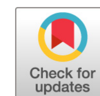
AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

MINIREVIEW



G-Quadruplex Structures in Bacteria: Biological Relevance and Potential as an Antimicrobial Target

Puja Yadav,[a] Nayun Kim,[b] Monika Kumari,[a] Shalini Verma,[a] Tarun Kumar Sharma,[c] Vikas Yadav,[d] Amit Kumar[e]

[a]Department of Microbiology, Central University of Haryana, Mahendergarh, Haryana, India

[b]Department of Microbiology and Molecular Genetics, The University of Texas Health Science Center at Houston, The University of Texas Graduate School of Biomedical Sciences, Houston, Texas, USA

[c]Translational Health Science and Technology Institute, Faridabad, Haryana, India

[d]School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

[e]Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Simrol, Indore, India

ABSTRACT DNA strands consisting of multiple runs of guanines can adopt a noncanonical, four-stranded DNA secondary structure known as G-quadruplex or G4 DNA. G4 DNA is thought to play an important role in transcriptional and translational regulation of genes, DNA replication, genome stability, and oncogene expression in eukaryotic genomes. In other organisms, including several bacterial pathogens and some plant species, the biological roles of G4 DNA and G4 RNA are starting to be explored. Recent investigations showed that G4 DNA and G4 RNA are generally conserved across plant species. *In silico* analyses of several bacterial genomes identified putative guanine-rich, G4 DNA-forming sequences in promoter regions. The sequences were particularly abundant in certain gene classes, suggesting that these highly diverse structures can be employed to regulate the expression of genes involved in secondary metabolite synthesis and signal transduction. Furthermore, in the pathogen *Mycobacterium tuberculosis*, the distribution of G4 motifs and their potential role in the regulation of gene transcription advocate for the use of G4 ligands to develop novel antitubercular therapies. In this review, we discuss the various roles of G4 structures in bacterial DNA and the application of G4 DNA as inhibitors or therapeutic agents to address bacterial pathogens.

KEYWORDS G-quadruplex, transcriptional and translational regulation, host-pathogen interaction, antigenic variation, homologous recombination, aptamers

In 1910, the ability of guanylic acid to form a gel at high concentrations was first reported (1). In 1960, fiber diffraction and biophysical methods revealed the formation of G-quartets via Hoogsteen-bonded guanines (2–5). Later, formation of more complex structures involving multiple G-quartets under nearly physiological conditions was observed *in vitro* with G-rich sequences from telomeres or immunoglobulin switching regions (6, 7). Correspondingly, guanine-rich sequences, which are widespread in DNA and RNA, possess an inherent propensity to fold into a four-stranded structure known as a guanine quadruplex, G-quadruplex, or G4 (8, 9). To form this secondary structure, multiple runs of guanines (at least four runs of guanines) are required. Four guanine bases are bound together by Hoogsteen bonding and present a square planar formation, making one G-quartet or G-tetrad (10). Two or more G-quartets stack on top of each other to form a G-quadruplex.

Based on the number of guanines present in each run or tract involved in quartet formation, G4 structures can be categorized into different types. G-quadruplexes with two or three guanine bases in each, which are able to form two or three G-quartets, respectively, are referred to as G2 or G3 type. Similarly, four G-quartets, consisting of four

Citation Yadav P, Kim N, Kumari M, Verma S, Sharma TK, Yadav V, Kumar A. 2021. G-quadruplex structures in bacteria: biological relevance and potential as an antimicrobial target. J Bacteriol 203:e00577-20. https://doi.org/10.1128/JB.00577-20.

Editor Thomas J. Silhavy, Princeton University

Copyright © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Puja Yadav, pujayadav@cuh.ac.in.

Accepted manuscript posted online 1 March 2021

Published 8 June 2021

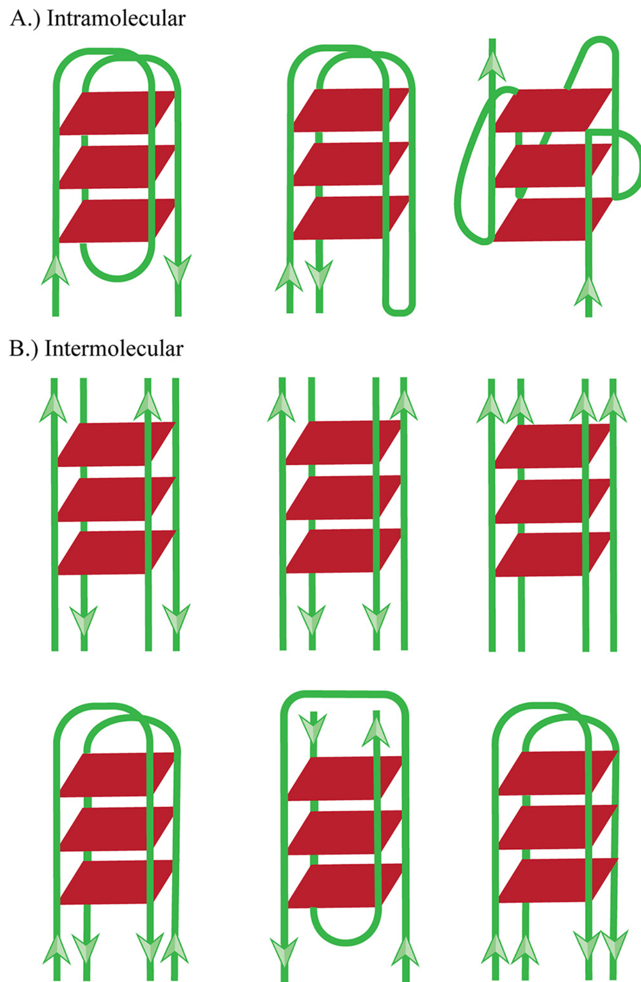


FIG 1 Schematic categorization of G4 DNA based on DNA strands. (A) Intramolecular type of G4 DNA (involving only one DNA strand in the formation of the quadruplex structure). (B) Intermolecular type of G4 DNA (involving two or more DNA strands in the formation of the quadruplex structure). The direction of the DNA strands is indicated by arrowheads.

guanine bases in each run stacked on top of each other, form a G4-type G-quadruplex. The guanine bases that are involved in quartet formation form the stem of the G4 structure, whereas the intervening sequences form loops of varying lengths and nucleotide composition. In addition, the presence of monovalent cations such as Na^+ and K^+ contributes to the folding and stability of these secondary structures (11). Depending on the number of DNA strands involved, G-quadruplexes can be categorized as intramolecular (a single DNA strand forming a G4) or intermolecular (multiple DNA strands involved in forming a G4) (12, 13) (Fig. 1). Based on the relative orientation of the DNA strands involved, a G-quadruplex is termed "parallel" if all strands are in the same orientation or "antiparallel" if one strand has a 5′ to 3′ direction, opposite the other DNA strand (14) (Fig. 1).

For *in silico* prediction of putative G-quadruplex-forming sequences (GQFSs) in the genome, there are several easily accessible, widely used algorithms, such as Quadparser, G4 calculator, QGRS Mapper, and QuadBase (15–18). Analyses of G4 motif distributions in human, *Saccharomyces cerevisiae*, and a number of prokaryotic genomes confirmed evolutionary conservation of the G4 motifs and their enrichment in certain functional regions (19–28). In a wide range of organisms, association of G4 motifs with specific genome features demonstrates that G4 DNA has *in vivo* functions that are under

evolutionary constraint. Also, in plant species, including *Arabidopsis thaliana*, *Zea mays*, *Oryza japonicum*, and *Oryza sativa*, the abundance of GQFSs upstream or downstream of the transcription start site of a gene suggests their role in the regulation of gene expression (22, 23). Gene ontology analyses have predicted the regulatory role of G-quadruplexes during cellular responses to DNA damage and other internal and external cues, such as sugar availability and metabolic and energy status (24).

In eukaryotes, G4s are particularly abundant within regulatory elements of both genomic DNA and mRNAs (20, 25). In addition, G4s are enriched at chromosomal telomeric sequences (26–28), promoter regions and transcription start sites (29–33), splice sites (34), and 5′ and 3′ untranslated regions (UTRs) of mRNAs (35–37). The presence of G4 at 5′ UTRs in mRNA suggests a regulatory role of G4s in gene expression, through either inhibiting or promoting translation (38), whereases G4s within 3′ UTRs of mRNA are hypothesized to obstruct translation and to regulate polyadenylation, as well as subcellular localization of mRNAs (39–42). At the level of transcription, G4 DNA appears to play a regulatory role in a position- and orientation-dependent manner (see Fig. 4) (43, 44). G-quadruplexes are also present in the promoter regions of numerous proto-oncogenes, including *bcl-2* (45), c-*myc* (29), *c-kit* (46, 47), *c-myb* (48), *VEGF* (49), *KRAS* (50, 51), and *HIF-1* (52), which further suggests their role in the development and progression of cancer.

The study of G4 DNA has emerged in the forefront of research because of its proposed role in several biological functions, ranging from physiology to pathology, in various forms of life, including humans, bacteria, protozoans, viruses, and plants (24). The role of G-quadruplexes in recombination and replication has been implicated and generally appreciated in recent years, although it is still relatively less understood (53–56). The significance of G4 DNA has been enhanced by recent technological advances allowing *in vivo* detection and verification of the presence of G4 DNA. In particular, the G4 DNA-specific antibodies 1H6 and BG4 emerged as valuable tools for studying G4 DNA in cells (57–59). The presence and formation of G4 structures in human and murine cells were first determined based on the intensity of 1H6 nuclear staining, which becomes stronger upon stabilization of G4 structures by treatment with agents, as well as in cells lacking a G4 DNA-specific helicase (FANCJ) (57). Additionally, DNA and RNA G-quadruplex structures within human cells were visualized by using a G-quadruplex-structure-specific antibody (BG4) (58, 59). Using the BG4 antibody, the same laboratory identified G4 DNA-containing genomic loci in human breast adenocarcinoma cells by chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) (60). In prokaryotes, the biological significance of G4 structures is yet to be significantly appreciated. Only sparse reports are available regarding the possible roles of G4 structures in prokaryotes in the context of either gene regulation or genome instability (61). In the current review, we attempt to highlight the presence, importance, and biological role of G4 DNA in bacterial genomes, along with the application of G4 aptamers in targeting bacterial pathogens.

