Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Trends in Microbiology

CellPress
REVIEWS

Review

G-Quadruplexes: More Than Just a Kink in Microbial Genomes

Nandhini Saranathan[1] and Perumal Vivekanandan[1,*]

G-quadruplexes (G4s) are noncanonical nucleic acid secondary structures formed by guanine-rich DNA and RNA sequences. In this review we aim to provide an overview of the biological roles of G4s in microbial genomes with emphasis on recent discoveries. G4s are enriched and conserved in the regulatory regions of microbes, including bacteria, fungi, and viruses. Importantly, G4s in hepatitis B virus (HBV) and hepatitis C virus (HCV) genomes modulate genes crucial for virus replication. Recent studies on Epstein–Barr virus (EBV) shed light on the role of G4s within the microbial transcripts as *cis*-acting regulatory signals that modulate translation and facilitate immune evasion. Furthermore, G4s in microbial genomes have been linked to radioresistance, antigenic variation, recombination, and latency. G4s in microbial genomes represent novel therapeutic targets for antimicrobial therapy.

Highlights

G4s display functional diversity among microbes. Their ability to influence molecular processes, including replication, transcription, translation, and recombination, has implications for the observed microbial phenotypes, including latency, virulence, rapid evolution, and radioresistance.

Quadruplexes are increasingly being recognized as novel therapeutic targets in microbes. Several reports convincingly demonstrate antimicrobial activity of quadruplex-binding ligands against clinically challenging pathogens, including HIV-1, HCV, Ebola virus, *Plasmodium falciparum*, and *Mycobacterium tuberculosis*.

Biological Role of G4s

G4s are nucleic acid secondary structures consisting of stacked planar G-tetrads. An intramolecular quadruplex is formed by four tracts of two or more guanines each, separated by nucleotide residues of one to seven bases in length [$G_n N_x G_n N_y G_n N_z G_n$; n = 2+, x,y,z $\geq$ 1 and $\leq$7] (Figure 1). An intermolecular quadruplex is formed by guanine runs present in two or four different nucleic acid strands. Adjacent guanines in a G-tetrad are connected via hydrogen bonds on their Hoogsteen faces [1]. The loop sequences (N_x, N_y, and N_z) connect the G-runs. Motifs with loop sequences that are over seven nucleotides long can also form G4s, albeit in a context-dependent manner [2,3].

The transient formation of G4s under thermodynamically favorable conditions has important regulatory roles dictated by the genomic location. G4s are ubiquitously found in the telomeres of eukaryotes [4]. The formation of G4s by the G-rich telomeric repeats inhibits extension of telomeres by telomerase; thus stabilization of G4s in telomeres with ligands represents a potential anticancer strategy [5]. Nearly 50% of human genes have a G4 motif in their promoter region [6]. Importantly, oncogenes like c-Myc, VEGF [7], and KRAS [8] are negatively regulated by their promoter-borne G4s [5]. G4 structures can also form in RNA. Quadruplexes formed in the 5′ UTR of the mRNA inhibit cap-dependent translation (e.g., NRAS and BCL-2) and enhance IRES-mediated cap-independent translation (e.g., hVEG-F and FGF2) [9,10]. Besides, G4s also influence other molecular mechanisms in RNA biology such as splicing, ribosomal frameshifting, mRNA localization, repeat-associated non-AUG (RAN) translation, and maturation of miRNAs (reviewed in [10]). Furthermore, formation of an intermolecular hybrid quadruplex (HQ) between nontemplate DNA and nascent mRNA acts as a transcription-termination signal [10]. In addition to gene expression, the spatial association of quadruplex motifs with the recombination hotspots in the human genome implicates quadruplexes in recombination [11].

[1]Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi, India

*Correspondence:
vperumal@bioschool.iitd.ac.in
(P. Vivekanandan).

 https://doi.org/10.1016/j.tim.2018.08.011

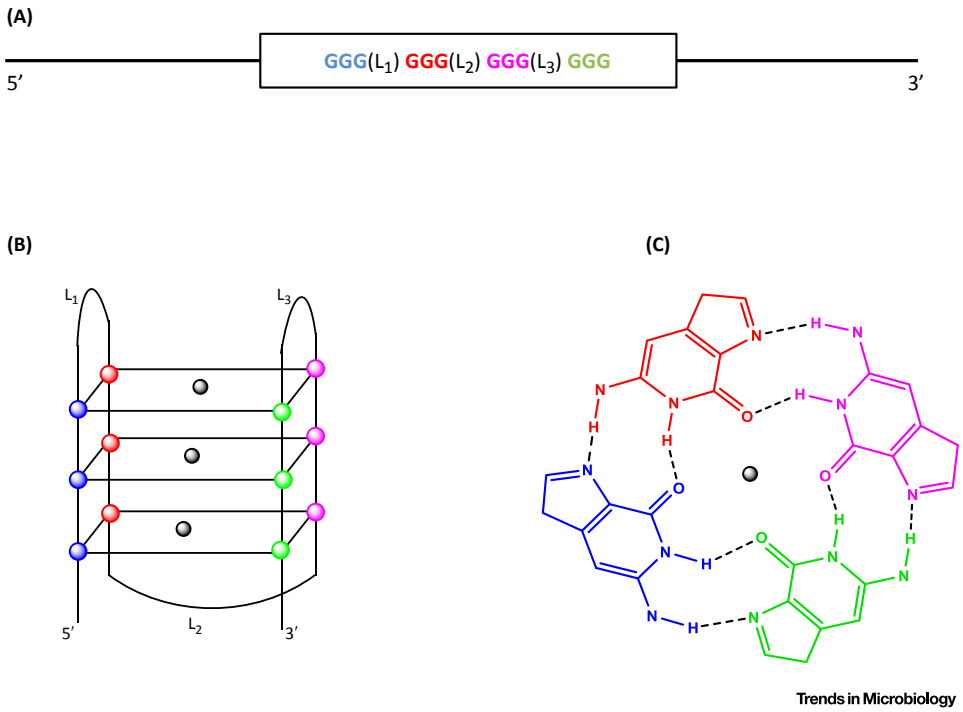


Figure 1. Structure of a G-Quadruplex (G4). (A) The nucleotide sequence of a G4 motif with varicoloured G-triplets separated by loop sequences (L_1, L_2, L3). (B) A 2D representation of a typical G4 fold showing the three planar quartets. The spheres at the vertices of the quartets represent one G from each of the four G-triplets. The black sphere at the centre denotes the central metal cation (Na^+, K^+) needed to stabilize the G4 structure. (C) A top view of a planar G-quartet showing the Hoogsteen bonds (dashed lines), the atoms thereof, and a cation in the central cavity. Figures are not drawn to scale.

