruplex formations in genomic DNA. At the same time, the different functional effects of these proteins shed light on possible biological roles of tetrahelical domains in DNA. As detailed below, quadruplex DNA interacting proteins can be classified by function into several distinct groups. Proteins of the first class bind preferentially and at high affinity to tetraplex DNA, and in some cases increase the stability of the bound tetrahelix. By also binding to the DNA, proteins of a second group greatly promote the formation of tetraplex structures. Binding proteins of a third type act conversely to destabilize quadruplex DNA in a non-catalytic fashion. Proteins of a fourth group are DNA helicases that preferentially catalyze the unwinding of DNA tetrahelices in ATP-hydrolysis dependent reaction. A fifth class of proteins consists of nucleases that specifically incise DNA at or next to tetraplex domains. Last, in a recent development, proteins are designed and engineered to carry out assigned interactions with tetraplex DNA. The multiplicity of tetraplex DNA interacting proteins and their diverse activities provide a strong indication for the dynamic nature of tetraplex DNA. A reasonable assumption is that regulated interaction of guanine-rich DNA sequences with different proteins alternatively causes augmentation of tetraplex formation, stabilization of DNA tetrads, their destabilization or unwinding or selective nucleolytic digestion of the tetrahelices.

4.1. Tetraplex DNA binding and stabilizing proteins

A compilation of proteins that bind different variants of tetraplex DNA is presented in Table 2. On first glance, one impressive feature is the wide evolutionary prevalence of the binding proteins - from a retrovirus, on to lower eukaryotes, to yeast and protozoa and to mammals. In fact, multiple non-telomeric tetraplex DNA binding proteins were detected but not characterized in extracts of *E. coli*, yeast, *Arabidopsis* and human cells (79). Second, different proteins display preference for tetraplex DNA over single-stranded or duplex DNA and many bind selectively to either intra- or intermolecular and to parallel or antiplarallel tetraplex structures. Most of the described proteins associate rather tightly with their preferred tetrahelical DNA form as reflected by dissociation constants of the resulting protein-DNA complexes that are commonly in the nanomolar (38, 40, 80, 81) and sub-nanomolar (56, 57) range. Notably, binding of some of the proteins increases the stability of tetraplex structures (40, 80, 82) or renders them more resistant to nucleolytic attack (81, 83).

Table 2. Tetraplex DNA binding proteins

Protein	Source	tetraplex DNA lignad BOUND[1]	Reference
Thrombin	Human	Antiparallel G'4 DNA	85, 86, 143, 144
MyoD	Human	Parallel G'2 DNA	38-40
QUAD	Murine	Parallel G4 DNA	82, 145
Macrophage scavenger receptor	Bovine	Parallel G4 DNA	84
Tetrahymena binding protein (TBP)	Tetrahymena thermophila	Parallel G4 DNA	146
G4p1	Saccharomyces cerevisiae	Parallel G4 and antiparallel G'2 DNA	147
G4p2/Stm1	Saccharomyces cerevisiae	Parallel G4 DNA	148-150
quadruplex Telomeric binding Protein 42 (qTBP42) [CBF-A]	Murine	Parallel G4 > G'2 DNA ~ G'4 DNA	80, 151
Unimolecular quadruplex Telomeric binding Protein 25 (uqTBP25)	Murine	Antiparallel G'4	81
LR1 (nucleolin and hnRNP D)	Murine	Parallel G4 DNA	56, 57
Gene 5 protein (g5p)	Bacteriophage fd/Ff	Inter- and intramolecular	152-155
Ku protein	Human	Antiparallel G'2 DNA	83
Nucleocapsid protein (NCp)	HIV-1	Parallel G4 and G'2 DNA	74
MutSα	Human	Parallel G4 DNA	59
Cytoplasmic intermediate filament proteins (cIF)	Murine	Parallel G4 > G'2 DNA > G'4 DNA	156

Listing of the tetraplex DNA binding proteins is according to their approximate chronological order of discovery. [1]In some cases the stoichiometry or the DNA strand orientation of the quadruplex structure were not defined.

Table 3. Tetraplex DNA formation promoting proteins

Protein	Source	tetraplex DNA Formed	Reference
Subunit β of telomere binding protein	*Oxytricha nova*	Parallel G4 DNA	87, 88, 90
RAP1	*Saccharomyces cerevisiae*	Parallel G4 DNA	91, 92
DNA topoisomerase I	Human	Parallel G4 DNA	95
Hop1	*Saccharomyces cerevisiae*	Parallel G4 DNA and antiparallel G'2 DNA	53
Thrombin	Human	Antiparallel G'4 DNA	97
Telomere end-binding proteins α and β	*Stylonychia lemnae*	Antiparallel G'2 DNA	27

Listing of the tetraplex DNA binding proteins is according to their approximate chronological order of discovery

Of prime interest are the biological implications of the tight association of proteins with specific tetraplex DNA variants. Plausibly, different functions may be served by the binding of specific proteins to tetrahelical structures of different guanine-rich sequences. Thus, for instance, the interaction of the HIV-1 zinc finger nucleocapsid protein with G4 and G'2 forms of tetraplex DNA may play a part in the compact assembly of the nucleocapsid (ref. 74), see also 3.3.6). In a different context, binding of the LR1 heterodimer and its constituent nucleolin and hnRNP D subunits, as well as the association of AID protein with tetraplex structures of the immunoglobulin class switch region is a likely indication of the role of tetraplex DNA and its binding protein in switch region recombination (refs. 56-58, see also 3.3.5 above). The tight binding of MyoD homodimers to tetraplex structures of regulatory sequences of muscle-specific genes (38, 40) was proposed to negatively regulate the transcription of these genes (40). Binding to and protection of intramolecular tetraplex forms of the guanine-rich telomeric DNA strand by proteins such as uqTBP25 (81) may stabilize the 3' end of telomeres and further inhibit their telomerase catalyzed extension.

In some cases the binding of quadruplex DNA to a protein is unlikely to reflect a specific *bona fide* biological function of the interaction. Such are the interaction with tetraplex DNA ligands of thrombin, a multifunctional serine protease, or of the macrophage scavenger receptors (84). However, the inhibition of thrombin-catalyzed fibrin clot formation by nanomolar concentrations of a tetraplex aptamer (85) raises the possibility that designed tetraplexes may serve as highly effective inhibitors of thrombin action. Engineering of potent quadruplex inhibitors can be based on our current detailed knowledge of the molecular structure of the aptamer (86). In an analogous line of thought, DNA aptamers that bind tightly to the HIV-1 integrase hold the potential of effectively acting as efficient anti-HIV agents (77, 78).

4.2. Proteins that promote the formation of tetraplex DNA

The discovery of proteins from various organisms that greatly promote the formation of tetraplex structures of guanine-rich sequences (Table 3), provides a sound argument for the existence of these structures in the cell and for their dynamic nature. The earliest described quadruplex promoting protein is the β subunit of the *Oxytricha nova* telomere-binding protein (87, 88). Monomers of 56-kDa (α) and 41-kDa (β) protein subunits bind specifically to single-stranded d(TTTTGGGGTTTTGGGG) termini of *Oxytricha* macronuclear telomeres and assemble into a heterodimeric telomere binding protein (89). The isolated β subunit or its basic carboxy terminal domain greatly stimulate the formation *in vitro* under physiological-like conditions of tetraplex structures of the *Oxtricha* and *Tetrahymena* sequence d(TTGGGGTTGGGG) (87). Whereas protein-independent formation of four-molecular parallel G4 tetraplex DNA is a second order reaction, it is enhanced 10^5 to 10^6-fold in a first-order reaction by saturating amounts of the protein (88). Raman and fluorescence spectroscopy analyses indicate that the β subunit first binds to and ruptures a guanine·guanine stabilized hairpin DNA motif, converting it into single-strands which go on to assemble into parallel four-molecular tetraplexes (90). The biological significance of the protein-telomere DNA interaction in the ciliate *Stylonychia lemnae* was recently probed with specific anti tetraplex DNA antibodies and RNA