G4 DNA-FORMING SEQUENCES IN BACTERIA

In yeast and mammalian systems, bioinformatic analyses of the genomes identified putative G4 sequences that were further confirmed by identification and mapping using the G4-specific antibodies. Those studies have provided information about the biology of G4s in transcriptional regulatory regions, which suggests that they are involved in a broad range of biological processes (24). However, information regarding the dynamics of G-quadruplexes in bacterial cells remained very sparse until recently, when *in silico* analyses of bacterial genomes gave insight about the occurrence of these secondary structures in the genomes, specifically in genes that are involved in virulence and pathogenesis. These studies in microbes together propose a regulatory role for G4s in gene expression.

In 2006, *in silico* analyses showed the enrichment of G4s in putative regulatory regions across 18 prokaryotic genomes, including that of *Mycobacterium tuberculosis*

(20). Those computational studies highlighted the need to experimentally confirm the *in vivo* existence of G4 DNA and further to evaluate the function(s) of G4s in biologically relevant contexts to gain insight about their potential as therapeutic targets.

In the past few years, genome-wide analyses of the occurrence and distribution of GQFSs in the genomes of bacterial human pathogens have become a prime interest, owing to their roles in survival, propagation, and pathogenesis (20, 62–64). The discovery of conserved G-quadruplex structures in various bacterial strains can lead to the development of an effective therapy for drug-resistant and susceptible strain alike to manage the infections caused by bacterial pathogens.

Computational analyses of 18 bacterial genomes identified enrichment of GQFSs in regulatory regions within 200 bp upstream of coding regions (20, 65). Moreover, cluster analyses of another 19 well-annotated bacterial species showed that gene groups with G4 DNA are distributed nonrandomly and are associated with specific functions (66). In *Escherichia coli*, G4 DNA is associated with target sites of the global regulators FIS and Lrp and the sigma factor RpoD (σ^{70}) (20). Overall, these *in silico* analyses found that, in multiple bacterial systems, GQFSs are enriched in the promoter regions of genes that are associated with transcription, secondary metabolite biosynthesis, and signal transduction (20). Such distribution suggests a specific regulatory role for GQFSs in prokaryotes. Additionally, GQFSs were present between the genes that encode two-component system response regulators, pyrophosphokinases, and diguanylate cyclases, with an average distance of $\sim$84 bp from the start of the following coding regions (20).

In general, GQFS frequencies were correlated with the GC contents of the genomes. An exception to this rule was observed in the phylum *Proteobacteria*, which includes a wide variety of pathogens such as *Escherichia*, *Salmonella*, *Vibrio*, *Helicobacter*, *Yersinia*, and *Legionellales* species (67). The genomes of these microbes have high GC contents ($>$50%) but exhibit relatively lower GQFS densities. In thermophiles, however, higher frequencies of GQFSs were observed in their genomes than expected by random chance (67). The highest density of GQFSs was found in the *Deinococcus-Thermus* phylum, with genomes with high GC contents ($>$60%), followed by *Actinobacteria* (67, 68). Also, in the *Deinococcus-Thermus* phylum, the GQFS distribution has evolved to be uniquely different from those of other closely related species. For instance, G-rich sequences were found to be randomly distributed in the order *Thermales*, whereas GQFSs were abundant in the order *Deinococcales* and enriched at locations proximal to the transcription start sites of genes. The occurrence and favored enrichment of GQFSs around regulatory regions such as those near or around the transcription start sites in the genomes of stress-resistant bacteria of *Deinococcales* suggest their role in gene regulation. In contrast, in other thermophiles such as *Hadesarchaea archaeon*, a random distribution of frequently occurring GQFSs may be associated with their extremophilic life at high temperatures (67). However, further work would be necessary to sufficiently compare G4 densities in acidophilic, thermophilic, halophilic, and psychrophilic organisms. Future studies could focus on the specific archaeal GQFS loci, for example, genes with high phylogenic conservation (69), answering questions regarding whether the G4-dependent regulatory mechanism is universal or confined to a given domain.

REGULATORY ROLE OF G4 DNA IN BACTERIA

In bacterial genomes, analyses of relationships between functional classes of genes and the presence of G-quadruplexes in gene promoters were conducted both at a genome-wide level and in individual genes. The occurrence of G4 DNA in the promoters of key genes is likely to impart species- and gene-specific functional attributes and may play a critical role in a certain cellular pathway (66).

(i) Functional variability of G-quadruplexes within promoter regions. As mentioned earlier, the order *Deinococcales*, which includes stress-resistant bacteria, exhibits favored enrichment of GQFSs around regulatory regions. *Deinococcus radiodurans* and *Deinococcus geothermalis* are able to withstand high levels of radiation (66, 70, 71). *In*

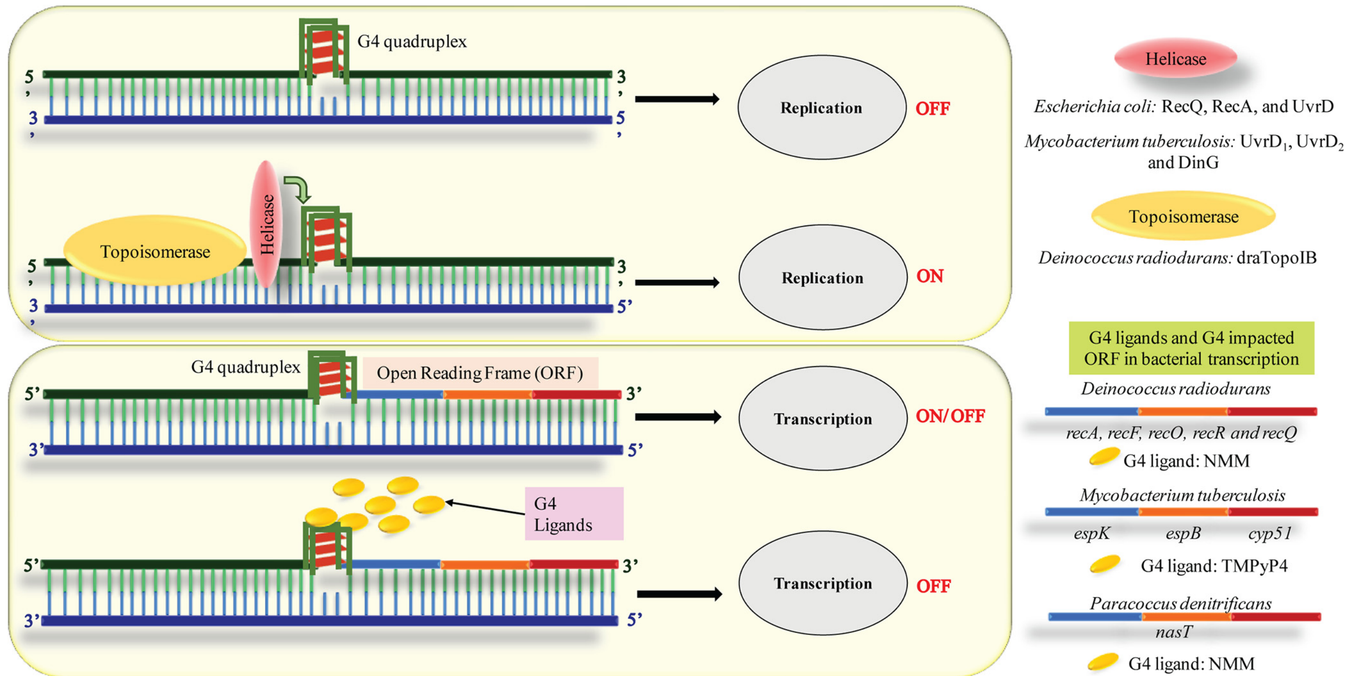


FIG 2 Graphical diagram representing the regulatory role of G-quadruplex sequences in replication and transcription (by either inhibiting or promoting expression). The presence of secondary structure G4 DNA stalls DNA replication in many bacterial species, while the destabilization of G4 DNA in the presence of helicase enzymes Rec Q, Rec A, and UvrD (studied in *E. coli*) and UvrD1, UvrD2, and DinG (investigated in *M. tuberculosis*) and topoisomerase DraTopoIB of *M. tuberculosis* results into the resumption of stalled DNA replication. The transcriptional regulation of many genes is observed to be either continued or hindered by G4 DNA in many bacteria. With the use of certain ligands such as NMM and TMPyP4, G4 DNA is stabilized and transcription of genes such as *recA*, *recO*, *recF*, *recR*, and *recQ* in *D. radiodurans*, *espK*, *espB*, and *cyp51* in *M. tuberculosis*, and *nasT* in *P. denitrificans* is observed to be obstructed.

silico analyses of the genome of *D. radiodurans* confirmed enrichment of GQFSs in promoters of key genes, strongly suggesting a regulatory role of G4 in the response to radiation (66). For instance, the presence of G4 in the promoter regions of *recA*, *recF*, *recO*, *recR*, and *recQ* genes, which are important components of the RecF recombinational DNA double-strand-break repair pathway, suggested that promoter G4s influence selective functions and regulate expression of the DNA repair proteins required for radioresistance. In an effort to determine the functional relevance of G4 DNA in microorganisms, intracellular quadruplex-binding ligands such as *N*-methyl mesoporphyrin (NMM) have proved to be very useful tools. Stabilization of the G4 structure by NMM repressed the gamma radiation response of many DNA repair genes and attenuated the radioresistance of *D. radiodurans* and *D. geothermalis* (66) (Fig. 2 and 3). However, the precise mechanism underlying the role of G4 in radioresistance is not yet known and would be worth further investigating (71). In a Gram-negative soil bacterium, *Paracoccus denitrificans*, G-quadruplexes were identified in the promoter region that activates transcription of the *nas* genes (assimilatory nitrate/nitrite reductase system) (72). Stabilization of this secondary structure using a G4 ligand resulted in attenuation of gene transcription, indicating G4-mediated regulation of bacterial nitrate/nitrite metabolism (Fig. 2). In *M. tuberculosis*, G4-mediated positive transcriptional regulation has been demonstrated for *zwf1* (glucose-6-phosphate dehydrogenase 1), *clpx* (ATP-dependent Clp protease), *mosR* (oxidation-sensing regulator transcription factor), and *ndhA* (membrane NADH dehydrogenase) (73). Additionally, in a recent study of G-quadruplexes in *M. tuberculosis*, G4 ligands inhibited growth with MICs in the low micromolar range (64). These data support the *in vivo* formation and functional relevance of G4 DNA in *M. tuberculosis* and their potential role in the regulation of gene transcription, and they suggest the use of G4 ligands for developing novel antitubercular agents (61).