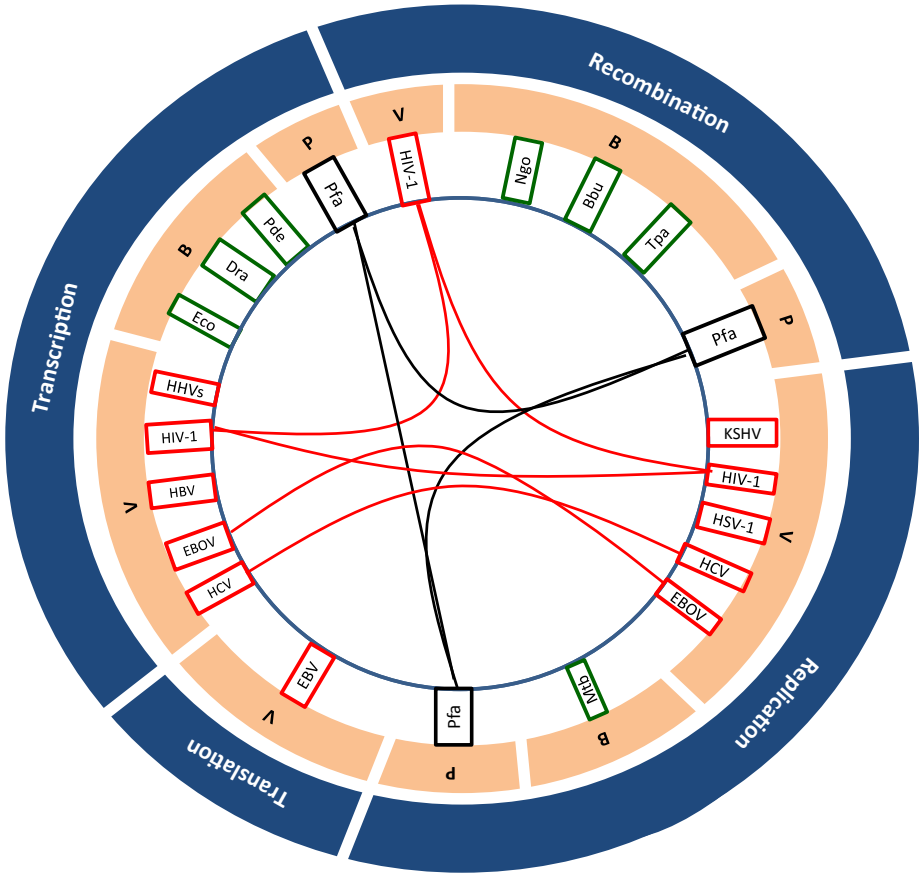
The repertoire of cellular proteins binding G4s is both structurally and functionally diverse; it comprises a number of zinc-finger transcription factors (SP1, MAZ, PARP, CNBP), splicing factors (U2AF), proteins of the shelterin complex, RNA-binding proteins such as hnRNPs and RHAU, and RGG-box-containing multifunctional proteins, including nucleolin and FMRP [12–14]. Persistence of G4 structures can dysregulate the cellular activities they control and also compromise genomic integrity [15,16]. Helicases, including FANCJ, Pif1, DHX36 [17], BLM, and WRN, can unwind the G4s in eukaryotic genomes. Extensive research on G4s has led to the identification of G4 ligands which are compounds that can specifically bind to these nucleic acid secondary structures [18].

Not much was known about G4s in microbial genomes about a decade ago. In the last few years, the roles of G4s in microbes have been increasingly recognized (Figure 2). In this review, we aim to provide insights on the biological role of quadruplexes in microbial genomes with an emphasis on recent findings.

Role in Virulence of Pathogens

In this section, we discuss three virulence-related microbiological features regulated by G4s.

Virulence factors are biomolecules produced by the pathogen that enable them to successfully establish infection and multiply in the host. They include adherence factors, immunomodulators, drug efflux pumps, and toxins.



V	Viruses
B	Bacteria
P	Parasites
HIV-1	Human immunodeficiency virus 1
EBOV	Ebola virus
HHVs	Human herpes viruses
HCV	Hepatitis C virus
EBV	Epstein–Barr virus
KSHV	Kaposi's sarcoma-associated herpesvirus
HSV-1	Herpes simplex virus-1
Eco	*Escherichia coli*
Dra	*Deinococcus radiodurans*
Pde	*Paracoccus denitrificans*
Ngo	*Neisseria gonorrhoeae*
Bbu	*Borrelia burgdorferi*
Tpa	*Treponema pallidum*
Mtb	*Mycobacterium tuberculosis*
Pfa	*Plasmodium falciparum*

Trends in Microbiology

Figure 2. Molecular Mechanisms Affected by G-Quadruplexes (G4s) in Microorganisms. The functional roles of G4s identified in microbes are included in this graph. The segments in the outermost circle (blue) denote the molecular biological processes regulated by G4s. The segments of the intermediate circle (orange) indicate the types of microbe affected for each biological process (see the legend beside the graph). The specific microorganism under each type is denoted either by an abbreviation (for viruses) or a three-letter notation (for bacteria and parasites) in the vertical bars in the innermost circle (see the legend beside the graph). The vertical bars in the innermost circle are colour-coded for bacteria (green), viruses (red), and parasites (black). The curved lines connect a given microbe with multiple quadruplex-regulated biological processes. Red curved lines connect viruses, and blue curved lines connect parasites.

Antigenic Variation

Depending on the tissue environment, pathogens have specific surface adaptations, including pili, fimbriae (in bacteria) and host cell receptor-binding glycoproteins (in viruses and parasites), all of which enable entry into the host. The surface-exposed proteins are at the interface of host–microbe interaction and are highly antigenic. The surface proteins of some pathogens are continuously altered by antigenic and phase variation to overcome host adaptive immune responses. Besides sequence mutation and natural competence to DNA transformation [19], the molecular basis of antigenic variation (Av) also involves recombination of genomic segments leading to the production of altered surface proteins [20]. Interestingly, G4 motifs have been identified at the recombination sites associated with Av in bacteria and parasites [21,22].

Intramolecular quadruplexes have been identified to play potential roles in the Av of (i) *pilE* in *Neisseria gonorrhoeae*, the bacterium that causes gonorrhea, (ii) *vlsE* in the Lyme disease agent *Borrelia burgdorferi*, and (iii) *tprK* in *Treponema pallidum* [23–25]. Conventionally, Av by gene conversion involves the unidirectional transfer of genetic segments from the donor loci, a tandem array of silent alleles of the surface protein, to a downstream recipient locus that actively expresses the gene encoding the surface protein; this process is assisted by recombinases. Pili are hair-like appendages made up of pilin proteins which are present on the bacterial cell surface and exhibit Av. In *N. gonorrhoeae*, it has been demonstrated experimentally that the G4 formed near the *pilE*, a pilin-expression locus, binds RecA and provides a topological advantage for the nicking process essential to initiate recombination [26]. Deletion of this quadruplex motif suppresses Av in *Neisseria*.

Lyme disease is caused by a tick-borne bacterium belonging to the genus *Borrelia*. The *vls* locus in *B. burgdorferi* is associated with Av. The coding strand of the *vls* locus has over a 100-fold enrichment of guanine (G)-runs of at least three nucleotides or more despite the preference for AT-rich codons [24]. These G-rich sequences form G4s and are suggested to play a role in recombination-mediated Av in *Borrelia* species.

T. pallidum is a spirochete that causes syphilis. TprK is a surface protein that undergoes Av in *T. pallidum*. G4-forming sequences were identified proximal to the TprK gene, indicating a possible role for these DNA secondary structures in Av among treponemes [25]. However, the proposed functional roles for quadruplexes identified in the Av loci of the two spirochetes are not supported by experimental evidence.

The family of erythrocyte membrane proteins-1 (PfEMP-1) is an important virulence factor of the malarial parasite, *Plasmodium falciparum*. Symptoms of malaria appear in about a week after exposure when the parasite enters red blood cells (RBCs) and digests hemoglobin [27]. Proteins of the PfEMP-1 family are expressed on the surface of infected erythrocytes during the asexual life cycle in man and are encoded by *var*, a family of 60 genes which are predominantly present in the subtelomeric region of chromosomes [28]. The *var* genes undergo recombination (indels and translocations) to facilitate sequence variation in PfEMP1 and immune evasion. Interestingly, about a quarter of all putative quadruplex motifs in the *P. falciparum* genome are associated with the promoters of the *var* genes [29]. Stanton *et al.* identified a close association between the recombination breakpoints and G4 motifs in *P. falciparum* [30]. Breakpoints were found to occur proximal to quadruplexes, especially in subtelomeic regions, indicating that G4s have a role in *var*-associated recombination. Although recombination in *P. falciparum* genomes occurred proximal to quadruplex motifs, the median distance of a G4 from a breakpoint was about 16 kb. The specific mechanisms underlying G4-assisted *var* gene recombination are not fully understood; nonetheless, a potential role for DNA repair has been speculated.