Table 4. Tetraplex DNA unwinding helicases

HELICASE	Source	unfolded tetraplex DNA lignad	Reference
Resolvase (DEHX protein)	Human	Parallel G4 DNA	157, 158
Large T-antigen	SV40	Parallel G4 DNA and antiparallel G'2 DNA	159
RecQ helicases			
• WRN	Human	Antiparallel G'2 DNA	69, 160, 161
• BLM	Human	Parallel G4 DNA	98, 102, 161
• Sgs1	*Saccharomyces cerevisiae*	Parallel G4 DNA and antiparallel G'2 DNA	101, 102, 162

The order of listing of tetraplex DNA unwinding helicases is in the approximate order of their discovery

interference. It was reported that TEBβ promotes the formation *in vivo* of antiparallel rather than parallel tetrahelical telomeric DNA as expected when two nanochromosomes are joined end-to-end by the pairing of two (TTTTGGGGTTTTGGGG) overhangs (27). Formation of the tetraplex structures, which requires the presence of an intact C-terminal domain of TEBβ, is cell cycle dependent such that no tetraplex is detected in replicating DNA. However, the telomere tetrahelical DNA structure is maintained in replicating DNA when phosphorylation of TEBβ is inhibited *in vivo* (27). It is likely, therefore, that TEBβ is phosphorylated upon initiation of the S-phase, causing the dissociation of the TEBαβ heterodimer from telomeres and the unfolding of the quadruplex structure.

Proteins in additional species were also found to promote the formation of tetraplex DNA structures. The *Saccharomyces cerevisiae* Repressor Activator Protein 1 (RAP1) which is essential for telomere maintenance, binds in a sequence-specific fashion to the double-stranded telomeric DNA and to its d(TG_{1-3}) single-strand. In binding to the guanine-rich single-strand DNA sequence, RAP1 promotes its assembly into an intermolecular quadruplex structure (91). Circular dichroism analysis indicates that RAP1 specifically stimulates the formation of parallel-stranded G4 DNA (92). Another yeast protein, Hop1 which is a required component of the meiosis synaptonemal complex, was reported to enhance the formation of quadruplex DNA. This protein binds to G4 DNA, (K_d = 0.2 nM), 1000-fold more tightly than to duplex DNA and transforms single-stranded oligomer that includes three clusters of contiguous guanines into an intermolecular G4 tetraplex (53). A subsequent report demonstrated that Hop1 is also able to promote the formation of a parallel tetraplex between four-strands of two duplex DNA molecules that contain arrays of clusters of mismatched guanine residues (93). Hence, it may be that the quadruplex promoting activity of Hop1 is necessary for the formation of inter-chromosomal tetraplex synapses that are proposed to mediate homolog chromosome pairing and recombination (6, 94). Human topoisomerase I (Topo I) which alters the topological arrangement of DNA by cleaving one strand in duplex DNA was also found to promote formation of tetraplex DNA and to bind to preformed intra- and intermolecular quadruplexes (95). Topo I cleaving activity was found to be inhibited by its binding to tetraplex-forming oligonucleotides and to some oligomers that do not form quadruplex structures (96). Thus association of the enzyme with such non-cleaved DNA sequences and structures may negatively regulate the double-stranded DNA cleaving activity of the enzyme. Last, in a recent work thrombin, the tetraplex aptamer binding protein was reported to promote the formation of this intramolecular tetrahelix in the absence of K^+ ions which are normally necessary for protein-independent generation of the quadruplex (97). Although thrombin may act as a chaperone that facilitates the arrangement of DNA into tetraplex conformation, it is clearly a non-physiological reaction. Hence, one should keep in mind that the mere binding of a protein to quadruplex DNA or even the promotion of its formation may not necessarily reflect a true biological function.

4.3. Tetraplex DNA unwinding helicases

The existence of cellular proteins that facilitate the folding and assembly of DNA into tetraplex configurations and of proteins that increase the stability of DNA tetrahelices suggest a dynamic state of formation of these structures. That quadruplex DNA is likely to be in a transient state is further highlighted by the discovery of helicases; DNA unwinding enzymes that preferentially melt tetrahelical DNA structures. As summarized in Table 4, helicases of such selectivity were identified in Simian virus 40 (SV40), yeast and human cells. It is notable that three of the described enzymes are members of the RecQ family of helicases whose member proteins are found in a broad range of organisms - from *E. coli* to humans. All the tetraplex unwinding helicases act catalytically and all require for their activity hydrolysis of nucleotide triphosphate, normally ATP, and the presence of Mg^{2+} ions. Further, the quadruplex DNA substrates are unwound with a 3'→5' polarity and the enzymes require that the tetraplex structure possesses a short 3' single-stranded tail that serves as a "loading dock" for the enzymes. It should be stressed, however, that none of the described helicases unwinds tetraplex DNA exclusively and all the enzymes are also able to unfold, albeit at a lower efficiency, other DNA structures such as duplex DNA, holiday junctions or triplex DNA.

Most of the information on tetraplex DNA unwinding helicases is based on experiments *in vitro*. However, some of the pathologies associated with genetic loss of a specific helicase may possibly be attributed to the absence of nuclear tetraplex DNA unwinding activity. The helicases may well be responsible for the unfolding of aberrant quadruplex structures that are likely to be formed in the course of normal DNA transactions such as replication or recombination. A tetraplex structure in template DNA was shown to obstruct the progression *in vitro* of the yeast replicative DNA polymerase δ. WRN helicase collaborates selectively with polymerase δ in unwinding the blocking quadruplex structure thus allowing replication to progress (69). This observation is in line with the reported blocked and collapsed replication forks in Werner syndrome cells which lack functional WRN protein. In a related case, the

Table 5. Tetraplex DNA destabilizing proteins

Protein	Source	tetraplex DNA unfolded	Reference
CArG-box binding protein A (CBF-A)	Murine	Antiparallel G'2 DNA	104, 106, 163
hnRNP D	Murine	G'4, G'2 and G4 DNA	107, 108
hnRNP A2	Human	Antiparallel G'2 DNA	106
Cdc13p	Saccharomyces cerevisiae	Parallel G4 DNA	117
Unwinding Protein 1 (UP1)	Murine	G'4 DNA	111, 112
Protection Of Telomere 1 (POT1)	Human	G'4 DNA	118
Nucleocapsid protein (NCp)	HIV-1	Antiparallel G'4 DNA	75
Thrombin	Human	Antiparallel G'4 DNA	164
Replication protein A (RPA)	Human	G'4 DNA	122

In some cases the strand orientation of the unfolded DNA substrates was not defined

absence of the G4 DNA unwinding BLM helicase in Bloom syndrome can be correlated with the genomic instability and the raised frequency of sister chromatid exchange in patients' cells (98). Last, aberrant metabolism of telomeres in Werner syndrome cells may be explained by the function of WRN helicase in the regulated unfolding of quadruplex (99) and/or of a D-loop structure of telomeric DNA (100) that forms by the invasion of the 3' guanine-rich single-strand into the telomeric double-stranded repeat.

The unwinding of tetraplex DNA by some RecQ helicases was found to be effectively inhibited by some quadruplex DNA stabilizing small molecules. For instance, the quadruplex unfolding activity of yeast Sgs1 helicase is blocked by the tetraplex interactive drug N,N'-Bis[1-piperidino)]-3,4,9,10-perylenetetracarboxylic Diimide (101), and G4 DNA unwinding by Sgs1 and BLM helicases is inhibited by N-methyl mesoporphyrin IX (102). It was proposed recently that molecules designed to inhibit the unwinding of specific tetraplex domains in genomic DNA, such in telomeres, could potentially be used as inhibitors of cell proliferation and thus as anti cancer agents (103).