(ii) Significant role of G-quadruplexes in essential genes. A study of all 160 representative *M. tuberculosis* genomes found in the NCBI database revealed the occurrence

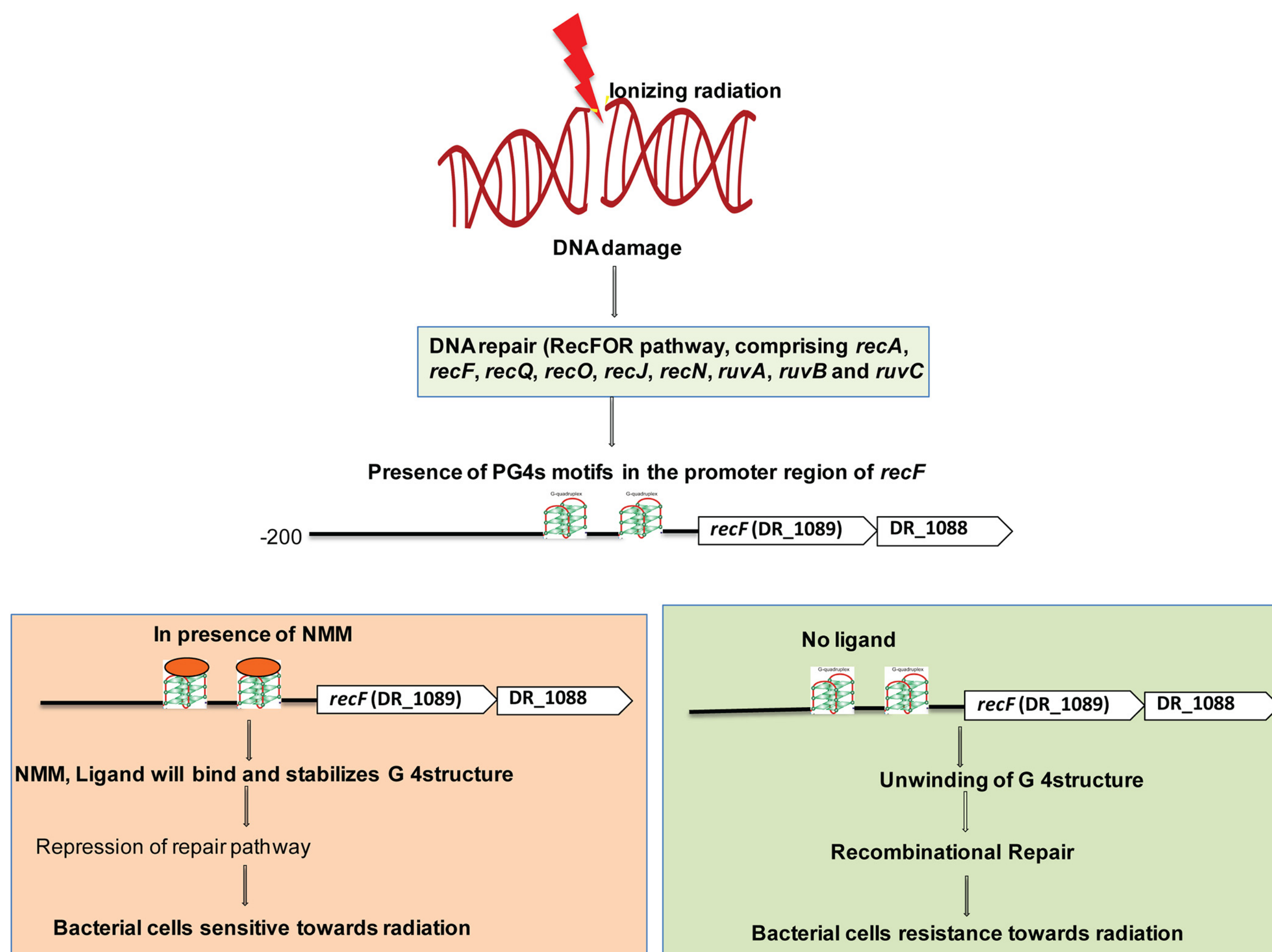


FIG 3 G-quadruplex-mediated regulatory mechanism of radioresistance in *D. radiodurans*.

of highly conserved G-quadruplexes in three essential genes, i.e., *espK*, *espB*, and *cyp51* (64). Functional assays (e.g., polymerase inhibition and quantitative reverse transcription-PCR assays) and biophysical characterization (e.g., NMR, circular dichroism, and gel electrophoresis) of these G-quadruplexes in the presence of the G4-stabilizing agent tetra-(*N*-methyl-4-pyridyl)porphyrin (TMPyP4) revealed the formation of stable intramolecular parallel G4 DNA structures. In this experiment, it was also shown that TMPyP4 inhibited the intracellular transcription of these G-quadruplex-containing genes (Fig. 2). The occurrence of these highly conserved G-quadruplexes suggests their conserved role in bacterial survival and pathogenesis and demonstrates that these G-quadruplexes can be considered as potential drug targets for the development of effective antituberculosis therapeutics.

In *Streptococcus pneumoniae*, detailed analysis of G4 structures in all 39 completely sequenced strains available revealed the presence of highly conserved, putative GQFSs in three essential genes (*recA*, *pmrA* and *hsdS*) (74). All three genes are virulence genes that are important in recombination repair, drug efflux, and host-pathogen interactions. The conservation of G-quadruplex motifs in these three virulence genes among various *S. pneumoniae* strains promises a formulation of a universal drug target against *S. pneumoniae* infection.

(iii) Impact of strand orientation and position of G4 on bacterial gene expression. In eukaryotic systems, several reports indicate strand-orientation-dependent effects of G4 DNA in regulating gene expression at the transcriptional or translational level (75). Since major differences in transcription and translation

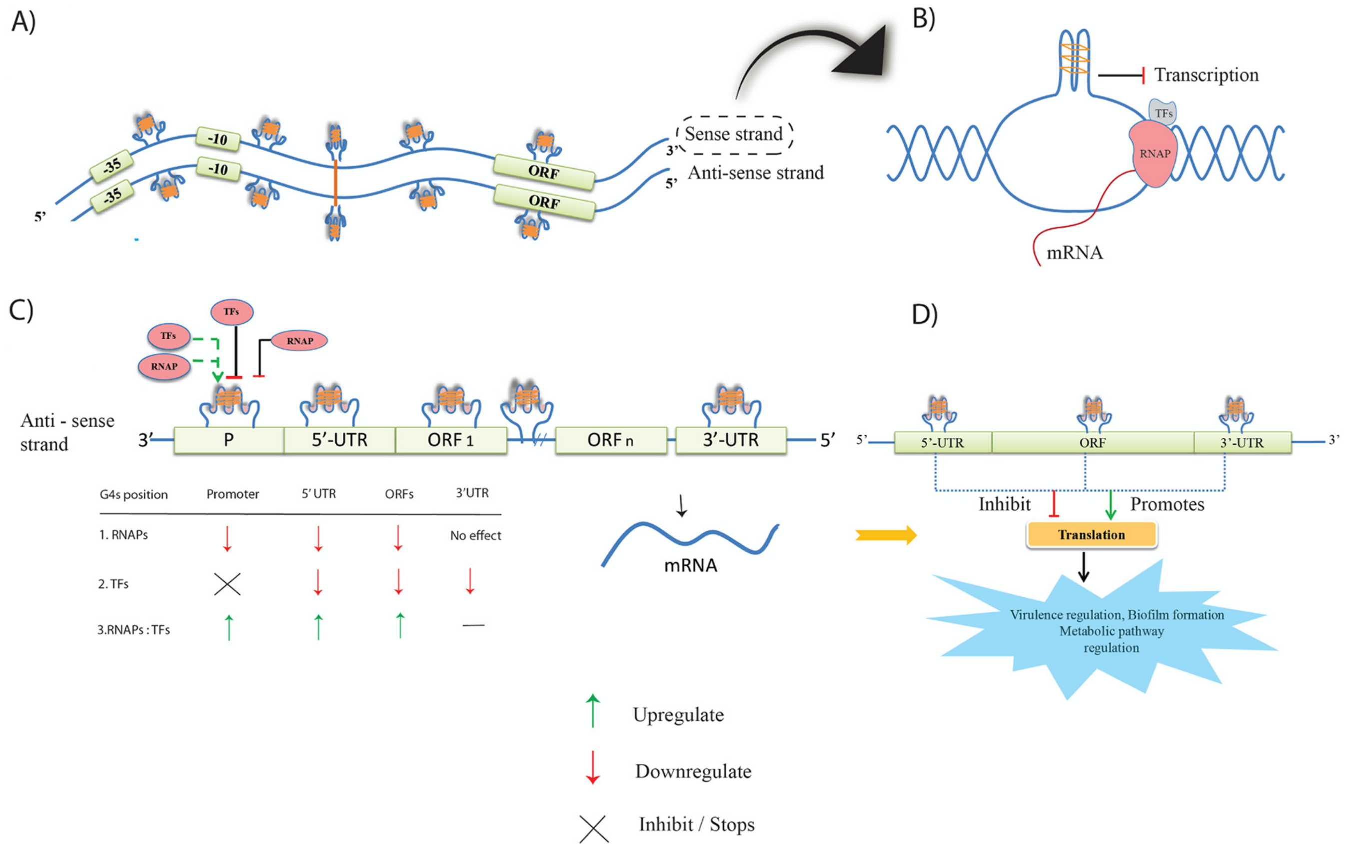


FIG 4 Effects of G-quadruplexes with respect to positioning and strand orientation. (A) G4 DNA formed in different regions on the sense and antisense strands. (B) The presence of G4 quadruplexes on the sense strand leads to inhibition of gene expression by formation of a DNA-RNA hybrid. (C) In the presence of these secondary structures on the antisense strand, both mRNA expression and inhibition occur. In the presence of transcription factors (TFs) and RNAP in the upstream or downstream promoter region, near the 5′ UTR of mRNA, or within genes, mRNA expression is downregulated; the presence of a G-quadruplex in the 3′ UTR of mRNA shows no significant effect. RNAPs can also upregulate gene expression. (D) G4 sequences present in the 5′ UTR of mRNA, in the gene body, within genes, and in the 3′ UTR of mRNA can either inhibit, upregulate, or downregulate protein expression. P, promoter.

machinery exist between prokaryotes and eukaryotes, data generated in eukaryotic systems regarding the orientation-dependent effect of G4 DNA cannot be readily applied to gene regulation in prokaryotic systems. Seminal work performed by two independent groups to understand the functional significance of G4 in bacterial biology and the impact of G4 orientation on gene regulation in bacterial systems has been summarized in Fig. 4 (20, 44). A key question regarding the G4-mediated regulation of gene expression in bacterial systems was asked by Holder and Hartig using *E. coli* as a model system (44). To check the positional and strand orientation effects of G-quadruplexes on gene expression, a reporter gene construct containing a GQFS at various positions inserted into the promoter, 5′ UTR, or 3′ UTR was developed. The result of this experiment showed that the presence of G4 sequences within the promoter and in proximity to the ribosome-binding site exhibited strong inhibitory effects on gene expression, which led to the conclusion that the exact position and strand orientation of these secondary structures significantly influence the transcription and translation machinery. In bacterial systems, G4-mediated bacterial gene regulation occurs at both the transcriptional level and the translational level, as insertion of G-rich sequences on the antisense strand of the core promoter resulted in a decrease in gene expression at the transcriptional level.