Recombination-Mediated Microbial Evolution

In addition to contributing to antigenic diversity in microbes (discussed above), G4s facilitate generation of genetic heterogeneity and evolution of HIV-1, the causative agent of AIDS. The recombination rate of HIV-1 stems from the ability of the reverse transcriptase (RT) to switch between RNA templates and generate chimeric proviral DNA. Recombinant HIV-1 strains have been associated with increased transmission efficiency and resistance to anti-HIV therapy [31]. The two positive-sense RNA strands of HIV-1 are held together by hairpin loops in the dimerization site (DIS) at the 5′ end of each of the RNAs [32]. This allows for strand transfer by RT, making the region a recombination hotspot. Similar to the hairpin loops, intermolecular G4s tether the recombining segments, thus bringing them into each other's proximity to

promote initiation of recombination. Independent studies identified intermolecular G4 motifs in three regions of the HIV-1 RNA (i) a 130-nt region comprising the DIS and 5′ portion of the *gag*, (ii) central polypurine tract (cPPT), and (iii) the U3 region on either termini of RNA [33–36]. Under *in vitro* conditions, synthetic RNA oligonucleotides corresponding to these G4s caused pausing of RT in the presence of potassium ions. Moreover, in an *in vitro* strand transfer assay, efficient switching over of RT between templates was observed under conditions that promote quadruplex formation (presence of potassium ions). These studies suggest that quadruplexes allow for dimerization of the HIV-1 genome at multiple loci along its length, thus contributing to recombinogenicity and rapid evolution of HIV-1.

In addition to HIV-1 evolution, other functional aspects of G4s discussed in this review, and the selective retention or exclusion of G4s from specific genomic loci in bacteria and yeast, indicate that G4s play a role in the evolution of microbes (Figure 3) [30,37–42].

Gene Expression and Packaging of Virions

Long terminal repeats (LTRs) present on either termini of the HIV-1 genome enable integration of the HIV-1 provirus into the host genome and contain within them the necessary genomic elements for expression and control of HIV-1 genes. Three overlapping quadruplex motifs were identified in the U3 promoter region of the LTR between positions −105 and −48 [43]. The motifs encompass the binding sites of the two transcription factors, NF-KB and SP1. These

Figure 3. The Link between G-Quadruplexes (G4s) and Microbial Evolution. The conservation or elimination of G4s from specific genomic locations within a species or across related species of microbes implicates these nucleic acid secondary structures in microbial evolution. Besides, G4s directly participate in mechanisms that generate diversity in the microbial populations, such as antigenic variation and recombination. *E. coli*, *Escherichia coli*; *S. cerevisiae*, *Saccharomyces cerevisiae*.

G4s negatively regulate the activity of the LTR promoter and hence the replication of HIV-1. Interestingly, the presence of the quadruplex in the LTR promoter is not restricted to HIV-1 but is evolutionarily conserved among the primate lentiviruses [41].

Human herpesviruses (HHVs) are large double-stranded DNA (ds-DNA) viruses that infect a variety of tissues. In HHVs, putative promoter regions have higher G4 densities than the coding regions, suggesting a regulatory role for G4s in gene expression [44]. G4s in the promoters of UL2, UL24, and K18, all of which have previously established roles in virulence, were found to be negative regulators of promoter activity (Figure 4A) [45–48]. Herpesvirus genes are divided into immediate early (IE), early (E), and late (L) based on the time at which they are expressed in the replication cycle. IE genes act as *trans*-activators or *trans*-repressors of the E and L genes. IE genes are expressed within a few hours of virus entry into the host cells. Interestingly, the regulatory regions of IE genes were particularly enriched for G4 motifs [44].

Besides transcription, a recent report also implicates G4s in the packaging of herpesvirus genomes [49]. An earlier study reports that, following concatemeric replication of HHV-1, the cleavage of unit length genomes and their encapsidation is achieved by the binding of virus proteins to a DNA secondary structure formed by a DNA packaging sequence (*pac-1*) [50]. It has now been identified that the DNA secondary structure formed by *pac-1* is a G4 (Figure 4B) [49]. In fact, the *pac-1* sequences of all the eight human herpesviruses contain a highly conserved G4 motif that predominantly forms intermolecular quadruplexes.

Human papilloma virus (HPV) is a DNA virus that causes warts and cervical cancer. Tluckova *et al.* identified the presence of three-tetrad G4 motifs in the long control region (LCR) and in the coding regions of E1, E2/E4, and L2 proteins of eight HPV types [51]. Interestingly, two-tetrad quadruplexes were identified in the same genomic regions (i.e., E1, E2/E4 and LCR) of manatee papilloma viruses [52]. In papilloma viruses, the LCR contains a number of *cis*-acting regulatory elements for virus replication and transcription that play a role in determining tissue tropism [53]. The early proteins (E1–E7) are nonstructural proteins and have pivotal roles in the modulation of the host regulatory network while the late proteins (L1, L2) are required for virion assembly [54]. The specific biological roles for the G4s in papilloma viruses remain to be discovered; however, potential functional roles in gene expression have been speculated for these DNA secondary structures based on their key genomic locations in certain HPV types.

Host G4-Binding Proteins Encoded by Microbes

Severe acute respiratory syndrome-coronavirus (SARS-CoV) is an enveloped virus with a positive-sense single-strand RNA genome. Nonstructural protein (nsp3) is a multidomain protein that is a part of the replication/transcription complex (RTC) of the virus [55]. The SARS-unique domain (SUD), exclusively present in the nsp3 of SARS-CoV, is believed to contribute to the higher pathogenicity of SARS-CoV as compared to other human coronaviruses [55]. Interestingly, SUD was identified to bind G-runs and the more ordered G4s, in both DNA and RNA [56]. The G4-binding property is mapped to the M domain nested within the SUD and is indispensable for the replication and transcription of the virus [57]. Putative G-rich targets of SUD include host mRNAs encoding proteins that regulate key cellular processes such as apoptosis and cell proliferation. Therefore, G4-binding microbial proteins may potentially play a role in the modulation of key cellular proteins and signaling pathways in the host [55,56].

Role in Virus Latency

The latency programme of viruses allows them to survive inside the host and protects them against the immune surveillance of the host. The expression of latency-associated genes leads