4.4. Tetraplex DNA destabilizing proteins

Additional evidence for the dynamic nature of tetraplex DNA is provided by the quadruplex DNA destabilizing activity of some cellular proteins. Isolated from diverse species, proteins of this class disrupt different types of tetrahelical DNA (Table 5). Unlike tetraplex unwinding helicases, all the destabilizing proteins unwind tetrahelical DNA in a non-catalytic fashion and the unfolding reaction does not require ATP hydrolysis. Interestingly, several of the destabilizing proteins are members of the hnRNP family. One of the earliest hnRNPs to be identified as a quadruplex DNA destabilizing protein is the CArG-box binding protein A (CBF-A). Initially identified as a G'2 quadruplex telomeric DNA binding and stabilizing murine protein that was originally termed qTBP42 (80) it was subsequently recognized as the previously defined hnRNP-related transcription regulating protein CBF-A (104). This protein possesses two highly conserved RNA Recognition Motifs, RRMs, (also called RNA Binding Domains; RBDs) that are characteristic of members of the hnRNP family. As in many other hnRNPs, CBF-A contains an RRM1 domain composed of $RNP2_1$ and $RNP1_1$ boxes and an RRM2 motif consisting of $RNP2_2$ and $RRM1_2$ boxes. Also present is an ATP/GTP binding fold which has been defined by its partial homology to functional NTP binding domain in other proteins but which does not appear to function as such in CBF-A (105). Whereas CBF-A stabilizes a bimolecular G'2 form of the telomeric repeat sequence $d(TTAGGG)_2$, it efficiently unfolds G'2 structure of $d(CGG)_n$. One of the two conserved motifs of CBF-A; $RNP1_1$ or the ATP/GTP binding fold is necessary for its tetraplex DNA destabilization activity. One of the same two motifs is also required for destabilization of quadruplex DNA by additional members of the hnRNP family; hnRNP A2 and a genetically modified variant of hnRNP A1 (106). Intra- and intermolecular quadruplex structures of the telomeric sequence $d(TTAGGG)_{1-4}$ were found to be destabilized *in vitro* by hnRNP D, another member of the same family of proteins (107, 108). Also, the single-stranded binding protein UP1, a proteolytically generated N-terminal fragment of hnRNP A1 (109, 110) binds and destabilizes intramolecular tetraplex forms of the hypervariable minisatellite sequence $d(GGCAG)_5$ and of the telomeric sequence $d(TTAGGG)_4$ (111). This protein was also reported to unfold an undefined non-quadruplex secondary structures of the $d(CGG)_7$ and $d(CGG)_{16}$ sequences (112). In addition to their regulatory roles in the splicing of pre mRNA and in the packaging and transport of mRNA, proteins of the hnRNP family; hnRNP A1, its derivative UP1 and hnRNP D were implicated in the maintenance of telomeres (107, 113, 114). It is tempting to propose that the capacity of proteins of this family to destabilize quadruplex structures of telomeric DNA is a reflection of their involvement in telomere maintenance. Another potential function of the tetraplex destabilizing activity of hnRNPs could be the removal of tetrahelical structures that aberrantly form in non-telomeric genomic DNA sequences.

Non-enzymatic destabilization of tetrahelical DNA structures was also described for proteins other than hnRNPs. The *Saccharomyces cerevisiae* protein Cdc13p binds telomeres *in vivo* and is required for their replication and protection (115, 116). This protein was shown to partially denature a G4 quadruplex structure of the yeast $d(TG_{1-3})$ telomeric repeat (117), implying that inter-conversion of single-stranded and quadruplex DNA may be involved in telomere maintenance. That such structural transformation of DNA in telomeres may be a common process is demonstrated by the finding of two additional eukaryotic tetraplex telomeric DNA disrupting proteins. One is the Protection Of Telomere 1 protein (POT1) which binds to the guanine-rich single-stranded telomeric overhangs. By binding to quadruplex telomeric DNA ends *in vitro*, human POT1 enables core telomerase to extend the DNA. It is not yet clear, however, whether POT1 facilitates DNA extension by actively disrupting the tetrahelical structure or by capturing a small proportion of unfolded

Table 6. Nucleases specific for tetraplex DNA

Nuclease	Source	DNA substrate	Reference
KEM1 gene product	*Saccharomyces cerevisiae*	Single-strand 5' to G4 DNA	123-125
G Quartet Nuclease 1 (GQN1)	Human	Single-strand 5' to G'2 and G4 DNA	127
Mre11	*Saccharomyces cerevisiae*	Single-stranded and G'4 DNA	126

DNA in an equilibrium population of DNA molecules, driving the binding to completion (118). Replication Protein A (RPA), a highly conserved heterotrimeric eukaryotic single-stranded DNA binding protein was found to be essential for the maintenance of yeast telomeres (119) and to stimulate the *in vitro* action of telomerase (120, 121). The enhancement of telomerase activity by human RPA is plausibly due to its recently described capacity to efficiently destabilize *in vitro* an intramolecular G'4 structure of the telomeric sequence $d(TTAGGG)_4$ (122). In summary, the identification of several quadruplex telomeric DNA destabilizing proteins underscores the likelihood of the function of dynamic formation and dissolution of tetrahelical DNA structures in telomere maintenance.

4.5. Tetraplex DNA selective nucleases

DNA recombination requires cleavage of strands in the recombining chromosomes. The identification of nucleases that incise DNA within or close to tetraplex DNA domains strengthens the proposed role of G4 DNA in the formation of interchromosomal synapses. Table 6 lists three proteins that were identified as nucleases that cleave DNA next to or within tetraplex domains. The earliest protein identified is the product of the *Saccharomyces cerevisiae* gene *KEM1* (also called *SEP1*, *DST2*, *XRN1* and *RAR5*). Initially detected in yeast cell extracts as a telomere binding protein, KEM1 was found to cleave DNA that includes a four-stranded G4 domain but to exhibit low or no nucleolytic activity toward single- or double-stranded DNA substrates (123). Following its purification to homogeneity and its identification as the *KEM1* gene product, the enzyme was shown to cut single-stranded DNA most efficiently when it is positioned 5' to a G4 region (124). In attempting to elucidate the biological function of KEM1 it was found that homozygous deletions of the KEM1 gene block meiosis at the pachytene stage and that they are associated with cellular senescence and telomere shortening (125). It was hypothesized, therefore, that KEM1 participates in recombination by excising DNA next to tetraplex synapses between paired homologous chromosomes and that its activity is also required for telomere maintenance. Another recently described yeast G4 DNA selective endonuclease is the *Saccharomyces cerevisiae* protein Mre11 which binds to G4 DNA more tightly than to single- or double-stranded DNA or to G'2 quadruplex DNA and cleaves the G4 DNA (126). A nuclease that cuts single-stranded DNA 5' to a G4 domain was also isolated from human cells. This enzyme, named G Quartet Nuclease 1 (GQN1) does not digest single- or double-stranded DNA, Holiday junctions or tetraplex RNA and it specifically cuts single-stranded DNA located few nucleotides 5' to either G'2 or G4 domains (127). Notably, in failing to incise tetraplex RNA, GQN1 is distinguished from a previously described Mouse Cytoplasmic Exoribonuclease (mXRN1p) which cleaves G4 RNA (128) Elevation of GQN1 activity in B cells raised the possibility that this enzyme is involved in immunoglobulin heavy chain class switch recombination (127).