Binding of the RNA polymerase (RNAP) to the promoter region is essential for initiation of transcription. The sigma70 factor of the RNAP, which binds to the −10 and −35 regions, plays an important role in promoter recognition and subsequent unwinding of double-stranded DNA (dsDNA) or promoter melting (76). After initiation of transcription, a 10-nucleotide transcript is generated, which causes release of the sigma70

factor from the promoter and initiation of the elongation phase (77, 78). Adoption of the DNA secondary structure by G-rich sequences in the promoter regions of the antisense strand hinders the binding of sigma70 factor and thus inhibits transcription; in the downstream region of the promoter, RNAP actively separates the dsDNA, which in turn would facilitate the formation of G4 DNA. The antisense strand would be in a single-stranded confirmation, supporting and helping to separate the dsDNA, which would contribute to increases in gene expression (77, 78).

Moreover, the occurrence of G-rich sequences in the open reading frame (ORF) region of the gene causes inhibition of initiation of translation and elongation in *E. coli* (79). *In silico* analysis of the *E. coli* K-12 genome led to the identification of 46 GQFSs that coincide with the Shine-Delgarno (SD) sequences on the coding strand. Formation of these secondary structures at SD regions putatively complicates the binding of ribosome machinery to the SD region and thereby decreases protein expression (80). An experiment employing an artificial reporter showed that G4 DNA located at the ribosomal binding site within mRNA inhibited protein expression in *E. coli* and the level of repression was dependent on the thermodynamic stability of the G4 structures (80). A more recent work demonstrated that the presence of G-quadruplexes in mRNA can lead to ribosome stalling and a −1 ribosomal frameshift (81).

Analogous to the role of G4 DNA in plants, where G4 DNA regulates the expression of genes involved in several pathophysiological conditions, including responses to biotic and abiotic stresses as well as DNA damage, G-rich sequences present in bacterial systems might have the propensity to form these secondary structures under certain physiological conditions. During osmotic shock and general stress responses, the intracellular potassium (K^+) concentration increases, which is a condition known to facilitate G-quadruplex formation (82, 83). Differential gene regulation mediated by these secondary structures is thereby postulated to be a potential mechanism to cope with adverse conditions. In this context, putative G-quadruplexes have been identified in the 5′ UTR of several stress-related genes, suggesting that these highly diverse structures in microbes can be employed to regulate gene expression under stress response conditions involving the expression of *oxyR*, *relA*, and/or *rseA*. OxyR is a transcriptional regulator in the oxidative and nitrosative stress responses (84, 85). The enzyme encoded by *relA* is involved in the responses to adverse conditions (86). RseA is involved in heat shock, osmotic shock, and other stress responses (87). These findings strongly support the role of G4 motifs as regulatory elements involved in responses to adverse condition or environmental changes. For pathogenic microbes, adaptation to different conditions during host-microbe interactions is regulated by several complex and overlapping pathways, which makes it difficult to clearly deduce the role of these transient structures formed *in vivo*.

ROLE OF G4 DNA IN PATHOGENESIS

In many bacterial pathogens, such as *S. pneumoniae* and *M. tuberculosis*, G4 DNA appears to be an important element in the mechanism of transcriptional regulation (64, 73, 74). As described above, G-quadruplexes are found to be highly conserved in the promoters of multiple essential genes in these organisms. In this section, we discuss additional reasons why G-quadruplexes might contribute to the virulence of bacterial pathogens.

(i) Antigenic variability in pathogenic microbes. In some pathogenic microbes, G4 DNA-mediated programmed recombination facilitates antigenic variation of certain surface-expressed immunogenic proteins. This process of antigenic variation helps the pathogens evade the host immune system and has been implicated as playing an important role during host-pathogen interactions. In *Neisseria gonorrhoeae*, a causative agent of gonorrheal infection, a 16-nucleotide sequence containing multiple runs of guanine (5′-GGGTGGGTTGGGTGGG-3′) is located upstream of the transcription start site for the pilin protein gene (*pilE*) (62). In addition, a small noncoding RNA (sRNA) promoter was found adjacent to G4 DNA and upstream of the *pilE* promoter in an

opposite orientation of transcription. *N. gonorrhoeae* contains multiple silent pilin donor cassettes (*pilS*), which are used for *pilE* antigenic variation via a homologous recombination process. The antigenic variation ability of *N. gonorrhoeae* was abolished by mutations disrupting the G4 DNA folding, indicating that secondary DNA structure formation is a key step in the homologous recombination between *pilE* and *pilS*. The antigenic variation is thought to initiate upon the formation of G4 DNA upstream of the transcription start site for *pilE*, causing nicks in the DNA (88–90). The G4-mediated DNA break is then further processed by the recombination process, which uses one of the *pilS* donor cassettes as a template, leading to a change in *pilE* sequence (91). The strand orientation and direction of the G4 motif present at the promoter of the *pilE* gene have a significant effect on the recombination process required for antigenic variation. When the G4 motif is located on the nontranscribed strand, the DNA-RNA hybrid would be generated on the C-rich transcribed strand, which in turn would facilitate the adoption of the secondary structure (i.e., G4 DNA) on the nontranscribed strand (91). However, when G4 motifs were placed in different configurations i.e., on the transcribed strand (inverted), in the reverse direction on the nontranscribed strand (reversed), or in the reverse direction on the transcribed strand (reversed and inverted), pilin antigenic variation was abolished, which suggests that both the correct orientation and the correct direction are required for the G4 structure to form and to function. Similarly, for the sRNA with the G4 motif acting in *cis*, the orientation and direction of the sRNA at the endogenous locus were critical for its function. These results suggest that the G4-mediated break is a key step in the process of pilin antigenic variation. Replacement of G4 DNA by an I-SceI cut at the *pilE* G4 sequence did not yield any antigenic variants, which emphasizes that *pilE* G4 requires either a different type of break, a nick, or more complex interactions with other factors to stimulate this programmed recombination system to promote pilin antigenic variation (92). In addition, transcription of *cis*-acting RNA (G4-sRNA) encoded within G4 DNA of *pilE* was also found to be essential for pilin antigenic variation (93). Significant changes in the size and nucleotide composition of the loop with the core guanine unchanged decreased or abrogated pilin antigenic variation due to the decrease in the thermal stability of *pilE* G4, which demonstrates the importance of folding kinetics and stability of G4 structures for pilin antigenic variation (94).

Similar to *N. gonorrhoeae*, the human pathogen *Borrelia burgdorferi*, the causative agent of Lyme disease, also undergoes the modification of a surface-exposed protein of unknown function, VlsE, which allows the pathogen to escape detection by the host adaptive immune system (95–97). This process is also putatively accomplished through G4-mediated homologous recombination between VlsE and one of the many silent cassettes with homologous sequences. Also, in *Treponema pallidum*, a G-rich sequence motif (d[(G4CT)3G4]) is potentially involved in the antigenic variation of the surface-exposed antigen TprK protein, which plays an important role in both treponemal immune evasion and persistence. Nonreciprocal gene conversion between the *tprK* expression site and donor sites causes *tprK* heterogeneity. Comparative genomic analyses of the donor sites, as well as the *tprK* expression sites, among different *T. pallidum* isolates were performed to understand the diversification of *tprK* during infection (98). The *T. pallidum* d[(G4CT)3G4] sequence motif was found to be highly overrepresented (in total, 58 times) in a diversity of bacterial genomes, including *Burkholderia*, *Frankia*, *Salmonella*, *Shigella*, and *Geobacter* strains. The d[(G4CT)3G4] pattern, which has a stringent selectivity for K^+ as the G4-stabilizing ion *in vitro*, was equally distributed at the ORFs and the UTRs (99).

(ii) RNA G-quadruplexes associated with bacterial pathogenicity and virulence. G-rich RNA sequences can adopt diverse RNA G-quadruplex (rG4) structures and are involved in various biological functions and cellular processes (100). In eukaryotes, rG4s are associated with telomere elongation, recombination, and transcription, as well as RNA posttranscriptional mechanisms (including pre-mRNA processing and mRNA turnover, targeting, and translation), and have a role in cancer and

neurodegeneration (101, 102). However, the presence and role of rG4s in prokaryotes, especially in bacterial species that are human pathogens, are not much appreciated.

A 2016 study conducted by Guo and Bartel showed the occurrence of fewer rG4s in bacterial systems, compared to eukaryotes (103). Interestingly, only one rG4 site was found in *Pseudomonas putida*, an opportunistic human pathogen with a high GC content. In contrast, thousands of RNA regions in eukaryotic genomes have repetitive G-rich sequences that can fold into rG4s *in vitro*. *In vivo*, these regions of RNA were mostly unfolded, as indicated by their accessibility to dimethyl sulfate modification (103). Transcriptome-wide rG4 sequencing analyses and subsequent biophysical, functional, and phenotypic characterizations in a wide range of bacterial species revealed that rG4s are abundant RNA secondary structures (104). Interestingly, these rG4 sites were enriched among genes involved in virulence, gene regulation, cell envelope synthesis, and metabolism in pathogenic bacteria. Such an indication of rG4-mediated regulation of bacterial pathogenicity and metabolic pathways in a wide range of bacterial species suggests that rG4s hold the potential to serve as therapeutic targets.