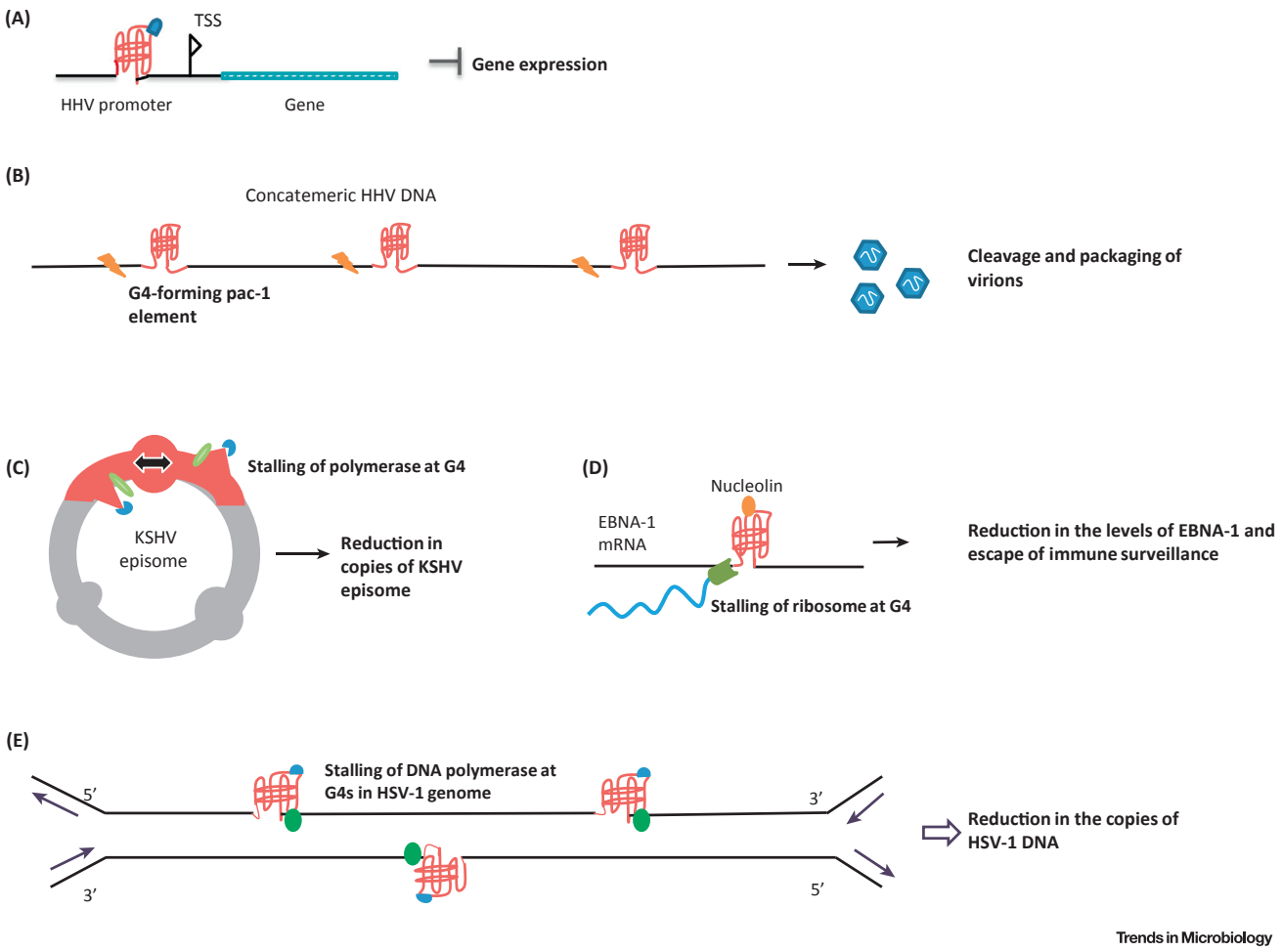


Figure 4. Roles of G-Quadruplexes (G4s) in Human Herpesviruses (HHVs). (A) Quadruplexes (red) present in the promoter regions of the HHVs negatively regulate gene expression. Addition of G4-binding ligand (blue) further augments the inhibition of gene expression. (B) HHV genomes exist as concatemers during replication. A quadruplex formed at the *pac-1* signal (packaging signal) acts as a scaffold for the protein machinery that cleaves and encapsidates unit-length genomes. The sites of cleavage proximal to the quadruplex-forming *pac-1* signals are shown in orange. (C) Binding of ligands (blue objects) to G4s (red triangular kinks) present near the origin of bidirectional episomal replication (red bulge with two-headed arrow) in the terminal repeat region (red portion) of Kaposi's sarcoma-associated herpesvirus (KSHV) episome, prevents progression of DNA polymerase (green ovals). The slippage of the polymerase leads to replicative stress that activates dormant origins (grey bulges on either sides). The net effect of the ligand binding to quadruplexes is a reduction in the number of KSHV episomal copies. (D) Quadruplex (red) formed in the Epstein–Barr virus (EBV) nuclear antigen-1 (EBNA-1) mRNA causes stalling of the ribosome-nascent chain complex (green and blue), resulting in inhibition of translation. EBV is thus able to maintain the levels of EBNA-1 below the detection threshold, evading immune response during latency. The binding of nucleolin (orange) to the quadruplex further reduces EBNA-1 protein levels. (E) Ligands (blue objects) bind to G4s (red) in the herpes simplex virus-1 (HSV-1) genome and stabilize them, causing polymerase (green ovals) stalling on the leading and lagging strands, inhibiting HSV-1 DNA replication. Arrows indicate direction of replication. Figures A through E are not drawn to scale.

to heterochromatinization of the virus genome and the inhibition of proteins necessary for virus replication [58,59]. HHVs are an example of viruses capable of causing latent infections. During latency, their large linear ds-DNA genome circularizes to form an episome. The episome is replicated and segregated between the daughter cells with every cell division in the host, leading to persistence of the virus.

Herpesvirus genomes consist of unique and repeat regions. Multiple reiterations of G-rich repeat units capable of forming G4s (known as 'repetitive G-quadruplex motifs' – RGQMs) has

been noted in the eight HHVs [44]. Such G4-forming repeats have recently been identified to be functionally relevant in the latency of Kaposi's sarcoma-associated herpesvirus (KSHV) or HHV-8. KSHV is an oncogenic virus that has acquired a number of its genes by molecular piracy. The terminal repeat (TR) region of KSHV genome is enriched for quadruplex-forming sequence motifs with each repetitive element (about 800 bp), having 12 and 16 putative quadruplex-forming sequences in the top and the bottom strand respectively [60]. The TR region harbours the only origin for replication of viral episomes. The preponderance of quadruplexes in the TR region is relevant in the regulation of episomal replication. Stabilization of the G4s with PhenDC3 (2,N9-bis(1-methylquinolin-3-yl)-1,10-phenanthroline-2,9-dicarboxamide) or TmPyP4 (5,10,15,20-tetrakis(1-methylpyridin-1-ium-4-yl)-21,22-dihydroporphyrin) halts the replication forks at the boundary of the TR; replicative stress ensues and results in the firing of the otherwise dormant origins in the viral episome. Consequently, the two replication parameters – the number of replication forks and the number of origins – increase in a dose-dependent manner. Furthermore, a transient replication assay also indicated inhibition of KSHV DNA replication and a decrease in the number of copies of the viral episome, post-treatment with PhenDC3 (Figure 4C). Interestingly, even upon removal of the ligand, the number of KSHV genome copies was lower as compared to that in cells without the ligand, indicating reduction of virus episomes.

Integration into the host genome is another strategy employed by herpesviruses for latent survival. Human herpesvirus 6A, the causative agent of roseola infantum, undergoes stable integration at the telomeric ends of the host chromosomes by homologous recombination. About 1% of the human population has the congenital presence of HHV-6a due to integration of the virus in germline cells [61]. The formation of G4s by the telomeric repeats (TTAGGG) is well documented [4]. A recent study demonstrated that the stabilization of telomeric G4s by BRACO-19 (N-[9-[4-(dimethylamino)anilino]-6-(3-pyrrolidin-1-ylpropanoylamino)acridin-3-yl]-3-pyrrolidin-1-ylpropanamide) significantly reduced the integration frequency of HHV-6A in telomerase-expressing cells [61]. The authors argue that stabilization of the telomeric G4s interferes with telomerase activity, resulting in reduced chromosomal integration of HHV-6a.

Epstein–Barr virus (EBV) is an oncogenic herpesvirus that is associated with B cell lymphoma and nasopharyngeal carcinoma. During latency, EBNA1, a genome-maintenance protein (GMP), tethers the circular EBV episome to cellular chromatin and ensures its transmission to daughter cells on completion of each cell cycle [62]. EBNA-1 also regulates host and viral transcription. It is imperative that the synthesis of the latency proteins is tightly controlled lest they are processed and presented to MHCs as antigens, defeating the primary biological function of this group of proteins.

Murat *et al.* identified putative quadruplex motifs in EBNA1 and GMPs encoded by other gamma herpesviruses [63]. Furthermore, they also describe quadruplex-mediated repression of EBNA1 protein levels. The translation of several oncogenic proteins in humans is known to be controlled by quadruplexes in the 5′UTR or coding region [9]. EBNA-1 is a key player in EBV-induced oncogenesis [64,65]. Interestingly, the level of the EBV oncogene, EBNA-1, is maintained by the folding and unfolding dynamics of its mRNA-borne quadruplex. The segment of EBNA1 mRNA that encodes its glycine–alanine repeat (GAr) domain was found to be rich in G4 motifs. Polysome distribution profiling, *in vitro* translation, and cell culture experiments identified that formation of quadruplexes obstructs the progress of ribosome machinery, resulting in low levels of EBNA1 and a concomitant decrease in presentation to T cells (Figure 4D). In addition, nucleolin was identified to bind the G4 formed in the GAr domain of EBNA-1 mRNA [66]. This interaction limits the synthesis of EBNA-1 to levels that allow

persistence of the virus and evasion of the host immune system. Mutation of the EBNA-1 mRNA quadruplex or absence of nucleolin resulted in alleviation of translation inhibition and increased presentation of EBNA-1.