4.6. Selection and design of tetraplex DNA interacting proteins

The wide assortment of cellular proteins that interact with tetraplex DNA structures raises the appealing option of selecting for or engineering proteins that will affect quadruplex DNA transactions *in vivo* more efficiently than the naturally occurring proteins. Such proteins can be useful for instance for restricting telomere length and inhibiting cancer cell proliferation or, alternatively, for the modulation of DNA recombination. An early step in this direction was the application of phage display technique to select among zinc finger proteins for tight binders of a unimolecular quadruplex structure of the telomeric repeat sequence $d(TTAGGG)_5$ (129). Clones expressing three-finger proteins that were selected for their ability to bind the G'4 telomeric DNA in sequence- and structure-specific fashion were shown to possess a strong amino acid consensus, indicating common requirement for the selective recognition of the quadruplex ligand (129). By binding and stabilizing the telomeric G'4 DNA structure, one of the selected proteins, Gq1, arrests the progression of *E. coli* Klenow DNA polymerase along a tetraplex-containing template and inhibits the activity of telomerase rather effectively (130). In a recent work the selectivity of tetraplex binding by Gq1 was investigated by substituting its zinc fingers with those of the transcription factor Zif268 (131). This study revealed that the zinc fingers of Gq1 allow tighter fit with parallel-stranded G'4 DNA than those of Zif268. Moreover, modeling identified key zinc finger residues in Gq1 that maintain the protein in optimum conformation for interaction with the tetrahelical DNA (131).

5. PERSPECTIVES

Reflecting on the body of experimental evidence that was accumulated in the recent years, it becomes quite evident that the study of tetrahelical DNA has come to age. Much is already known on the fine structural details of the plethora of *in vitro* generated quadruplex DNA variants. Moreover, evidence is highly suggestive that tetrahelical DNA exists in the genomes of various organisms. This claim is supported by the identification of multiple interacting proteins that increase the stability of diverse tetrahelical DNA structures, promote their formation or unwind or cleave them. What does the future hold? A likely major direction of research is the expansion of our understanding of the biological functions of tetraplex DNA. In this context there are two main areas of study. One is the exploration of the roles of tetrahelical DNA structures in the regulation of gene expression and another is the investigation of functions of quadruplex structures of telomeric DNA in the maintenance of telomere length and integrity. Obviously, both tetraplex DNA interacting small molecules and proteins have to be used to probe the effects of structural inter-conversions of quadruplex forming DNA sequences in living cells. Insights into the biology of

tetraplex DNA is likely to be gained by applying such reagents to increase or diminish the stability of DNA tetrahelices or to cleave them *in vivo*. Such approach calls for the development of reagents of higher specificity than the ones now available. Obtaining meaningful results depends on the application of reagents that interact with defined sequences and structures of selected tetraplexes at high specificity. Hence, an understanding of the role of a quadruplex domain in a regulatory region of a chosen gene will be attained only if a small molecule or protein interacts with the investigated tetraplex in an exclusive or near-exclusive fashion. Thus, the design and engineering of highly specific agents is a major challenge. The fine structural details of quadruplex formations of DNA sequences of interest, as learned through NMR analysis and X-ray crystal diffraction studies, already provide a necessary foundation for the development of tetraplex DNA interacting reagents of greater selectivity. This is illustrated by the pioneering attempts to select and construct proteins that interact with a particular quadruplex DNA structure at increased specificity (129-131).

Surpassing their utility as tools for the study of the biological functions of tetrahelical structures of DNA, agents that interact specifically with chosen quadruplex DNA structures could be used in the future as therapeutic agents. Indeed, rather remarkable first steps have already been made in using tetraplex-interacting low molecular size drugs to inhibit the expression of the *c-myc* gene or to modulate telomere extension (see refs 44-48 for reviews). Considering the significant advances that were made in the last two decades in the study of quadruplex DNA and its interacting proteins and small molecules, it is realistic to expect that similar or greater progress will be made in the foreseeable future.

6. ACKNOWLEDGEMENTS

Research in the author's laboratory is supported by grants from the Israel Science Foundation, the United States-Israel Binational Science foundation and the Conquer Fragile X Foundation, USA

7. REFERENCES

1. Garcia, H. G., P. Grayson, L. Han, M. Inamdar, J. Kondev, P. C. Nelson, R. Phillips, J. Widom & P. A. Wiggins: Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity. *Biopolymers*, 85, 115-130 (2007)
2. Frank-Kamenetskii, M. D. & S. M. Mirkin: Triplex DNA structures. *Annu Rev Biochem*, 64, 65-95 (1995)
3. Rogers, F. A., J. A. Lloyd & P. M. Glazer: Triplex-forming oligonucleotides as potential tools for modulation of gene expression. *Curr Med Chem Anticancer Agents*, 5, 319-26 (2005)
4. Rich, A. & S. Zhang: Timeline: Z-DNA: the long road to biological function. *Nat Rev Genet*, 4, 566-72 (2003)
5. Phan, A. T., V. Kuryavyi & D. J. Patel: DNA architecture: from G to Z. *Curr Opin Struct Biol*, 16, 288-98 (2006)
6. Sen, D. & W. Gilbert: Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, 334, 364-6 (1988)
7. Sundquist, W. I. & A. Klug: Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature*, 342, 825-9 (1989)
8. Williamson, J. R., M. K. Raghuraman & T. R. Cech: Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, 59, 871-80 (1989)
9. Neidle, S. & G. N. Parkinson: The structure of telomeric DNA. *Curr Opin Struct Biol*, 13, 275-83 (2003)
10. Phan, A. T. & J. L. Mergny: Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. *Nucleic Acids Res*, 30, 4618-25 (2002)
11. Kang, C., X. Zhang, R. Ratliff, R. Moyzis & A. Rich: Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature*, 356, 126-31 (1992)
12. Smith, F. W. & J. Feigon: Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature*, 356, 164-8 (1992)
13. Wang, Y. & D. J. Patel: Solution structure of a parallel-stranded G-quadruplex DNA. *J Mol Biol*, 234, 1171-83 (1993)
14. Kettani, A., R. A. Kumar & D. J. Patel: Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J Mol Biol*, 254, 638-56 (1995)
15. Parkinson, G. N., M. P. Lee & S. Neidle: Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, 417, 876-80 (2002)
16. Keniry, M. A.: Quadruplex structures in nucleic acids. *Biopolymers*, 56, 123-46 (2000)
17. Simonsson, T.: G-quadruplex DNA structures--variations on a theme. *Biol Chem*, 382, 621-8 (2001)
18. Hazel, P., G. N. Parkinson & S. Neidle: Topology variation and loop structural homology in crystal and simulated structures of a bimolecular DNA quadruplex. *J Am Chem Soc*, 128, 5480-7 (2006)
19. Phan, A. T., Kuryavyi, V., Luu, K.N. and Patel, D.J.: Structural Diversity of G-Quadruplex Scaffolds. In: Quadruplex Nucleic Acids. Ed: S. a. B. Neidle, S. RSC Publishing, London (2006)
20. Rawal, P., V. B. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das & S. Chowdhury: Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res*, 16, 644-55 (2006)
21. Todd, A. K., M. Johnston & S. Neidle: Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res*, 33, 2901-7 (2005)
22. Eddy, J. & N. Maizels: Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res*, 34, 3887-96 (2006)
23. Brown, B. A., 2nd, Y. Li, J. C. Brown, C. C. Hardin, J. F. Roberts, S. C. Pelsue & L. D. Shultz: Isolation and characterization of a monoclonal anti-quadruplex DNA antibody from autoimmune "viable motheaten" mice. *Biochemistry*, 37, 16325-37 (1998)
24. Brown, J. C., B. A. Brown, 2nd, Y. Li & C. C. Hardin: Construction and characterization of a quadruplex DNA selective single-chain autoantibody from a viable motheaten mouse hybridoma with homology to telomeric DNA binding proteins. *Biochemistry*, 37, 16338-48 (1998)
25. Schaffitzel, C., I. Berger, J. Postberg, J. Hanes, H. J. Lipps & A. Pluckthun: *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with