In addition to rG4 formation, DNA-RNA hybrid G4 structure (HQ) formation during transcription was reported for an *E. coli* plasmid. HQs are a new type of G-quadruplexes that form during *in vitro* transcription and are more prevalent and abundant than DNA G-quadruplex (DQ)-forming motifs in genes (105). A unique distribution of HQs in animal genomes and their correlation with transcription suggest a role of HQs in the regulation of transcription. It was recently suggested that the formation of HQs dominates over that of DQs in cells and that conversion/competition between HQs and DQs may regulate transcription (106).

RESOLUTION OF G-QUADRUPLEXES IN BACTERIA

(i) Role of topoisomerase. The molecular mechanisms underlying the regulation of G4 DNA formation and resolution have not been sufficiently studied in bacterial pathogens and need to be further explored. In yeast genomes, the role of topoisomerase I in the inhibition of genomic instability associated with highly transcribed G4 motifs has been extensively studied (107). Similarly, the role of human topoisomerase I in the resolution of G4 DNA structures has been reported (108). In *D. radiodurans*, K^+-dependent type IB topoisomerase, DraTopoIB, mediates the resolution of G4 DNA (109). In the standard *Taq* DNA polymerase stop assay, DNA synthesis past the G4 motif in the template is hampered in the presence of G4-binding drugs (110). However, incubation of the G4 DNA-containing template with recombinant DraTopoIB expressed and purified from *E. coli* led to the full-length product, which suggests that DraTopoIB is capable of destabilizing these secondary structures. Treatment with the topoisomerase I inhibitor camptothecin restored the polymerization block at G4 DNA, suggesting that the catalytic activity of DraTopoIB is required to resolve the parallel G4 DNA. Future studies are needed to test whether such a role of DraTopoIB as a G4 resolvase has significant relevance in the regulation of gene expression under G4 DNA control.

(ii) Role of DNA helicases. In eukaryotes, several different classes of DNA helicases have been proposed as the major G4 DNA resolvases. These classes include the RecQ family (human BLM and WRN and yeast Sgs1), XPD family helicases with 5′ to 3′ directionality (human FANCJ and *Caenorhabditis elegans* DOG-1), and the Pif1 family (32, 111–115). In *S. cerevisiae*, Pif1 is essential for genome integrity and helps in resolving G4 DNA. In *Schizosaccharomyces pombe*, the absence of Pfh1, which is an ortholog of Pif1, leads to replication fork slowing and chromosomal fragility around G4 sites, although Pfh1 is not involved in replication fork progression in general (114, 116–118). RecQ family helicases are a large family of proteins found in both microbes and eukaryotes and include enzymes such as BLM, WRN, and RECQ4 in humans and Sgs1 in yeast. In *E. coli*, bacterial RecQ helicase resolves G4 structures in an ATP-dependent manner in the presence of divalent cations as cofactors (119). NMM, a G4 ligand, is a highly specific inhibitor of RecQ helicase activity on G4 substrates but has no inhibitory effect on RecQ helicase activity on non-G4 duplex DNA (119). In addition to RecQ helicases, UvrD, which is a multifunctional helicase/translocase, plays a role in resolving G4

structures in *E. coli* (120). Recently, in *E. coli*, the G4-unwinding activity of Rep helicase and Rep-X (an enhanced version of Rep) has been observed *in vitro* (121). The unwinding capacity of Rep helicase is significantly greater than that of the closely related UvrD helicase. In addition to Rep helicase and Rep-X, RecA recombinase can disrupt G4 structures and remove G4-stabilizing ligands *in vitro*. The *M. tuberculosis* genome encodes two orthologs of *E. coli* UvrD helicase, UvrD1 and UvrD2. Both helicases are proficient in resolving tetramolecular G4 structures in an ATP-dependent manner and in maintaining genome integrity (122). Another helicase, DinG, which is a member of the iron-sulfur family of helicases, also unwinds G4 tetraplexes in *E. coli* and *M. tuberculosis* (123). Discovery of small inhibitory molecules specific for these G4-interacting helicases may be useful for the development of novel antibacterial regimens.

G4 APTAMERS TO TARGET BACTERIAL PATHOGENS

Owing to its versatile nature, G4 DNA has been identified in a number of random DNA library screens as synthetic functional nucleic acid molecules, also known as aptamers (124, 125). Aptamers selectively recognize and bind various targets, ranging from small molecules to proteins to whole cells. Due to the extraordinary ability of G4 aptamers to withstand harsh environments such as serum nucleases, they are being used in diagnostics and therapeutics (126–128). A large number of G4 aptamers have been identified for small molecules, proteins, and enzymes (126). One classic and well-studied G4 aptamer is the antithrombin aptamer that was generated through the systematic evolution of ligands by exponential enrichment (SELEX) process. This thrombin-binding aptamer is an antiparallel G4 aptamer that binds to exosite I of thrombin (129). In recent years, G4 aptamers have also been developed to tackle bacterial pathogenesis. Recently, Shum et al. identified a G4 DNA aptamer against *M. tuberculosis* polyphosphate kinase 2 (PPK2), which regulates the intracellular metabolism of inorganic polyphosphate (polyP). PolyP is vital for *M. tuberculosis* maintaining its virulence (130). Therefore, blocking PPK2 with a G4 aptamer is detrimental to *M. tuberculosis*. This PPK2-specific aptamer was highly potent and exhibited a 50% inhibitory concentration (IC_{50}) of 40 nM, with a dissociation constant (K_d) of 870 nM. That study clearly demonstrated the potential of G4 aptamers in controlling bacterial infections by targeting vital enzymes. In another recent work, Kalra et al. developed two high-affinity G4 aptamers (HupB-4T and HupB-13T) against *M. tuberculosis* HupB protein (131). HupB is an essential histone-like protein of *M. tuberculosis* that plays an important role in bacterial entry and survival in host cells. Both of these aptamers have submicromolar K_d values and bind to two distinct sites on HupB, inhibit the DNA-binding function of HupB, and block *M. tuberculosis* entry into host monocytic cells by targeting the surface-associated HupB. Additionally, the theranostic potential of a G4-forming DNA aptamer targeting malate synthase (MS) of *M. tuberculosis*, an enzyme in the glyoxylate pathway, has been demonstrated (132). MS also plays an important role as an adhesin and helps bacteria invade the host cell. In that study, through SELEX and post-SELEX optimization, an 11-mer parallel G4-forming aptamer with picomolar affinity for MS was identified. The G4 aptamer exhibited robust inhibition of the catalytic function of MS, with an IC_{50} of 251 nM and a K_i of 230 nM. Furthermore, this aptamer blocked mycobacterial entry into host cells by targeting the adhesin function of MS (132). Additionally, this G4 aptamer was successfully used in the diagnosis of tuberculous meningitis with cerebrospinal fluid specimens, with sensitivity and specificity values of $>$97%, demonstrating the potential of G4 aptamers for the diagnosis of bacterial infections (132). Another possible advantage of G4 aptamers could be that, if a panel of G4 aptamers were carefully developed, it could be used for multidrug-resistant pathogens if it were used in a cocktail in which each aptamer binds to a unique target. Overall, the recent investigations described here fully support the effectiveness of G4 aptamers as a new class of therapeutic agents that can complement the existing drug regimens to treat bacterial infections.

Finally, we have a cautionary note regarding the potential therapeutic applications of aptamers. Broadly, aptamers are nonimmunogenic; in recent years, however, it has been seen that CpG oligonucleotides can induce innate immune responses through their binding to some Toll-like receptors (TLRs). Therefore, as a caution, exposed CpG sequences in aptamers should be avoided. One intelligent way of doing this would be truncation of aptamers or methylation of cytosines in CpG regions, which could potentially overcome the problem of aptamer toxicity (133, 134).

FUTURE PERSPECTIVE

To date, understanding of the association of G-quadruplex motifs with intrinsic cellular functions in bacterial systems lags significantly behind that in eukaryotic microbes or in protozoan systems. However, studies are being conducted to understand the roles of G4 structures in the pathogenesis of several important microbial pathogens. Computational analyses of diverse prokaryotes suggest that highly disparate microorganisms can use G4 DNA in transcriptional regulation. Furthermore, the occurrence of highly conserved DNA secondary structures within the bacterial genome suggests their regulatory role in bacterial virulence and pathogenesis. Overall, these findings suggest the potential of G-quadruplexes either as drug targets or as therapeutic aptamers in pharmaceutical applications. In addition, helicases with G4 substrates may be promising drug targets for antibacterial agents to combat specific pathogenic bacteria. In the future, for G-quadruplexes beyond the current status of the computational predictions (65), adaptation of the high-throughput, genome-wide methods used for G4 detection in higher eukaryotic organisms to bacterial systems would provide a significant advance in the understanding of the role of G4 in gene regulation and host-microbe interactions.

ACKNOWLEDGMENTS

This work was supported in part by a grant to P.Y. (grant ECR/2015/000431) from the Science and Engineering Research Board, Department of Science and Technology, Government of India, and a grant to V.Y. (Ramalingaswami Fellowship BT/RLF/RE-ENTRY/29/2014) from the Department of Biotechnology, Government of India. This work was supported by grants from the National Institutes of Health (grant R01 GM116007) and from the Welch Foundation (grant AU1875) to N.K.

P.Y., N.K., and V.Y. contributed to the conception of the review article. P.Y., M.K., and V.Y. drafted the work. P.Y., N.K., M.K., S.V., T.K.S., V.Y., and A.K. wrote the review article. P.Y., N.K., and V.Y. revised the article critically. P.Y., M.K., and T.K.S. helped in literature searches. P.Y., N.K., T.K.S., and V.Y. gave final approval of the version to be published.

We declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Bang I. 1910. Untersuchungen über die Guanylsäure. Biochem Z 26:293–311.
2. Sen D, Gilbert W. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334:364–366. https://doi.org/10.1038/334364a0.
3. Arnott S, Chandrasekaran R, Marttila CM. 1974. Structures for polyinosinic acid and polyguanylic acid. Biochem J 141:537–543. https://doi.org/10.1042/bj1410537.
4. Zimmerman SB, Cohen GH, Davies DR. 1975. X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid. J Mol Biol 92:181–192. https://doi.org/10.1016/0022-2836(75)90222-3.
5. Gellert M, Lipsett MN, Davies DR. 1962. Helix formation by guanylic acid. Proc Natl Acad Sci U S A 48:2013–2018. https://doi.org/10.1073/pnas.48.12.2013.
6. Sen D, Gilbert W. 1990. A sodium-potassium switch in the formation of four-stranded G4-DNA. Nature 344:410–414. https://doi.org/10.1038/344410a0.
7. Sundquist WI, Klug A. 1989. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature 342:825–829. https://doi.org/10.1038/342825a0.
8. Bochman ML, Paeschke K, Zakian VA. 2012. DNA secondary structures: stability and function of G-quadruplex structures. Nat Rev Genet 13:770–780. https://doi.org/10.1038/nrg3296.
9. Sun Z-Y, Wang X-N, Cheng S-Q, Su X-X, Ou T-M. 2019. Developing novel G-quadruplex ligands: from interaction with nucleic acids to interfering with nucleic acid–protein interaction. Molecules 24:396. https://doi.org/10.3390/molecules24030396.
10. Huppert JL, Balasubramanian S. 2005. Prevalence of quadruplexes in the human genome. Nucleic Acids Res 33:2908–2916. https://doi.org/10.1093/nar/gki609.
11. Largy E, Mergny J-L, Gabelica V. 2016. Role of alkali metal ions in G-quadruplex nucleic acid structure and stability. Met Ions Life Sci 16:203–258. https://doi.org/10.1007/978-3-319-21756-7_7.
12. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. 2006. Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res 34:5402–5415. https://doi.org/10.1093/nar/gkl655.
13. Hazel P, Parkinson GN, Neidle S. 2006. Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures. Nucleic Acids Res 34:2117–2127. https://doi.org/10.1093/nar/gkl182.

14. Zhou J, Bourdoncle A, Rosu F, Gabelica V, Mergny J-L. 2012. Tri-G-quadruplex: controlled assembly of a G-quadruplex structure from three G-rich strands. Angew Chem Int Ed Engl 51:11002–11005. https://doi.org/10.1002/anie.201205390.
15. Huppert JL, Balasubramanian S. 2007. G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res 35:406–413. https://doi.org/10.1093/nar/gkl1057.
16. Eddy J, Maizels N. 2006. Gene function correlates with potential for G4 DNA formation in the human genome. Nucleic Acids Res 34:3887–3896. https://doi.org/10.1093/nar/gkl529.
17. Dhapola P, Chowdhury S. 2016. QuadBase2: web server for multiplexed guanine quadruplex mining and visualization. Nucleic Acids Res 44:W277–W283. https://doi.org/10.1093/nar/gkw425.
18. Wong HM, Stegle O, Rodgers S, Huppert JL. 2010. A toolbox for predicting G-quadruplex formation and stability. J Nucleic Acids 2010:564946. https://doi.org/10.4061/2010/564946.
19. Capra JA, Paeschke K, Singh M, Zakian VA. 2010. G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. PLoS Comput Biol 6:e1000861. https://doi.org/10.1371/journal.pcbi.1000861.
20. Rawal P, Kummarasetti VBR, Ravindran J, Kumar N, Halder K, Sharma R, Mukerji M, Das SK, Chowdhury S. 2006. Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. Genome Res 16:644–655. https://doi.org/10.1101/gr.4508806.
21. Todd AK, Johnston M, Neidle S. 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res 33:2901–2907. https://doi.org/10.1093/nar/gki553.
22. Mullen MA, Olson KJ, Dallaire P, Major F, Assmann SM, Bevilacqua PC. 2010. RNA G-quadruplexes in the model plant species *Arabidopsis thaliana*: prevalence and possible functional roles. Nucleic Acids Res 38:8149–8163. https://doi.org/10.1093/nar/gkq804.
23. Andorf CM, Kopylov M, Dobbs D, Koch KE, Stroupe ME, Lawrence CJ, Bass HW. 2014. G-quadruplex (G4) motifs in the maize (*Zea mays* L.) genome are enriched at specific locations in thousands of genes coupled to energy status, hypoxia, low sugar, and nutrient deprivation. J Genet Genomics 41:627–647. https://doi.org/10.1016/j.jgg.2014.10.004.
24. Yadav V, Hemansi, Kim N, Tuteja N, Yadav P. 2017. G quadruplex in plants: a ubiquitous regulatory element and its biological relevance. Front Plant Sci 8:1163. https://doi.org/10.3389/fpls.2017.01163.
25. Ravichandran S, Kim Y-E, Bansal V, Ghosh A, Hur J, Subramani VK, Pradhan S, Lee MK, Kim KK, Ahn J-H. 2018. Genome-wide analysis of regulatory G-quadruplexes affecting gene expression in human cytomegalovirus. PLoS Pathog 14:e1007334. https://doi.org/10.1371/journal.ppat.1007334.
26. Parkinson GN, Lee MPH, Neidle S. 2002. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 417:876–880. https://doi.org/10.1038/nature755.
27. Blackburn EH. 1994. Telomeres: no end in sight. Cell 77:621–623. https://doi.org/10.1016/0092-8674(94)90046-9.
28. Smith JS, Chen Q, Yatsunyk LA, Nicoludis JM, Garcia MS, Kranaster R, Balasubramanian S, Monchaud D, Teulade-Fichou M-P, Abramowitz L, Schultz DC, Johnson FB. 2011. Rudimentary G-quadruplex–based telomere capping in *Saccharomyces cerevisiae*. Nat Struct Mol Biol 18:478–485. https://doi.org/10.1038/nsmb.2033.
29. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. 2002. 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 99:11593–11598. https://doi.org/10.1073/pnas.182256799.
30. Evans T, Schon E, Gora-Maslak G, Patterson J, Efstratiadis A. 1984. S1-hypersensitive sites in eukaryotic promoter regions. Nucleic Acids Res 12:8043–8058. https://doi.org/10.1093/nar/12.21.8043.
31. Kilpatrick MW, Torri A, Kang DS, Engler JA, Wells RD. 1986. Unusual DNA structures in the adenovirus genome. J Biol Chem 261:11350–11354. https://doi.org/10.1016/S0021-9258(18)67390-9.
32. Gray LT, Vallur AC, Eddy J, Maizels N. 2014. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. Nat Chem Biol 10:313–318. https://doi.org/10.1038/nchembio.1475.
33. Fernando H, Sewitz S, Darot J, Tavaré S, Huppert JL, Balasubramanian S. 2009. Genome-wide analysis of a G-quadruplex-specific single-chain antibody that regulates gene expression. Nucleic Acids Res 37:6716–6722. https://doi.org/10.1093/nar/gkp740.
34. Huang H, Zhang J, Harvey SE, Hu X, Cheng C. 2017. RNA G-quadruplex secondary structure promotes alternative splicing via the RNA-binding protein hnRNPF. Genes Dev 31:2296–2309. https://doi.org/10.1101/gad.305862.117.
35. Kumari S, Bugaut A, Huppert JL, Balasubramanian S. 2007. An RNA G-quadruplex in the 5′ UTR of the *NRAS* proto-oncogene modulates translation. Nat Chem Biol 3:218–221. https://doi.org/10.1038/nchembio864.
36. Bugaut A, Balasubramanian S. 2012. 5′-UTR RNA G-quadruplexes: translation regulation and targeting. Nucleic Acids Res 40:4727–4741. https://doi.org/10.1093/nar/gks068.
37. Beaudoin J-D, Perreault J-P. 2010. 5′-UTR G-quadruplex structures acting as translational repressors. Nucleic Acids Res 38:7022–7036. https://doi.org/10.1093/nar/gkq557.
38. Morris MJ, Negishi Y, Pazsint C, Schonhoft JD, Basu S. 2010. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. J Am Chem Soc 132:17831–17839. https://doi.org/10.1021/ja106287x.
39. Chen E, Sharma MR, Shi X, Agrawal RK, Joseph S. 2014. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. Mol Cell 54:407–417. https://doi.org/10.1016/j.molcel.2014.03.023.
40. Rouleau S, Glouzon J-PS, Brumwell A, Bisaillon M, Perreault J-P. 2017. 3′ UTR G-quadruplexes regulate miRNA binding. RNA 23:1172–1179. https://doi.org/10.1261/rna.060962.117.
41. Ishiguro A, Kimura N, Watanabe Y, Watanabe S, Ishihama A. 2016. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. Genes Cells 21:466–481. https://doi.org/10.1111/gtc.12352.
42. Subramanian M, Rage F, Tabet R, Flatter E, Mandel J-L, Moine H. 2011. G–quadruplex RNA structure as a signal for neurite mRNA targeting. EMBO Rep 12:697–704. https://doi.org/10.1038/embor.2011.76.
43. Kim N. 2019. The interplay between G-quadruplex and transcription. Curr Med Chem 26:2898–2917. https://doi.org/10.2174/0929867325666171229132619.
44. Holder IT, Hartig JS. 2014. A matter of location: influence of G-quadruplexes on *Escherichia coli* gene expression. Chem Biol 21:1511–1521. https://doi.org/10.1016/j.chembiol.2014.09.014.
45. Dexheimer TS, Sun D, Hurley LH. 2006. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the *bcl*-2 P1 promoter. J Am Chem Soc 128:5404–5415. https://doi.org/10.1021/ja0563861.
46. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. 2006. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. Biochemistry 45:7854–7860. https://doi.org/10.1021/bi0601510.
47. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S. 2005. Putative DNA quadruplex formation within the human c-*kit* oncogene. J Am Chem Soc 127:10584–10589. https://doi.org/10.1021/ja050823u.
48. Palumbo SL, Memmott RM, Uribe DJ, Krotova-Khan Y, Hurley LH, Ebbinghaus SW. 2008. A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. Nucleic Acids Res 36:1755–1769. https://doi.org/10.1093/nar/gkm1069.
49. Sun D, Guo K, Rusche JJ, Hurley LH. 2005. 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 33:6070–6080. https://doi.org/10.1093/nar/gki917.
50. Cogoi S, Paramasivam M, Filichev V, Géci I, Pedersen EB, Xodo LE. 2009. Identification of a new G-quadruplex motif in the *KRAS* promoter and design of pyrene-modified G4-decoys with antiproliferative activity in pancreatic cancer cells. J Med Chem 52:564–568. https://doi.org/10.1021/jm800874t.
51. Cogoi S, Xodo LE. 2006. G-quadruplex formation within the promoter of the *KRAS* proto-oncogene and its effect on transcription. Nucleic Acids Res 34:2536–2549. https://doi.org/10.1093/nar/gkl286.
52. De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW. 2005. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α promoter. Biochemistry 44:16341–16350. https://doi.org/10.1021/bi051618u.
53. Sinden RR. 1994. DNA structure and function. Academic Press, San Diego, CA.
54. Pérez-Martín J, de Lorenzo V. 1997. Clues and consequences of DNA bending in transcription. Annu Rev Microbiol 51:593–628. https://doi.org/10.1146/annurev.micro.51.1.593.