Besides regulating synthesis, G4s also come into play in the functional aspects of EBNA-1. As a GMP, EBNA-1 is involved in episomal replication and attachment to metaphase chromosomes. These functions are carried out by the linking regions LR1 and LR2 present in EBNA-1. LR1 and LR2 bind cellular RNA-quadruplexes to recruit origin recognition complex (ORC) to OriP, the origin of episomal replication in EBV [67]. BRACO-19 outcompetes EBNA-1 in binding to the intermediary RNA quadruplex, thus inhibiting the replication of EBV in latently infected cells. Consequently, a reduction in the EBV copy number and its attachment to metaphase chromosomes was observed by q-PCR and flow cytometry, respectively.

Viruses exploit the molecular machinery of the host for successful infection. They utilize the quadruplex-binding abilities of host proteins to regulate the dynamics of quadruplex formation in their genome and the downstream effects thereof. The quadruplex in the LTR promoter of HIV-1 provirus binds the host protein nucleolin [68]. Nucleolin stabilizes the quadruplex and represses the transcriptional activity of the LTR promoter, allowing the virus to enter latency. Interestingly, heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2/B1), a host protein, binds and unfolds the quadruplex in the LTR promoter of HIV-1 provirus, leading to enhanced transcription [69]. Taken together, these results suggest that G4s in virus genomes may interact with host proteins not only to facilitate virus latency but also to revoke viruses from latency. These studies on G4–protein interactions highlight how G4s contribute to the molecular milieu of host–pathogen interaction.

G4s as Antimicrobial Targets

The emergence of antimicrobial resistance is a major limiting factor in the management of infectious diseases such as AIDS and tuberculosis. Indiscriminate use of antimicrobial agents and patient noncompliance contribute to the emergence of antimicrobial resistance [70]. The need to develop or identify novel therapies as well as novel therapeutic targets to tackle antimicrobial resistance has been increasingly recognized.

The identification of G4s in microbial genomes as targets of antimicrobial therapy has led to the identification of novel antimicrobial agents. G4s have been shown to inhibit the transcription or translation of structural and nonstructural proteins in viruses, deleteriously affecting the virus loads and their pathogenicity; the stabilization of these quadruplexes with ligands has been investigated as a potential mechanism for targeting viruses. For example, the HIV-1 *nef* gene contains quadruplex motifs that inhibit synthesis of the Nef protein [71]. The addition of TmPyP4, a quadruplex-binding ligand, further lowered the expression of this protein. The Nef protein is required for efficient viral entry, integration of provirus into host genome, and replication in the host cells [72]. It also modulates a number of cellular immunity factors like CD4 and MHC I to enhance the survival of the virus [73]. Defects in the *nef* gene or its deletion from the virus genome affect the infectivity of the virus and delay the progression to AIDS [74,75].

HCV is an enveloped positive-sense RNA virus. Chronic HCV infection is a major cause of hepatocellular carcinoma (HCC). A quadruplex motif in the core gene inhibits the synthesis of core (capsid) protein and replication of HCV [76]. Stabilization of this quadruplex in the HCV core gene with ligands results in stalling of the viral RNA-dependent RNA polymerase (RdRp) at the G4 motif, resulting in decreased HCV core protein levels.

Ebola virus, a negative-sense RNA virus, causes hemorrhagic fevers and represents one of the well studied zoonotic filoviruses. A 27-nt long G4 motif was identified in the L gene of Ebola virus that encodes the RdRp [77]. Stabilization of this G4 motif in the L gene with a quadruplex-binding ligand led to reduced transcription of the L gene. The RdRp encoded by the L gene is indispensable for the life cycle of Ebola virus. The stabilization of the G4 motif in the L gene with ligands reduces the replication competence of Ebola virus. Furthermore, the authors report G4 ligands as more potent antiviral agents as compared to ribavirin.

Negative regulation of virus transcription, translation or replication by quadruplex motifs in virus genomes forms the basis of using G4-binding ligands as antiviral agents. Considering that the G4s that negatively regulate virus replication are retained in virus genomes during evolution, it is likely that viruses may stand to benefit from these G4s in their genomes. Further research in this area may help us better understand this conundrum. In the last few years, the antimicrobial activity of quadruplex-binding ligands has been demonstrated in bacteria and parasites in addition to viruses (Table 1).

Although G4-binding ligands appear to be promising as potential antimicrobial agents, an important but often ignored aspect is specificity. It is very likely that G4 ligands will bind several host G4 motifs, which outnumber the microbial quadruplex motifs. Studies investigating the undesired interaction of G4-binding ligands with G-quartets in the host genome may help us to better understand the therapeutic potential of this class of drugs.

Across different types of microbes, the modulation of transcription by G4s and its cascading effect on specific microbial phenotypes appears to be a common theme (Figure 5).

Table 1. Antimicrobial Activity of G4-Binding Ligands

Pathogen	G4 ligand	Suggested mode of action[a]	Refs
Herpes simplex virus-I	BRACO-19, c-exNDI-2	Inhibition of HSV-1 DNA replication (Figure 4E)	[78,79]
HIV-1	BRACO-19	Inhibition of reverse transcription and transcription by binding to G4 in the U3 region of RNA and proviral DNA, respectively	[80]
	TmPyP4	Inhibition of Nef-dependent HIV replication	[71]
	c-exNDI-2	Negative regulation of HIV-1 transcription by binding to the G4 in the LTR	[81]
Mycobacterium tuberculosis	BRACO-19, c-exNDI	Inhibition of bacterial growth (no specific mechanism is elucidated)	[82]
Plasmodium falciparum	Quarfloxin	Deregulated expression of G4-associated genes and inhibition of ring-stage parasites	[83]
Ebola virus	TmPyP4	Inhibition of the L (polymerase) gene expression	[77]
Hepatitis C virus	PDP and TmPyP4	Inhibition of core gene expression	[76]

[a]Some of these require experimental validation.