Stylonychia lemnae macronuclei. *Proc Natl Acad Sci U S A*, 98, 8572-7 (2001)
26. Jonsson, F., J. Postberg, C. Schaffitzel & H. J. Lipps: Organization of the macronuclear gene-sized pieces of stichotrichous ciliates into a higher order structure via telomere-matrix interactions. *Chromosome Res*, 10, 445-53 (2002)
27. Paeschke, K., T. Simonsson, J. Postberg, D. Rhodes & H. J. Lipps: Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*. *Nat Struct Mol Biol*, 12, 847-54 (2005)
28. Chang, C. C., I. C. Kuo, I. F. Ling, C. T. Chen, H. C. Chen, P. J. Lou, J. J. Lin & T. C. Chang: Detection of quadruplex DNA structures in human telomeres by a fluorescent carbazole derivative. *Anal Chem*, 76, 4490-4 (2004)
29. Duquette, M. L., P. Handa, J. A. Vincent, A. F. Taylor & N. Maizels: Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev*, 18, 1618-29 (2004)
30. Dexheimer, T. S., Fry, M., and Hurley, L. H.: DNA Quadruplexes and Gene Regulation. In: Quadruplex Nucleic Acids. Ed: S. Neidle, and Balasubramanian, S. RSC Publishing, London (2006)
31. Simonsson, T., P. Pecinka & M. Kubista: DNA tetraplex formation in the control region of *c-myc*. *Nucleic Acids Res*, 26, 1167-72 (1998)
32. Siddiqui-Jain, A., C. L. Grand, D. J. Bearss & L. H. Hurley: 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 U S A*, 99, 11593-8 (2002)
33. Phan, A. T., Y. S. Modi & D. J. Patel: Propeller-type parallel-stranded G-quadruplexes in the human *c-myc* promoter. *J Am Chem Soc*, 126, 8710-6 (2004)
34. Lemarteleur, T., D. Gomez, R. Paterski, E. Mandine, P. Mailliet & J. F. Riou: Stabilization of the *c-myc* gene promoter quadruplex by specific ligands' inhibitors of telomerase. *Biochem Biophys Res Commun*, 323, 802-8 (2004)
35. Seenisamy, J., E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain & L. H. Hurley: The dynamic character of the G-quadruplex element in the *c-MYC* promoter and modification by TMPyP4. *J Am Chem Soc*, 126, 8702-9 (2004)
36. Kumar, P., A. Verma, S. Maiti, R. Gargallo & S. Chowdhury: Tetraplex DNA transitions within the human *c-myc* promoter detected by multivariate curve resolution of fluorescence resonance energy transfer. *Biochemistry*, 44, 16426-34 (2005)
37. Halder, K. & S. Chowdhury: Kinetic resolution of bimolecular hybridization versus intramolecular folding in nucleic acids by surface plasmon resonance: application to G-quadruplex/duplex competition in human *c-myc* promoter. *Nucleic Acids Res*, 33, 4466-74 (2005)
38. Walsh, K. & A. Gualberto: MyoD binds to the guanine tetrad nucleic acid structure. *J Biol Chem*, 267, 13714-8 (1992)
39. Yafe, A., S. Etzioni, P. Weisman-Shomer & M. Fry: Formation and properties of hairpin and tetraplex structures of guanine-rich regulatory sequences of muscle-specific genes. *Nucleic Acids Res*, 33, 2887-900 (2005)
40. Etzioni, S., A. Yafe, S. Khateb, P. Weisman-Shomer, E. Bengal & M. Fry: Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. *J Biol Chem*, 280, 26805-12 (2005)
41. Zahler, A. M., J. R. Williamson, T. R. Cech & D. M. Prescott: Inhibition of telomerase by G-quartet DNA structures. *Nature*, 350, 718-20 (1991)
42. Sun, D., C. C. Lopez-Guajardo, J. Quada, L. H. Hurley & D. D. Von Hoff: Regulation of catalytic activity and processivity of human telomerase. *Biochemistry*, 38, 4037-44 (1999)
43. Oganesian, L., I. K. Moon, T. M. Bryan & M. B. Jarstfer: Extension of G-quadruplex DNA by ciliate telomerase. *Embo J*, 25, 1148-59 (2006)
44. Neidle, S., R. J. Harrison, A. P. Reszka & M. A. Read: Structure-activity relationships among guanine-quadruplex telomerase inhibitors. *Pharmacol Ther*, 85, 133-9 (2000)
45. Neidle, S. & M. A. Read: G-quadruplexes as therapeutic targets. *Biopolymers*, 56, 195-208 (2000)
46. Kerwin, S. M.: G-Quadruplex DNA as a target for drug design. *Curr Pharm Des*, 6, 441-78 (2000)
47. Cairns, D., R. J. Anderson, P. J. Perry & T. C. Jenkins: Design of telomerase inhibitors for the treatment of cancer. *Curr Pharm Des*, 8, 2491-504 (2002)
48. Pendino, F., I. Tarkanyi, C. Dudognon, J. Hillion, M. Lanotte, J. Aradi & E. Segal-Bendirdjian: Telomeres and telomerase: Pharmacological targets for new anticancer strategies? *Curr Cancer Drug Targets*, 6, 147-80 (2006)
49. Yanez, G. H., S. J. Khan, A. M. Locovei, I. M. Pedroso & T. M. Fletcher: DNA structure-dependent recruitment of telomeric proteins to single-stranded/double-stranded DNA junctions. *Biochem Biophys Res Commun*, 328, 49-56 (2005)
50. Shukla, A. K. & K. B. Roy: Rec A-independent homologous recombination induced by a putative fold-back tetraplex DNA. *Biol Chem*, 387, 251-6 (2006)
51. Anuradha, S. & K. Muniyappa: Molecular aspects of meiotic chromosome synapsis and recombination. *Prog Nucleic Acid Res Mol Biol*, 79, 49-132 (2005)
52. Kironmai, K. M., K. Muniyappa, D. B. Friedman, N. M. Hollingsworth & B. Byers: DNA-binding activities of Hop1 protein, a synaptonemal complex component from *Saccharomyces cerevisiae*. *Mol Cell Biol*, 18, 1424-35 (1998)
53. Muniyappa, K., S. Anuradha & B. Byers: Yeast meiosis-specific protein Hop1 binds to G4 DNA and promotes its formation. *Mol Cell Biol*, 20, 1361-9 (2000)
54. Williams, M. & N. Maizels: LR1, a lipopolysaccharide-responsive factor with binding sites in the immunoglobulin switch regions and heavy-chain enhancer. *Genes Dev*, 5, 2353-61 (1991)
55. Williams, M., L. A. Hanakahi & N. Maizels: Purification and properties of LR1, an inducible DNA binding protein from mammalian B lymphocytes. *J Biol Chem*, 268, 13731-7 (1993)
56. Dempsey, L. A., H. Sun, L. A. Hanakahi & N. Maizels: 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, 1066-71 (1999)
57. Hanakahi, L. A., H. Sun & N. Maizels: High affinity interactions of nucleolin with G-G-paired rDNA. *J Biol Chem*, 274, 15908-12 (1999)