55. Pedersen AG, Jensen LJ, Brunak S, Staerfeldt HH, Ussery DW. 2000. A DNA structural atlas for *Escherichia coli*. J Mol Biol 299:907–930. https://doi.org/10.1006/jmbi.2000.3787.
56. Bacolla A, Wells RD. 2004. Non-B DNA conformations, genomic rearrangements, and human disease. J Biol Chem 279:47411–47414. https://doi.org/10.1074/jbc.R400028200.
57. Henderson A, Wu Y, Huang YC, Chavez EA, Platt J, Johnson FB, Brosh RM, Sen D, Lansdorp PM. 2014. Detection of G-quadruplex DNA in mammalian cells. Nucleic Acids Res 42:860–869. https://doi.org/10.1093/nar/gkt957.
58. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem 5:182–186. https://doi.org/10.1038/nchem.1548.
59. Biffi G, Tannahill D, Miller J, Howat WJ, Balasubramanian S. 2014. Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. PLoS One 9:e102711. https://doi.org/10.1371/journal.pone.0102711.
60. Lam EYN, Beraldi D, Tannahill D, Balasubramanian S. 2013. G-quadruplex structures are stable and detectable in human genomic DNA. Nat Commun 4:1796. https://doi.org/10.1038/ncomms2792.
61. Hatfield GW, Benham CJ. 2002. DNA topology-mediated control of global gene expression in *Escherichia coli*. Annu Rev Genet 36:175–203. https://doi.org/10.1146/annurev.genet.36.032902.111815.
62. Harris LM, Merrick CJ. 2015. G-quadruplexes in pathogens: a common route to virulence control? PLoS Pathog 11:e1004562. https://doi.org/10.1371/journal.ppat.1004562.
63. Saranathan N, Vivekanandan P. 2019. G-quadruplexes: more than just a kink in microbial genomes. Trends Microbiol 27:148–163. https://doi.org/10.1016/j.tim.2018.08.011.
64. Mishra SK, Shankar U, Jain N, Sikri K, Tyagi JS, Sharma TK, Mergny J-L, Kumar A. 2019. Characterization of G-quadruplex motifs in *espB*, *espK*, and *cyp51* genes of *Mycobacterium tuberculosis* as potential drug targets. Mol Ther Nucleic Acids 16:698–706. https://doi.org/10.1016/j.omtn.2019.04.022.
65. Yadav VK, Abraham JK, Mani P, Kulshrestha R, Chowdhury S. 2008. QuadBase: genome-wide database of G4 DNA occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. Nucleic Acids Res 36:D381–D385. https://doi.org/10.1093/nar/gkm781.
66. Beaume N, Pathak R, Yadav VK, Kota S, Misra HS, Gautam HK, Chowdhury S. 2013. Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of *D. radiodurans* involves G4 DNA-mediated regulation. Nucleic Acids Res 41:76–89. https://doi.org/10.1093/nar/gks1071.
67. Ding Y, Fleming AM, Burrows CJ. 2018. Case studies on potential G-quadruplex-forming sequences from the bacterial orders *Deinococcales* and *Thermales* derived from a survey of published genomes. Sci Rep 8:15679. https://doi.org/10.1038/s41598-018-33944-4.
68. Bartas M, Čutová M, Brázda V, Kaura P, Šťastný J, Kolomazník J, Coufal J, Goswami P, Červeň J, Pečinka P. 2019. The presence and localization of G-quadruplex forming sequences in the domain of bacteria. Molecules 24:1711. https://doi.org/10.3390/molecules24091711.
69. Brazda V, Fojta M, Bowater RP. 2020. Structures and stability of simple DNA repeats from bacteria. Biochem J 477:325–339. https://doi.org/10.1042/BCJ20190703.
70. Kota S, Dhamodharan V, Pradeepkumar PI, Misra HS. 2015. G-quadruplex forming structural motifs in the genome of *Deinococcus radiodurans* and their regulatory roles in promoter functions. Appl Microbiol Biotechnol 99:9761–9769. https://doi.org/10.1007/s00253-015-6808-6.
71. Mishra S, Chaudhary R, Singh S, Kota S, Misra HS. 2019. Guanine quadruplex DNA regulates gamma radiation response of genome functions in the radioresistant bacterium *Deinococcus radiodurans*. J Bacteriol 201:e00154-19. https://doi.org/10.1128/JB.00154-19.
72. Waller ZAE, Pinchbeck BJ, Buguth BS, Meadows TG, Richardson DJ, Gates AJ. 2016. Control of bacterial nitrate assimilation by stabilization of G-quadruplex DNA. Chem Commun (Camb) 52:13511–13514. https://doi.org/10.1039/c6cc06057a.
73. Perrone R, Lavezzo E, Riello E, Manganelli R, Palù G, Toppo S, Provvedi R, Richter SN. 2017. Mapping and characterization of G-quadruplexes in *Mycobacterium tuberculosis* gene promoter regions. Sci Rep 7:5743. https://doi.org/10.1038/s41598-017-05867-z.
74. Mishra SK, Jain N, Shankar U, Tawani A, Sharma TK, Kumar A. 2019. Characterization of highly conserved G-quadruplex motifs as potential drug targets in *Streptococcus pneumoniae*. Sci Rep 9:1791. https://doi.org/10.1038/s41598-018-38400-x.
75. Agarwal T, Roy S, Kumar S, Chakraborty TK, Maiti S. 2014. In the sense of transcription regulation by G-quadruplexes: asymmetric effects in sense and antisense strands. Biochemistry 53:3711–3718. https://doi.org/10.1021/bi401451q.
76. Feklistov A. 2013. RNA polymerase: in search of promoters. Ann N Y Acad Sci 1293:25–32. https://doi.org/10.1111/nyas.12197.
77. Mekler V, Minakhin L, Severinov K. 2011. A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex. J Biol Chem 286:22600–22608. https://doi.org/10.1074/jbc.M111.247080.
78. Roberts JW, Roberts CW. 1996. Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. Cell 86:495–501. https://doi.org/10.1016/S0092-8674(00)80122-1.
79. Endoh T, Kawasaki Y, Sugimoto N. 2013. Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability. Angew Chem Int Ed Engl 52:5522–5526. https://doi.org/10.1002/anie.201300058.
80. Wieland M, Hartig JS. 2007. RNA quadruplex-based modulation of gene expression. Chem Biol 14:757–763. https://doi.org/10.1016/j.chembiol.2007.06.005.
81. Endoh T, Sugimoto N. 2013. Unusual −1 ribosomal frameshift caused by stable RNA G-quadruplex in open reading frame. Anal Chem 85:11435–11439. https://doi.org/10.1021/ac402497x.
82. Epstein W, Schultz SG. 1965. Cation transport in *Escherichia coli*. J Gen Physiol 49:221–234. https://doi.org/10.1085/jgp.49.2.221.
83. Huo Y-X, Rosenthal AZ, Gralla JD. 2008. General stress response signalling: unwrapping transcription complexes by DNA relaxation via the sigma38 C-terminal domain. Mol Microbiol 70:369–378. https://doi.org/10.1111/j.1365-2958.2008.06412.x.
84. Anjem A, Varghese S, Imlay JA. 2009. Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. Mol Microbiol 72:844–858. https://doi.org/10.1111/j.1365-2958.2009.06699.x.
85. Seth D, Hausladen A, Wang Y-J, Stamler JS. 2012. Endogenous protein *S*-nitrosylation in *E. coli*: regulation by OxyR. Science 336:470–473. https://doi.org/10.1126/science.1215643.
86. Magnusson LU, Farewell A, Nyström T. 2005. ppGpp: a global regulator in *Escherichia coli*. Trends Microbiol 13:236–242. https://doi.org/10.1016/j.tim.2005.03.008.
87. Ades SE, Connolly LE, Alba BM, Gross CA. 1999. The *Escherichia coli* σ^E-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-σ factor. Genes Dev 13:2449–2461. https://doi.org/10.1101/gad.13.18.2449.
88. Cahoon LA, Seifert HS. 2009. An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. Science 325:764–767. https://doi.org/10.1126/science.1175653.
89. Kuryavyi V, Cahoon LA, Seifert HS, Patel DJ. 2012. RecA-binding *pilE* G4 sequence essential for pilin antigenic variation forms monomeric and 5′ end-stacked dimeric parallel G-quadruplexes. Structure 20:2090–2102. https://doi.org/10.1016/j.str.2012.09.013.
90. Vink C, Rudenko G, Seifert HS. 2012. Microbial antigenic variation mediated by homologous DNA recombination. FEMS Microbiol Rev 36:917–948. https://doi.org/10.1111/j.1574-6976.2011.00321.x.
91. Cahoon LA, Seifert HS. 2013. Transcription of a *cis*-acting, noncoding, small RNA is required for pilin antigenic variation in *Neisseria gonorrhoeae*. PLoS Pathog 9:e1003074. https://doi.org/10.1371/journal.ppat.1003074.
92. Prister LL, Xu J, Seifert HS. 2019. A double-strand break does not promote *Neisseria gonorrhoeae* pilin antigenic variation. J Bacteriol 201:e00256-19. https://doi.org/10.1128/JB.00256-19.
93. Prister LL, Ozer EA, Cahoon LA, Seifert HS. 2019. Transcriptional initiation of a small RNA, not R-loop stability, dictates the frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. Mol Microbiol 112:1219–1234. https://doi.org/10.1111/mmi.14356.
94. Prister LL, Yin S, Cahoon LA, Seifert HS. 2020. Altering the *Neisseria gonorrhoeae pilE* guanine quadruplex loop bases affects pilin antigenic variation. Biochemistry 59:1104–1112. https://doi.org/10.1021/acs.biochem.9b01038.
95. Zhang J-R, Norris SJ. 1998. Genetic variation of the *Borrelia burgdorferi* gene *vlsE* involves cassette-specific, segmental gene conversion. Infect Immun 66:3698–3704. https://doi.org/10.1128/IAI.66.8.3698-3704.1998.
96. Lin T, Gao L, Edmondson DG, Jacobs MB, Philipp MT, Norris SJ. 2009. Central role of the Holliday junction helicase RuvAB in *vlsE* recombination and infectivity of *Borrelia burgdorferi*. PLoS Pathog 5:e1000679. https://doi.org/10.1371/journal.ppat.1000679.