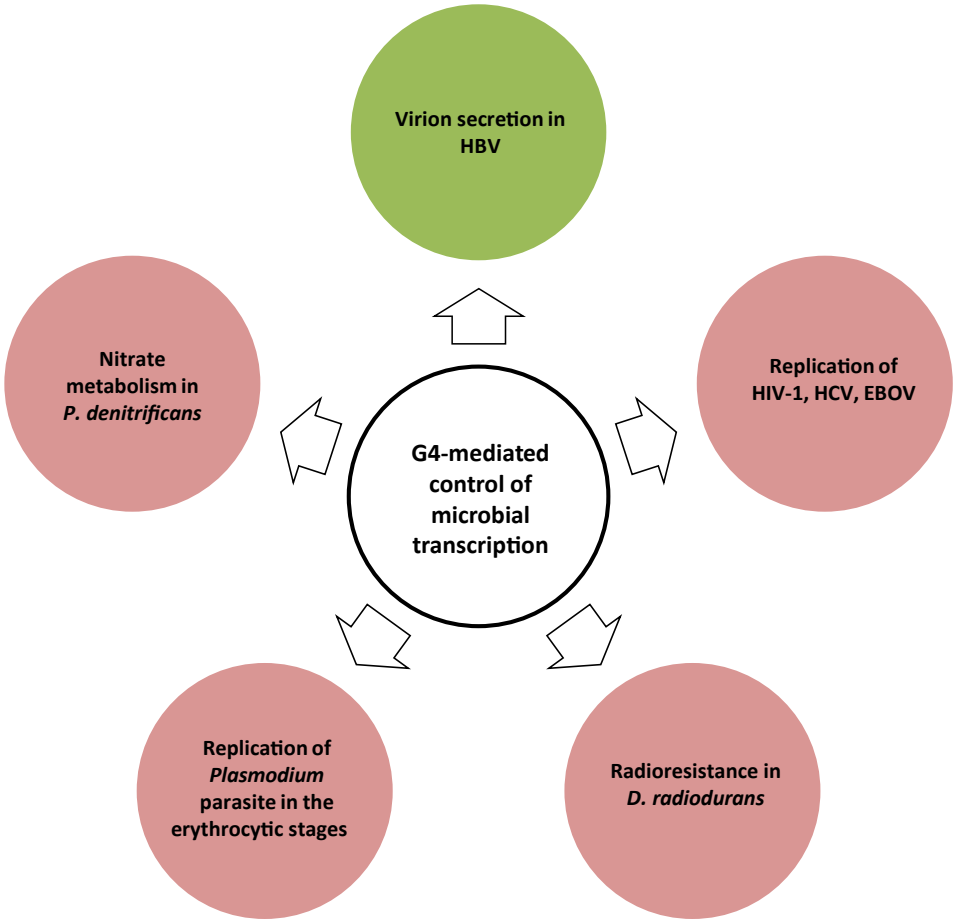


Figure 5. Phenotypic Effects of the Modulation of Transcription by G-Quadruplexes (G4s). G4-mediated control of microbial transcription appears to be a common theme leading to tangible differences in the phenotype of microbes. G4s in microbial genomes may regulate microbial transcription either negatively (pink circles) or positively (green circle). *D. radiodurans*, *Deinococcus radiodurans*; *P. denitrificans*, *Paracoccus denitrificans*; EBOV, Ebola virus.

Role in Genotype-Specific Pathogenicity

HBV is an enveloped hepatotropic DNA virus that replicates with an RNA intermediate. Persistent infection with HBV can cause serious liver damage leading to cirrhosis and HCC. The HBV genome exhibits a high degree of genetic variability owing to the lack of proof-reading ability in the HBV reverse transcriptase. As a result, HBV is classified into 10 HBV genotypes (A through J) with an intergenotypic sequence variation of at least 8% [84]. The HBV genotypes differ in transmissibility, virus loads, response to antiviral therapy, and ability to cause liver disease [84,85]. However, genotype-specific regulatory mechanisms in HBV remain elusive. We had recently identified a G4 motif as a genotype-specific regulator of HBV replication [40]. A conserved three-tetrad G4 motif, 190 bp upstream of the transcription start site, was identified in the preS2/S promoter of HBV genotype B. This motif was virtually absent in the rest of the HBV genotypes. This quadruplex specifically enhanced the transcription of the preS2/S transcript and the production of HBV surface antigen (HBsAg). Point mutations disrupting the G4 motif in the preS2/S promoter of HBV genotype B led to a reduction in HBsAg production resulting in a fivefold reduction in virion secretion [26].

Role in Control of Radiation Resistance

Deinococcus radiodurans is an extremophilic bacterium tolerant to ionizing radiations such as gamma rays and UV rays. Beaume *et al.* analyzed the promoters of *D. radiodurans* and found that G4 motifs were particularly enriched within the 200 bp upstream region in genes that confer radiation resistance; these include *recA*, *recF*, *recO*, *recR*, *recQ*, and *mutL*, all of which are involved in recombinational DNA repair [86,87]. Interestingly, the addition of an intracellular G4-binding ligand led to a marked reduction in the expression of genes associated with radiation resistance, thus rendering *D. radiodurans* sensitive to radiation. The ability of G4 motifs to modulate radioresistance in *D. radiodurans* sheds light on how these DNA secondary structures contribute to microbial tolerance to environmental pressures by regulating the transcriptional machinery.

Role in Metabolism

Paracoccus denitrificans is a facultative anaerobe capable of metabolizing nitrogen, nitrate, and ammonia. Reduction of nitrate or nitrite to dinitrogen, a cellular process known as denitrification, is associated with the *nasABGHC* gene cluster in *P. denitrificans* [88]. This nitrate-assimilatory system (*nas*) is regulated by a two-component NasS–NasT system. NasT is an effector molecule that positively regulates transcription of *nas* genes by acting as an anti-termination signal [88]. The GC-rich genome *of P. denitrificans* contains a three-tetrad quadruplex motif 150 nucleotides upstream of the *nasT* gene [89]. Stabilization of the G4 by ligands (TmPyP4 and a benzophenoxazine ligand) or by cations (KCl) inhibited the transcription of the *nasT* gene. Similarly, the presence of G4-stabilizing ligands inhibited the growth of *P. denitrificans* in media containing nitrate as the sole source of nitrogen. This work on *P. denitrificans* highlights a role for G4-linked transcriptional control in modulating specific metabolic pathways.

Studies investigating bacterial and yeast genomes found an enrichment of G4s in promoters of genes involved in carbohydrate, amino acid, and nucleotide metabolism [37,86,90]. Although the functional significance of the G4s in the promoters of bacterial and yeast genes involved in metabolism remains to be demonstrated, it may not be too speculative to suggest a possible role for these DNA secondary structures in regulating the synthesis of macromolecules in bacteria and yeast by modulation of key metabolic pathways.

Role in RNA Editing

Trypanosoma brucei is a parasitic kinetoplastid that causes African sleeping sickness in humans. Mitochondrial transcripts of kinetoplastid organisms undergo extensive editing post-transcription; this mRNA editing involves deletion or insertion of 'U' residues at multiple locations specified by the anchoring of guide RNAs (gRNAs) encoded by the mitochondrial genome [91]. The nucleotide composition of the pre-mRNAs may be potentially altered by up to 50% as a result of editing, which is referred to as pan-editing [92]. Matthias-Leeder *et al.* analyzed nine mRNAs of *T. brucei* and found that the guanosine (G) content is lowered to about 19% from about 34% during pan-editing [93]. Importantly, the authors used computational methods to demonstrate the progressive decrease in G4 content during pan-editing. Therefore, pan-editing in African trypanosomes has been suggested as a G4-resolving process that leads to the generation of G4-free translatable ORFs. The authors also propose the formation of DNA/RNA hybrid G4s (HQ) between the nontemplate DNA strand and pre-edited transcripts. Furthermore, it is speculated that the formation of HQs is involved in the termination of transcription and the initiation of mitochondrial replication. Thus, quadruplexes may play a crucial role in switching between the two mutually exclusive processes of mitochondrial replication and transcription in trypanosomes.

Concluding Remarks

Among the microorganisms that contain a G4 in their genome, the over-representation of viruses associated with cancer, namely, KSHV, EBV, HCV, HBV, and HPV, is noteworthy [40,51,60,63,76]. The existence of these secondary structures in zoonotic agents such as Ebola virus and vector-borne pathogens such as Zika virus, *Plasmodium* spp., *B. burgdorferi*, and *T. brucei*, is particularly interesting [24,42,93,94]. From an evolutionary perspective, it may be of interest to identify G4-influenced adaptations, if any, that facilitate the survival of these microbes in different hosts.

Repeat regions in herpesviruses contain important regulatory elements for replication, packaging, latency, and reactivation [95,96]. The existence of RGQMs amplifies the G4 load of the genomes of HHVs manifold [44]. Such G4-forming iterative G-rich units also comprise the simple sequence repeats (SSRs) present in the noncoding regions of *Nostoc* sp. and *Xanthomonad* spp. [97]. Bacterial SSRs are known to be implicated in antigenic and phase variation. Given the functional significance of repeat sequences in microbial genomes it may be interesting to investigate the link between the tandem array of G4s and molecular processes related to microbial pathogenesis. Recent reports on G4 motifs in viruses infecting nonhuman hosts shed light on how G4s have been exploited by viruses for virulence and genome regulation throughout evolution [52,98].