58. Duquette, M. L., P. Pham, M. F. Goodman & N. Maizels: AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. *Oncogene*, 24, 5791-8 (2005)
59. Larson, E. D., M. L. Duquette, W. J. Cummings, R. J. Streiff & N. Maizels: MutSalpha binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. *Curr Biol*, 15, 470-4 (2005)
60. Verkerk, A. J., M. Pieretti, J. S. Sutcliffe, Y. H. Fu, D. P. Kuhl, A. Pizzuti, O. Reiner, S. Richards, M. F. Victoria, F. P. Zhang, B. E. Eussen, G. J. B. van Ommen, L. A. J. Blonden, G. J. Riggins, J. L. Chastain, C. B. Kunst, H. Galjaard, C. T. Caskey, D. L. Nelson, B. A. Oostra & S. bT. Warren.: Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, 65, 905-14 (1991)
61. Pieretti, M., F. P. Zhang, Y. H. Fu, S. T. Warren, B. A. Oostra, C. T. Caskey & D. L. Nelson: Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell*, 66, 817-22 (1991)
62. Fu, Y. H., D. P. Kuhl, A. Pizzuti, M. Pieretti, J. S. Sutcliffe, S. Richards, A. J. Verkerk, J. J. Holden, R. G. Fenwick, Jr., S. T. Warren, B. A. Oostra, D. L. Nelson & C. T. Caskey.: Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, 67, 1047-58 (1991)
63. La Spada, A. R., E. M. Wilson, D. B. Lubahn, A. E. Harding & K. H. Fischbeck: Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, 352, 77-9 (1991)
64. La Spada, A. R., D. B. Roling, A. E. Harding, C. L. Warner, R. Spiegel, I. Hausmanowa-Petrusewicz, W. C. Yee & K. H. Fischbeck: Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nat Genet*, 2, 301-4 (1992)
65. Chen, X., S. V. Mariappan, P. Catasti, R. Ratliff, R. K. Moyzis, A. Laayoun, S. S. Smith, E. M. Bradbury & G. Gupta: Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc Natl Acad Sci U S A*, 92, 5199-203 (1995)
66. Nadel, Y., P. Weisman-Shomer & M. Fry: The fragile X syndrome single strand d $(CGG)_n$ nucleotide repeats readily fold back to form unimolecular hairpin structures. *J Biol Chem*, 270, 28970-7 (1995)
67. Fry, M. & L. A. Loeb: The fragile X syndrome d $(CGG)_n$ nucleotide repeats form a stable tetrahelical structure. *Proc Natl Acad Sci U S A*, 91, 4950-4 (1994)
68. Usdin, K. & K. J. Woodford: CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis *in vitro*. *Nucleic Acids Res*, 23, 4202-9 (1995)
69. Kamath-Loeb, A. S., L. A. Loeb, E. Johansson, P. M. Burgers & M. Fry: Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d $(CGG)_n$ trinucleotide repeat sequence. *J Biol Chem*, 276, 16439-46 (2001)
70. Pataskar, S. S., D. Dash & S. K. Brahmachari: Progressive myoclonus epilepsy [EPM1] repeat d $(CCCCGCCCCGCG)_n$ forms folded hairpin structures at physiological pH. *J Biomol Struct Dyn*, 19, 293-305 (2001)
71. Saha, T. & K. Usdin: Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. *FEBS Lett*, 491, 184-7 (2001)
72. Sundquist, W. I. & S. Heaphy: Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. *Proc Natl Acad Sci U S A*, 90, 3393-7 (1993)
73. Lyonnais, S., C. Hounsou, M. P. Teulade-Fichou, J. Jeusset, E. Le Cam & G. Mirambeau: G-quartets assembly within a G-rich DNA flap. A possible event at the center of the HIV-1 genome. *Nucleic Acids Res*, 30, 5276-83 (2002)
74. Lyonnais, S., R. J. Gorelick, J. L. Mergny, E. Le Cam & G. Mirambeau: G-quartets direct assembly of HIV-1 nucleocapsid protein along single-stranded DNA. *Nucleic Acids Res*, 31, 5754-63 (2003)
75. Kankia, B. I., G. Barany & K. Musier-Forsyth: Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein. *Nucleic Acids Res*, 33, 4395-403 (2005)
76. de Soultrait, V. R., P. Y. Lozach, R. Altmeyer, L. Tarrago-Litvak, S. Litvak & M. L. Andreola: DNA aptamers derived from HIV-1 RNase H inhibitors are strong anti-integrase agents. *J Mol Biol*, 324, 195-203 (2002)
77. Phan, A. T., V. Kuryavyi, J. B. Ma, A. Faure, M. L. Andreola & D. J. Patel: An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. *Proc Natl Acad Sci U S A*, 102, 634-9 (2005)
78. Chou, S. H., K. H. Chin & A. H. Wang: DNA aptamers as potential anti-HIV agents. *Trends Biochem Sci*, 30, 231-4 (2005)
79. Kang, S. G. & E. Henderson: Identification of non-telomeric G4-DNA binding proteins in human, *E. coli*, yeast, and *Arabidopsis*. *Mol Cells*, 14, 404-10 (2002)
80. Sarig, G., P. Weisman-Shomer, R. Erlitzki & M. Fry: Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes. *J Biol Chem*, 272, 4474-82 (1997)
81. Erlitzki, R. & M. Fry: Sequence-specific binding protein of single-stranded and unimolecular quadruplex telomeric DNA from rat hepatocytes. *J Biol Chem*, 272, 15881-90 (1997)
82. Weisman-Shomer, P. & M. Fry: Stabilization of tetrahelical DNA by the quadruplex DNA binding protein QUAD. *Biochem Biophys Res Commun*, 205, 305-11 (1994)
83. Uliel, L., P. Weisman-Shomer, H. Oren-Jazan, T. Newcomb, L. A. Loeb & M. Fry: Human Ku antigen tightly binds and stabilizes a tetrahelical form of the Fragile X syndrome d $(CGG)_n$ expanded sequence. *J Biol Chem*, 275, 33134-41 (2000)
84. Pearson, A. M., A. Rich & M. Krieger: Polynucleotide binding to macrophage scavenger receptors depends on the formation of base-quartet-stabilized four-stranded helices. *J Biol Chem*, 268, 3546-54 (1993)
85. Bock, L. C., L. C. Griffin, J. A. Latham, E. H. Vermaas & J. J. Toole: Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, 355, 564-6 (1992)