97. Walia R, Chaconas G. 2013. Suggested role for G4 DNA in recombinational switching at the antigenic variation locus of the Lyme disease spirochete. PLoS One 8:e57792. https://doi.org/10.1371/journal.pone.0057792.
98. Giacani L, Brandt SL, Puray-Chavez M, Reid TB, Godornes C, Molini BJ, Benzler M, Hartig JS, Lukehart SA, Centurion-Lara A. 2012. Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J Bacteriol 194:4208–4225. https://doi.org/10.1128/JB.00863-12.
99. Rehm C, Holder IT, Groß A, Wojciechowski F, Urban M, Sinn M, Drescher M, Hartig JS. 2014. A bacterial DNA quadruplex with exceptional K^+ selectivity and unique structural polymorphism. Chem Sci 5:2809–2818. https://doi.org/10.1039/C4SC00440J.
100. Fay MM, Lyons SM, Ivanov P. 2017. RNA G-quadruplexes in biology: principles and molecular mechanisms. J Mol Biol 429:2127–2147. https://doi.org/10.1016/j.jmb.2017.05.017.
101. Wolfe AL, Singh K, Zhong Y, Drewe P, Rajasekhar VK, Sanghvi VR, Mavrakis KJ, Jiang M, Roderick JE, Van der Meulen J, Schatz JH, Rodrigo CM, Zhao C, Rondou P, de Stanchina E, Teruya-Feldstein J, Kelliher MA, Speleman F, Porco JA, Pelletier J, Rätsch G, Wendel H-G. 2014. RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. Nature 513:65–70. https://doi.org/10.1038/nature13485.
102. Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim M-S, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J. 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 507:195–200. https://doi.org/10.1038/nature13124.
103. Guo JU, Bartel DP. 2016. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353:aaf5371. https://doi.org/10.1126/science.aaf5371.
104. Shao X, Zhang W, Umar MI, Wong HY, Seng Z, Xie Y, Zhang Y, Yang L, Kwok CK, Deng X. 2020. RNA G-quadruplex structures mediate gene regulation in bacteria. mBio 11:e02926-19. https://doi.org/10.1128/mBio.02926-19.
105. Zhang J, Zheng K, Xiao S, Hao Y, Tan Z. 2014. Mechanism and manipulation of DNA:RNA hybrid G-quadruplex formation in transcription of G-rich DNA. J Am Chem Soc 136:1381–1390. https://doi.org/10.1021/ja4085572.
106. Wu R, Zheng K, Zhang J, Hao Y, Tan Z. 2015. Formation of DNA:RNA hybrid G-quadruplex in bacterial cells and its dominance over the intramolecular DNA G-quadruplex in mediating transcription termination. Angew Chem Int Ed Engl 54:2447–2451. https://doi.org/10.1002/anie.201408719.
107. Yadav P, Harcy V, Argueso JL, Dominska M, Jinks-Robertson S, Kim N. 2014. Topoisomerase I plays a critical role in suppressing genome instability at a highly transcribed G-quadruplex-forming sequence. PLoS Genet 10:e1004839. https://doi.org/10.1371/journal.pgen.1004839.
108. Arimondo PB, Riou J-F, Mergny J-L, Tazi J, Sun J-S, Garestier T, Hélène C. 2000. Interaction of human DNA topoisomerase I with G-quartet structures. Nucleic Acids Res 28:4832–4838. https://doi.org/10.1093/nar/28.24.4832.
109. Kota S, Misra HS. 2015. Topoisomerase IB of *Deinococcus radiodurans* resolves guanine quadruplex DNA structures in vitro. J Biosci 40:833–843. https://doi.org/10.1007/s12038-015-9571-z.
110. Jamroskovic J, Obi I, Movahedi A, Chand K, Chorell E, Sabouri N. 2019. Identification of putative G-quadruplex DNA structures in *S. pombe* genome by quantitative PCR stop assay. DNA Repair (Amst) 82:102678. https://doi.org/10.1016/j.dnarep.2019.102678.
111. Maizels N, Gray LT. 2013. The G4 genome. PLoS Genet 9:e1003468. https://doi.org/10.1371/journal.pgen.1003468.
112. Estep KN, Butler TJ, Ding J, Brosh RM. 2019. G4-interacting DNA helicases and polymerases: potential therapeutic targets. Curr Med Chem 26:2881–2897. https://doi.org/10.2174/0929867324666171116123345.
113. Cheung I, Schertzer M, Rose A, Lansdorp PM. 2002. Disruption of *dog-1* in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA. Nat Genet 31:405–409. https://doi.org/10.1038/ng928.
114. Paeschke K, Bochman ML, Garcia PD, Cejka P, Friedman KL, Kowalczykowski SC, Zakian VA. 2013. Pif1 family helicases suppress genome instability at G-quadruplex motifs. Nature 497:458–462. https://doi.org/10.1038/nature12149.
115. Brosh RM, Matson SW. 2020. History of DNA helicases. Genes 11:255. https://doi.org/10.3390/genes11030255.
116. Paeschke K, Capra JA, Zakian VA. 2011. DNA replication through G-quadruplex motifs is promoted by the *S. cerevisiae* Pif1 DNA helicase. Cell 145:678–691. https://doi.org/10.1016/j.cell.2011.04.015.
117. Lopes J, Piazza A, Bermejo R, Kriegsman B, Colosio A, Teulade-Fichou M-P, Foiani M, Nicolas A. 2011. G-quadruplex-induced instability during leading-strand replication. EMBO J 30:4033–4046. https://doi.org/10.1038/emboj.2011.316.
118. Sabouri N, McDonald KR, Webb CJ, Cristea IM, Zakian VA. 2012. DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the *S. pombe* Pfh1 helicase. Genes Dev 26:581–593. https://doi.org/10.1101/gad.184697.111.
119. Wu X, Maizels N. 2001. Substrate-specific inhibition of RecQ helicase. Nucleic Acids Res 29:1765–1771. https://doi.org/10.1093/nar/29.8.1765.
120. Shukla K, Thakur RS, Ganguli D, Rao DN, Nagaraju G. 2017. *Escherichia coli* and *Neisseria gonorrhoeae* UvrD helicase unwinds G4 DNA structures. Biochem J 474:3579–3597. https://doi.org/10.1042/BCJ20170587.
121. Paul T, Voter AF, Cueny RR, Gavrilov M, Ha T, Keck JL, Myong S. 2020. *E. coli* Rep helicase and RecA recombinase unwind G4 DNA and are important for resistance to G4-stabilizing ligands. Nucleic Acids Res 48:6640–6653. https://doi.org/10.1093/nar/gkaa442.
122. Saha T, Shukla K, Thakur RS, Desingu A, Nagaraju G. 2019. *Mycobacterium tuberculosis* UvrD1 and UvrD2 helicases unwind G-quadruplex DNA. FEBS J 286:2062–2086. https://doi.org/10.1111/febs.14798.
123. Thakur RS, Desingu A, Basavaraju S, Subramanya S, Rao DN, Nagaraju G. 2014. *Mycobacterium tuberculosis* DinG is a structure-specific helicase that unwinds G4 DNA. J Biol Chem 289:25112–25136. https://doi.org/10.1074/jbc.M114.563569.
124. Ku T-H, Zhang T, Luo H, Yen TM, Chen P-W, Han Y, Lo Y-H. 2015. Nucleic acid aptamers: an emerging tool for biotechnology and biomedical sensing. Sensors (Basel) 15:16281–16313. https://doi.org/10.3390/s150716281.
125. Platella C, Riccardi C, Montesarchio D, Roviello GN, Musumeci D. 2017. G-quadruplex-based aptamers against protein targets in therapy and diagnostics. Biochim Biophys Acta Gen Subj 1861:1429–1447. https://doi.org/10.1016/j.bbagen.2016.11.027.
126. Roxo C, Kotkowiak W, Pasternak A. 2019. G-quadruplex-forming aptamers: characteristics, applications, and perspectives. Molecules 24:3781. https://doi.org/10.3390/molecules24203781.
127. Kaur H, Bruno JG, Kumar A, Sharma TK. 2018. Aptamers in the therapeutics and diagnostics pipelines. Theranostics 8:4016–4032. https://doi.org/10.7150/thno.25958.
128. Sharma TK, Bruno JG, Dhiman A. 2017. ABCs of DNA aptamer and related assay development. Biotechnol Adv 35:275–301. https://doi.org/10.1016/j.biotechadv.2017.01.003.
129. Bock LC, Griffin LC, Latham JA, Vermaas EH, Toole JJ. 1992. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. Nature 355:564–566. https://doi.org/10.1038/355564a0.
130. Shum KT, Lui ELH, Wong SCK, Yeung P, Sam L, Wang Y, Watt RM, Tanner JA. 2011. Aptamer-mediated inhibition of *Mycobacterium tuberculosis* polyphosphate kinase. Biochemistry 50:3261–3271. https://doi.org/10.1021/bi2001455.
131. Kalra P, Mishra SK, Kaur S, Kumar A, Prasad HK, Sharma TK, Tyagi JS. 2018. G-quadruplex-forming DNA aptamers inhibit the DNA-binding function of HupB and *Mycobacterium tuberculosis* entry into host cells. Mol Ther Nucleic Acids 13:99–109. https://doi.org/10.1016/j.omtn.2018.08.011.
132. Dhiman A, Kumar C, Mishra SK, Sikri K, Datta I, Sharma P, Singh TP, Haldar S, Sharma N, Bansal A, Ahmad Y, Kumar A, Sharma TK, Tyagi JS. 2019. Theranostic application of a novel G-quadruplex-forming DNA aptamer targeting malate synthase of *Mycobacterium tuberculosis*. Mol Ther Nucleic Acids 18:661–672. https://doi.org/10.1016/j.omtn.2019.09.026.
133. Bruno JG. 2018. Potential inherent stimulation of the innate immune system by nucleic acid aptamers and possible corrective approaches. Pharmaceuticals (Basel) 11:62. https://doi.org/10.3390/ph11030062.
134. Mena A, Nichani AK, Popowych Y, Godson DL, Dent D, Townsend HGG, Mutwiri GK, Hecker R, Babiuk LA, Griebel P. 2003. Innate immune responses induced by CpG oligodeoxyribonucleotide stimulation of ovine blood mononuclear cells. Immunology 110:250–257. https://doi.org/10.1046/j.1365-2567.2003.01722.x.