The identification of G4s in microbial genomes has opened up new avenues for therapeutics; additional studies on the specificity of G4-binding ligands and their undesired effects may help us to better understand the therapeutic potential of this novel group of antimicrobial agents. Host protein–microbial G4 interaction or the host G4–microbial protein interactions at the molecular interface of the host and microbe during infection are fascinating and merit further investigation [55,56,66–69]. It would be interesting to understand if such interactions defend the host or demonstrate yet another mechanism of microbial pathogenesis. Intuitively, the threshold to transcend the thin line between these two opposing outcomes may be subject to complex regulation which may be important for an understanding of the therapeutic potential of targeting this host–microbe interaction.

The nucleotide sequences complementary to G4 motifs are cytosine-rich and may form i-motifs which are higher-order nucleic acid structures formed in near-neutral or acidic pH [99]. Recently, i-motifs were visualized in human cells [100]. It may be particularly interesting to study the molecular dynamics of G4s and i-motifs and its impact on microbial pathogenesis and evolution.

Outstanding Questions

Do G4-unwinding helicases encoded by the host interact with G4s in pathogens during infection? If so, are such host–microbe interactions double-edged swords that help the host to defend some infections and prove to be detrimental in others? It may also be interesting to study if pathogens modulate the expression profile of host-encoded G4 helicases during infections.

Do G4s represent structural elements for conservation of energy in bacterial metabolism? Is G4-mediated control of gene expression exercised only in the early stages of central dogma to conserve energy? Are the prokaryotic ORFs devoid of G4s to cut down on the energy cost of resolution of G4 for translation? Does the presence of G4s as an additional level of coordinated gene control in operons also signify the same?

As in *D. radiodurans,* are the unique adaptations of other extremophiles regulated by G4s?

The primary sequence of viruses coevolves with that of their hosts. Nonetheless, it is not known whether nucleic acid secondary structures, such as G4s, in virus genomes coevolve with host genomes.

As in *Nostoc* spp. and *Xanthomonas* spp., are the plasmids of other bacteria also devoid of G4s? Given that plasmids harbor important bacterial genes needed for virulence, antibiotic resistance etc., and the ability of G4s to pause replication, are they selectively eliminated from extrachromosomal elements to minimize interference, if any, in vertical transmission?

Do the pathogenicity islands in bacterial genomes have unique G4 profiles?

Are the differences in the density and distribution of G4s in pathogens influenced by the ecological relationship they share with the host? In other words, are there differences in the genomic densities of G4s among parasites, symbionts, commensals, and mutualists?

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tim.2018.08.011.

References

1. Lane, A.N. *et al.* (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* 36, 5482–5515
2. Chambers, V.S. *et al.* (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* 33, 877–881
3. Guedin, A. *et al.* (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* 38, 7858–7868
4. Tran, P.L. *et al.* (2011) Stability of telomeric G-quadruplexes. *Nucleic Acids Res.* 39, 3282–3294
5. Patel, D.J. *et al.* (2007) Human telomere, oncogenic promoter and 5′ UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* 35, 7429–7455
6. Rhodes, D. and Lipps, H.J. (2015) G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* 43, 8627–8637
7. Sun, D. *et al.* (2011) Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene. *Nucleic Acids Res.* 39, 1256–1265
8. Cogoi, S. and Xodo, L.E. (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.* 34, 2536–2549

9. Bugaut, A. and Balasubramanian, S. (2012) 5′-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* 40, 4727–4741
10. Fay, M.M. *et al.* (2017) RNA G-quadruplexes in biology: principles and molecular mechanisms. *J. Mol. Biol.* 429, 2127–2147
11. Mani, P. *et al.* (2009) Genome-wide analyses of recombination prone regions predict role of DNA structural motif in recombination. *PLoS One* 4, e4399
12. Brazda, V. *et al.* (2014) DNA and RNA quadruplex-binding proteins. *Int. J. Mol. Sci.* 15, 17493–17517
13. Kumar, P. *et al.* (2011) Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide. *Nucleic Acids Res.* 39, 8005–8016
14. McRae, K.S.E. *et al.* (2017) On Characterizing the interactions between proteins and guanine quadruplex structures of nucleic acids. *J. Nucleic Acids* Published online November 9, 2017. http://dx.doi.org/10.1155/2017/9675348
15. Sauer, M. and Paeschke, K. (2017) G-quadruplex unwinding helicases and their function *in vivo*. *Biochem. Soc. Trans.* 45, 1173–1182
16. Estep, K.N. *et al.* (2017) G4-interacting DNA helicases and polymerases: potential therapeutic targets. *Curr. Med. Chem.* Published online November 16, 2017. http://dx.doi.org/10.2174/0929867324666171116123345
17. Chen, M.C. *et al.* (2018) Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* 558, 465–469
18. Ruggiero, E. and Richter, S.N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.* 46, 3270–3283
19. Seifert, H.S. and So, M. (1988) Genetic mechanisms of bacterial antigenic variation. *Microbiol. Rev.* 52, 327–336
20. Deitsch, K.W. *et al.* (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol. Mol. Biol. Rev.* 61, 281–293
21. Harris, L.M. and Merrick, C.J. (2015) G-quadruplexes in pathogens: a common route to virulence control? *PLoS Pathog.* 11, e1004562
22. Seifert, H.S. (2018) Above and beyond Watson and Crick: guanine quadruplex structures and microbes. *Annu. Rev. Microbiol.* 72, 49–69
23. Cahoon, L.A. and Seafert, H.S. (2009) An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. *Science* 325, 764–767
24. Walia, R. and 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
25. Giacani, L. *et al.* (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
26. Kuryavyi, V. *et al.* (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
27. Mawson, A.R. (2013) The pathogenesis of malaria: a new perspective. *Pathog. Glob. Health* 107, 122–129
28. Pasternak, N.D. and Dzikowski, R. (2009) PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite *Plasmodium falciparum*. *Int. J. Biochem. Cell Biol.* 41, 1463–1466
29. Smargiasso, N. *et al.* (2009) Putative DNA G-quadruplex formation within the promoters of *Plasmodium falciparum var* genes. *BMC Genomics* 10, 362
30. Stanton, A. *et al.* (2016) Recombination events among virulence genes in malaria parasites are associated with G-quadruplex-forming DNA motifs. *BMC Genomics* 17, 859
31. Burke, D.S. (1997) Recombination in HIV: an important viral evolutionary strategy. *Emerg. Infect. Dis.* 3, 253–259
32. Moore, M. and Hu, W. (2009) HIV-1 RNA dimerization: it takes two to tango. *AIDS Rev.* 11, 91–102
33. Shen, W. *et al.* (2009) A recombination hot spot in HIV-1 contains guanosine runs that can form a G-quartet structure and promote strand transfer *in vitro*. *J. Biol. Chem.* 284, 33883–33893
34. Sundquist, W.I. and Heaphy, S. (1993) Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. *Proc. Natl. Acad. Sci. U. S. A.* 90, 3393–3397
35. Piekna-Przybylska, D. *et al.* (2013) Mechanism of HIV-1 RNA dimerization in the central region of the genome and significance for viral evolution. *J. Biol. Chem.* 288, 24140–24150
36. Piekna-Przybylska, D. *et al.* (2014) U3 region in the HIV-1 genome adopts a G-quadruplex structure in its RNA and DNA sequence. *Biochemistry* 53, 2581–2593
37. Rawal, P. *et al.* (2006) Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res.* 16, 644–655
38. Capra, J.A. *et al.* (2010) G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* 6, e1000861
39. Guo, J.U. and Bartel, D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* Published online September 23, 2016. http://dx.doi.org/10.1126/science.aaf5371
40. Biswas, B. *et al.* (2017) A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B. *Nucleic Acids Res.* 45, 11268–11280
41. Perrone, R. *et al.* (2017) Conserved presence of G-quadruplex forming sequences in the long terminal repeat promoter of lentiviruses. *Sci. Rep.* Published online May 17, 2017. http://dx.doi.org/10.1038/s41598-017-02291-1
42. Fleming, A.M. *et al.* (2016) Zika virus genomic RNA possesses conserved G-quadruplexes characteristic of the Flaviviridae family. *ACS Infect. Dis.* 2, 674–681
43. Perrone, R. *et al.* (2013) A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J. Med. Chem.* 56, 6521–6530
44. Biswas, B. *et al.* (2016) Genome-wide analysis of G-quadruplexes in herpes virus genomes. *BMC Genomics* 17, 949
45. Rochette, P.A. *et al.* (2015) Mutation of UL24 impedes the dissemination of acute herpes simplex virus 1 infection from the cornea to neurons of trigeminal ganglia. *J. Gen. Virol.* 96, 2794–2805
46. Jacobson, J.G. *et al.* (1998) Importance of the herpes simplex virus UL24 gene for productive ganglionic infection in mice. *Virology* 242, 161–169
47. Brinkmann, M.M. *et al.* (2007) Modulation of host gene expression by the K15 protein of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* 81, 42–58
48. Pyles, R.B. and Thompson, R.L. (1994) Evidence that the herpes simplex virus type 1 uracil DNA glycosylase is required for efficient viral replication and latency in the murine nervous system. *J. Virol.* 68, 4963–4972
49. Biswas, B. *et al.* (2018) Pac1 signals of human herpesviruses contain a highly conserved G-quadruplex motif. *ACS Infect. Dis.* 4, 744–751
50. Adelman, K. *et al.* (2001) Herpes simplex virus DNA packaging sequences adopt novel structures that are specifically recognized by a component of the cleavage and packaging machinery. *Proc. Natl. Acad. Sci. U. S. A.* 98, 3086–3091
51. Tluckova, K. *et al.* (2013) Human papillomavirus G-quadruplexes. *Biochemistry* 52, 7207–7216
52. Zahin, M. *et al.* (2018) Identification of G-quadruplex forming sequences in three manatee papillomaviruses. *PLoS One* 13, e0195625