86. Macaya, R. F., P. Schultze, F. W. Smith, J. A. Roe & J. Feigon: Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc Natl Acad Sci U S A*, 90, 3745-9 (1993)
87. Fang, G. & T. R. Cech: The beta subunit of *Oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell*, 74, 875-85 (1993)
88. Fang, G. & T. R. Cech: Characterization of a G-quartet formation reaction promoted by the beta-subunit of the *Oxytricha* telomere-binding protein. *Biochemistry*, 32, 11646-57 (1993)
89. Fang, G. & T. R. Cech: *Oxytricha* telomere-binding protein: DNA-dependent dimerization of the alpha and beta subunits. *Proc Natl Acad Sci U S A*, 90, 6056-60 (1993)
90. Laporte, L. & G. J. Thomas, Jr.: Structural basis of DNA recognition and mechanism of quadruplex formation by the beta subunit of the *Oxytricha* telomere binding protein. *Biochemistry*, 37, 1327-35 (1998)
91. Giraldo, R. & D. Rhodes: The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA. *Embo J*, 13, 2411-20 (1994)
92. Giraldo, R., M. Suzuki, L. Chapman & D. Rhodes: Promotion of parallel DNA quadruplexes by a yeast telomere binding protein: a circular dichroism study. *Proc Natl Acad Sci U S A*, 91, 7658-62 (1994)
93. Anuradha, S. & K. Muniyappa: Meiosis-specific yeast Hop1 protein promotes synapsis of double-stranded DNA helices via the formation of guanine quartets. *Nucleic Acids Res*, 32, 2378-85 (2004)
94. Venczel, E. A. & D. Sen: Synapsable DNA. *J Mol Biol*, 257, 219-24 (1996)
95. Arimondo, P. B., J. F. Riou, J. L. Mergny, J. Tazi, J. S. Sun, T. Garestier & C. Helene: Interaction of human DNA topoisomerase I with G-quartet structures. *Nucleic Acids Res*, 28, 4832-8 (2000)
96. Marchand, C., P. Pourquier, G. S. Laco, N. Jing & Y. Pommier: Interaction of human nuclear topoisomerase I with guanosine quartet-forming and guanosine-rich single-stranded DNA and RNA oligonucleotides. *J Biol Chem*, 277, 8906-11 (2002)
97. Baldrich, E. & C. K. O'Sullivan: Ability of thrombin to act as molecular chaperone, inducing formation of quadruplex structure of thrombin-binding aptamer. *Anal Biochem*, 341, 194-7 (2005)
98. Sun, H., J. K. Karow, I. D. Hickson & N. Maizels: The Bloom's syndrome helicase unwinds G4 DNA. *J Biol Chem*, 273, 27587-92 (1998)
99. Crabbe, L., R. E. Verdun, C. I. Haggblom & J. Karlseder: Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, 306, 1951-3 (2004)
100. Opresko, P. L., M. Otterlei, J. Graakjaer, P. Bruheim, L. Dawut, S. Kolvraa, A. May, M. M. Seidman & V. A. Bohr: The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. *Mol Cell*, 14, 763-74 (2004)
101. Han, H., R. J. Bennett & L. H. Hurley: Inhibition of unwinding of G-quadruplex structures by Sgs1 helicase in the presence of N,N'-bis[2- (1-piperidino)ethyl]-3,4,9,10-perylenetetracarboxylic diimide, a G-quadruplex-interactive ligand. *Biochemistry*, 39, 9311-6 (2000)
102. Huber, M. D., D. C. Lee & N. Maizels: G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res*, 30, 3954-61 (2002)
103. Sharma, S., K. M. Doherty & R. M. Brosh, Jr.: DNA helicases as targets for anti-cancer drugs. *Curr Med Chem Anticancer Agents*, 5, 183-99 (2005)
104. Weisman-Shomer, P., Y. Naot & M. Fry: Tetrahelical forms of the fragile X syndrome expanded sequence d $(CGG)_n$ are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins. *J Biol Chem*, 275, 2231-8 (2000)
105. Kamada, S. & T. Miwa: A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene*, 119, 229-36 (1992)
106. Khateb, S., P. Weisman-Shomer, I. Hershco, L. A. Loeb & M. Fry: Destabilization of tetraplex structures of the fragile X repeat sequence $(CGG)_n$ is mediated by homolog-conserved domains in three members of the hnRNP family. *Nucleic Acids Res*, 32, 4145-54 (2004)
107. Eversole, A. & N. Maizels: *In vitro* properties of the conserved mammalian protein hnRNP D suggest a role in telomere maintenance. *Mol Cell Biol*, 20, 5425-32 (2000)
108. Enokizono, Y., Y. Konishi, K. Nagata, K. Ouhashi, S. Uesugi, F. Ishikawa & M. Katahira: Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D. *J Biol Chem*, 280, 18862-70 (2005)
109. Riva, S., C. Morandi, P. Tsoulfas, M. Pandolfo, G. Biamonti, B. Merrill, K. R. Williams, G. Multhaup, K. Beyreuther, H. Werr, H. Henrich & K. P. Schäfer.: Mammalian single-stranded DNA binding protein UP1 is derived from the hnRNP core protein A1. *Embo J*, 5, 2267-73 (1986)
110. Merrill, B. M., M. B. LoPresti, K. L. Stone & K. R. Williams: High pressure liquid chromatography purification of UP1 and UP2, two related single-stranded nucleic acid-binding proteins from calf thymus. *J Biol Chem*, 261, 878-83 (1986)
111. Fukuda, H., M. Katahira, N. Tsuchiya, Y. Enokizono, T. Sugimura, M. Nagao & H. Nakagama: Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. *Proc Natl Acad Sci U S A*, 99, 12685-90 (2002)
112. Fukuda, H., M. Katahira, E. Tanaka, Y. Enokizono, N. Tsuchiya, K. Higuchi, M. Nagao & H. Nakagama: Unfolding of higher DNA structures formed by the d (CGG) triplet repeat by UP1 protein. *Genes Cells*, 10, 953-62 (2005)
113. LaBranche, H., S. Dupuis, Y. Ben-David, M. R. Bani, R. J. Wellinger & B. Chabot: Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat Genet*, 19, 199-202 (1998)
114. Dallaire, F., S. Dupuis, S. Fiset & B. Chabot: Heterogeneous nuclear ribonucleoprotein A1 and UP1 protect mammalian telomeric repeats and modulate telomere replication *in vitro*. *J Biol Chem*, 275, 14509-16 (2000)
115. Qi, H. & V. A. Zakian: The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes Dev*, 14, 1777-88 (2000)

116. Pennock, E., K. Buckley & V. Lundblad: Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell*, 104, 387-96 (2001)
117. Lin, Y. C., J. W. Shih, C. L. Hsu & J. J. Lin: Binding and partial denaturing of G-quartet DNA by Cdc13p of *Saccharomyces cerevisiae*. *J Biol Chem*, 276, 47671-4 (2001)
118. Zaug, A. J., E. R. Podell & T. R. Cech: Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension *in vitro*. *Proc Natl Acad Sci U S A*, 102, 10864-9 (2005)
119. Smith, J., H. Zou & R. Rothstein: Characterization of genetic interactions with RFA1: the role of RPA in DNA replication and telomere maintenance. *Biochimie*, 82, 71-8 (2000)
120. Schramke, V., P. Luciano, V. Brevet, S. Guillot, Y. Corda, M. P. Longhese, E. Gilson & V. Geli: RPA regulates telomerase action by providing Est1p access to chromosome ends. *Nat Genet*, 36, 46-54 (2004)
121. Cohen, S., E. Jacob & H. Manor: Effects of single-stranded DNA binding proteins on primer extension by telomerase. *Biochim Biophys Acta*, 1679, 129-40 (2004)
122. Salas, T. R., I. Petruseva, O. Lavrik, A. Bourdoncle, J. L. Mergny, A. Favre & C. Saintome: Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res* (2006)
123. Liu, Z., J. D. Frantz, W. Gilbert & B. K. Tye: Identification and characterization of a nuclease activity specific for G4 tetrastranded DNA. *Proc Natl Acad Sci U S A*, 90, 3157-61 (1993)
124. Liu, Z. & W. Gilbert: The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex DNA: implication of *in vivo* functions for this novel DNA structure. *Cell*, 77, 1083-92 (1994)
125. Liu, Z., A. Lee & W. Gilbert: Gene disruption of a G4-DNA-dependent nuclease in yeast leads to cellular senescence and telomere shortening. *Proc Natl Acad Sci U S A*, 92, 6002-6 (1995)
126. Ghosal, G. & K. Muniyappa: 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, 4692-703 (2005)
127. Sun, H., A. Yabuki & N. Maizels: A human nuclease specific for G4 DNA. *Proc Natl Acad Sci U S A*, 98, 12444-9 (2001)
128. Bashkirov, V. I., H. Scherthan, J. A. Solinger, J. M. Buerstedde & W. D. Heyer: A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *J Cell Biol*, 136, 761-73 (1997)
129. Isalan, M., S. D. Patel, S. Balasubramanian & Y. Choo: Selection of zinc fingers that bind single-stranded telomeric DNA in the G-quadruplex conformation. *Biochemistry*, 40, 830-6 (2001)
130. Patel, S. D., M. Isalan, G. Gavory, S. Ladame, Y. Choo & S. Balasubramanian: Inhibition of human telomerase activity by an engineered zinc finger protein that binds G-quadruplexes. *Biochemistry*, 43, 13452-8 (2004)
131. Ladame, S., J. A. Schouten, J. Roldan, J. E. Redman, S. Neidle & S. Balasubramanian: Exploring the recognition of quadruplex DNA by an engineered Cys2-His2 zinc finger protein. *Biochemistry*, 45, 1393-9 (2006)
132. Simonsson, T., M. Pribylova & M. Vorlickova: A nuclease hypersensitive element in the human *c-myc* promoter adopts several distinct i-tetraplex structures. *Biochem Biophys Res Commun*, 278, 158-66 (2000)
133. Murchie, A. I. & D. M. Lilley: Retinoblastoma susceptibility genes contain 5' sequences with a high propensity to form guanine-tetrad structures. *Nucleic Acids Res*, 20, 49-53 (1992)
134. Xu, Y. & H. Sugiyama: Highly efficient photochemical 2'-deoxyribonolactone formation at the diagonal loop of a 5-iodouracil-containing antiparallel G-quartet. *J Am Chem Soc*, 126, 6274-9 (2004)
135. Xu, Y. & H. Sugiyama: Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res*, 34, 949-54 (2006)
136. Cogoi, S., F. Quadrifoglio & L. E. Xodo: G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. *Biochemistry*, 43, 2512-23 (2004)
137. Cogoi, S. & L. E. Xodo: G-quadruplex formation within the promoter of the *KRAS* proto-oncogene and its effect on transcription. *Nucleic Acids Res*, 34, 2536-49 (2006)
138. Sun, D., K. Guo, J. J. Rusche & L. H. Hurley: Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res*, 33, 6070-80 (2005)
139. De Armond, R., S. Wood, D. Sun, L. H. Hurley & S. W. Ebbinghaus: Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1alpha promoter. *Biochemistry*, 44, 16341-50 (2005)
140. Fernando, H., A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle & S. Balasubramanian: A conserved quadruplex motif located in a transcription activation site of the human *c-kit* oncogene. *Biochemistry*, 45, 7854-60 (2006)
141. Rankin, S., A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian & S. Neidle: Putative DNA quadruplex formation within the human *c-kit* oncogene. *J Am Chem Soc*, 127, 10584-9 (2005)
142. Dexheimer, T. S., D. Sun & L. H. Hurley: Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the *bcl-2* P1 promoter. *J Am Chem Soc*, 128, 5404-15 (2006)
143. Tasset, D. M., M. F. Kubik & W. Steiner: Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol*, 272, 688-98 (1997)
144. Kelly, J. A., J. Feigon & T. O. Yeates: Reconciliation of the X-ray and NMR structures of the thrombin-binding aptamer d (GGTTGGTGTGGTTGG). *J Mol Biol*, 256, 417-22 (1996)
145. Weisman-Shomer, P. & M. Fry: QUAD, a protein from hepatocyte chromatin that binds selectively to guanine-rich quadruplex DNA. *J Biol Chem*, 268, 3306-12 (1993)

146. Schierer, T. & E. Henderson: A protein from *Tetrahymena thermophila* that specifically binds parallel-stranded G4-DNA. *Biochemistry*, 33, 2240-6 (1994)
147. Frantz, J. D. & W. Gilbert: A novel yeast gene product, G4p1, with a specific affinity for quadruplex nucleic acids. *J Biol Chem*, 270, 20692-7 (1995)
148. Frantz, J. D. & W. Gilbert: A yeast gene product, G4p2, with a specific affinity for quadruplex nucleic acids. *J Biol Chem*, 270, 9413-9 (1995)
149. Hayashi, N. & S. Murakami: *STM1*, a gene which encodes a guanine quadruplex binding protein, interacts with CDC13 in *Saccharomyces cerevisiae*. *Mol Genet Genomics*, 267, 806-13 (2002)
150. Van Dyke, M. W., L. D. Nelson, R. G. Weilbaecher & D. V. Mehta: Stm1p, a G4 quadruplex and purine motif triplex nucleic acid-binding protein, interacts with ribosomes and subtelomeric Y' DNA in *Saccharomyces cerevisiae. J Biol Chem*, 279, 24323-33 (2004)
151. Sarig, G., P. Weisman-Shomer & M. Fry: Telomeric and tetraplex DNA binding properties of qTBP42: a homologue of the CArG box binding protein CBF-A. *Biochem Biophys Res Commun*, 237, 617-23 (1997)
152. Oliver, A. W., I. Bogdarina, E. Schroeder, I. A. Taylor & G. G. Kneale: Preferential binding of fd gene 5 protein to tetraplex nucleic acid structures. *J Mol Biol*, 301, 575-84 (2000)
153. Oliver, A. W. & G. G. Kneale: Structural characterization of DNA and RNA sequences recognized by the gene 5 protein of bacteriophage fd. *Biochem J*, 339 (Pt 3), 525-31 (1999)
154. Wen, J. D., C. W. Gray & D. M. Gray: SELEX selection of high-affinity oligonucleotides for bacteriophage Ff gene 5 protein. *Biochemistry*, 40, 9300-10 (2001)
155. Wen, J. D. & D. M. Gray: The Ff gene 5 single-stranded DNA-binding protein binds to the transiently folded form of an intramolecular G-quadruplex. *Biochemistry*, 41, 11438-48 (2002)
156. Tolstonog, G. V., G. Li, R. L. Shoeman & P. Traub: Interaction *in vitro* of type III intermediate filament proteins with higher order structures of single-stranded DNA, particularly with G-quadruplex DNA. *DNA Cell Biol*, 24, 85-110 (2005)
157. Harrington, C., Y. Lan & S. A. Akman: The identification and characterization of a G4-DNA resolvase activity. *J Biol Chem*, 272, 24631-6 (1997)
158. Vaughn, J. P., S. D. Creacy, E. D. Routh, C. Joyner-Butt, G. S. Jenkins, S. Pauli, Y. Nagamine & S. A. Akman: 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, 38117-20 (2005)
159. Baran, N., L. Pucshansky, Y. Marco, S. Benjamin & H. Manor: The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. *Nucleic Acids Res*, 25, 297-303 (1997)
160. Fry, M. & L. A. Loeb: Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d $(CGG)_n$. *J Biol Chem*, 274, 12797-802 (1999)
161. Mohaghegh, P., J. K. Karow, R. M. Brosh Jr, Jr., V. A. Bohr & I. D. Hickson: The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res*, 29, 2843-9 (2001)
162. Sun, H., R. J. Bennett & N. Maizels: The *Saccharomyces cerevisiae* Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res*, 27, 1978-84 (1999)
163. Weisman-Shomer, P., E. Cohen & M. Fry: Distinct domains in the CArG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d $(CGG)_n$. *Nucleic Acids Res*, 30, 3672-81 (2002)
164. Basnar, B., R. Elnathan & I. Willner: Following aptamer-thrombin binding by force measurements. *Anal Chem*, 78, 3638-42 (2006)
165. Hammond-Kosack, M. C., B. Dobrinski, R. Lurz, K. Docherty & M. W. Kilpatrick: The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res*, 20, 231-6 (1992)
166. Hammond-Kosack, M. C. & K. Docherty: A consensus repeat sequence from the human insulin gene linked polymorphic region adopts multiple quadriplex DNA structures *in vitro*. *FEBS Lett*, 301, 79-82 (1992)
167. Hammond-Kosack, M. C., M. W. Kilpatrick & K. Docherty: The human insulin gene-linked polymorphic region adopts a G-quartet structure in chromatin assembled *in vitro*. *J Mol Endocrinol*, 10, 1-6 (1993)
168. Woodford, K. J., R. M. Howell, & K. Usdin: A novel K^+-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *J Biol Chem*, 269, 27029-35 (1994)
169. Howell, R. M., K. J. Woodford, M. N. Weitzmann, & K. Usdin, The chicken beta-globin gene promoter forms a novel "cinched" tetrahelical structure. *J Biol Chem*, 271, 5208-14 (1996)

Abbreviations: G'4 DNA: unimolecular (intramolecular) quadruplex DNA; G'2 DNA: bimolecular (intermolecular) quadruplex DNA; G4 DNA: four-molecular (intermolecular) quadruplex DNA

Key Words: Tetraplex DNA; Quadruplex DNA; Tetrahelical DNA; Guanine quartets; Telomere DNA; DNA Recombination; Transcriptional regulation; Nucleotide expansion disorders, Review

Send correspondence to: Dr Michael Fry, Department of Biochemistry, Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, P.O.B. 9649, Bat Galim, Haifa 31096 Israel, Tel: 972-4-829-5328, Fax: 972-4-851-0735, E-mail: mickey@tx.technion.ac.il

http://www.bioscience.org/current/vol12.htm