53. Bernard, H. (2013) Regulatory elements in the viral genome. *Virology* 445, 197–204
54. Graham, S.V. (2010) Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. *Future Microbiol.* 5, 1493–1506
55. Tan, J. *et al.* (2007) The 'SARS-unique domain' (SUD) of SARS coronavirus is an oligo(G)-binding protein. *Biochem. Biophys. Res. Commun.* 364, 877–882
56. Tan, J. *et al.* (2009) The SARS-unique domain (SUD) of SARS coronavirus contains two macrodomains that bind G-quadruplexes. *PLoS Pathog.* 5, e1000428
57. Kusov, Y. *et al.* (2015) A G-quadruplex-binding macrodomain within the 'SARS-unique domain' is essential for the activity of the SARS-coronavirus replication-transcription complex. *Virology* 484, 313–322
58. Cary, D.C. *et al.* (2016) Molecular mechanisms of HIV latency. *J. Clin. Invest.* 126, 448–454
59. Nicoll, M.P. *et al.* (2012) The molecular basis of herpes simplex virus latency. *FEMS Microbiol. Rev.* 36, 684–705
60. Madireddy, A. *et al.* (2016) G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Res.* 44, 3675–3694
61. Gilbert-Girard, S. *et al.* (2017) Stabilization of telomere G-quadruplexes interferes with human herpes virus 6A chromosomal integration. *J. Virol.* 91, e00402-17
62. Kang, M.S. and Kieff, E. (2015) Epstein–Barr virus latent genes. *Exp. Mol. Med.* 47, e131
63. Murat, P. *et al.* (2014) G-quadruplexes regulate Epstein–Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.* 10, 358–364
64. Schulz, T.F. and Cordes, S. (2009) Is the Epstein–Barr virus EBNA-1 protein an oncogen? *Proc. Natl. Acad. Sci. U. S. A.* 106, 2091–2092
65. Wood, V.H. *et al.* (2007) Epstein–Barr virus-encoded EBNA1 regulates cellular gene transcription and modulates the STAT1 and TGFbeta signaling pathways. *Oncogene* 26, 4135–4147
66. Lista, M.J. *et al.* (2017) Nucleolin directly mediates Epstein–Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat. Commun.* 8, 16043
67. Norseen, J. *et al.* (2009) Role for G-quadruplex RNA binding by Epstein–Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment. *J. Virol.* 83, 10336–10346
68. Tosoni, E. *et al.* (2015) Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res.* 43, 8884–8897
69. Scalabrin, M. *et al.* (2017) The cellular protein hnRNP A2/B1 enhances HIV-1 transcription by unfolding LTR promoter G-quadruplexes. *Sci. Rep.* 7, 45244
70. Michael, C.A. *et al.* (2014) The antimicrobial resistance crisis: causes, consequences, and management. *Front. Public Health* 2, 145
71. Perrone, R. *et al.* (2013) Formation of a unique cluster of G-quadruplex structures in the HIV-1 Nef coding region: implications for antiviral activity. *PLoS One* 8, e73121
72. Basmaciogullari, S. and Pizzato, M. (2014) The activity of Nef on HIV-1 infectivity. *Front. Microbiol.* 5, 232
73. Garcia, J.V. and Miller, A.D. (1991) Serine phosphorylation-independent downregulation of cell-surface CD4 by *nef*. *Nature* 350, 508–511
74. Rhodes, D. *et al.* (2000) Characterization of three nef-defective human immunodeficiency virus type 1 strains associated with long-term nonprogression. Australian Long-Term Nonprogressor Study Group. *J. Virol.* 74, 10581–10588
75. Miller, M.D. *et al.* (1994) The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages. *J. Exp. Med.* 179, 101–113
76. Wang, S.R. *et al.* (2016) A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti–hepatitis C target. *Sci. Adv.* 2, e1501535
77. Wang, S.R. *et al.* (2016) Chemical targeting of a G-quadruplex RNA in the Ebola virus L gene. *Cell Chem. Biol.* 23, 1113–1122
78. Artusi, S. *et al.* (2015) The herpes simplex virus-1 genome contains multiple clusters of repeated G-quadruplex: implications for the antiviral activity of a G-quadruplex ligand. *Antiviral Res.* 118, 123–131
79. Callegaro, S. *et al.* (2017) A core extended naphtalene diimide G-quadruplex ligand potently inhibits herpes simplex virus 1 replication. *Sci. Rep.* 7, 2341
80. Perrone, R. *et al.* (2014) Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. *J. Antimicrob. Chemother.* 69, 3248–3258
81. Perrone, R. *et al.* (2015) Synthesis, binding and antiviral properties of potent core-extended naphthalene diimides targeting the HIV-1 long terminal repeat promoter G-quadruplexes. *J. Med. Chem.* 58, 9639–9652
82. Perrone, R. *et al.* (2017) Mapping and characterization of G-quadruplexes in *Mycobacterium tuberculosis* gene promoter regions. *Sci. Rep.* 7, 5743
83. Harris, L.M. and Monsell, K.R. *et al.* (2018) G-quadruplex DNA motifs in the malaria parasite *Plasmodium falciparum* and their potential as novel antimalarial drug targets. *Antimicrob. Agents Chemother.* 62 (3),
84. Lin, C.L. and Kao, J.H. (2015) Hepatitis B virus genotypes and variants. *Cold Spring Harb. Perspect. Med.* 5, a021436
85. Sunbul, M. (2014) Hepatitis B virus genotypes: global distribution and clinical importance. *World J. Gastroenterol.* 20, 5427–5434
86. Beaume, N. *et al.* (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
87. Kota, S. *et al.* (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
88. Luque-Almagro, V.M. *et al.* (2017) Transcriptional and translational adaptation to aerobic nitrate anabolism in the denitrifier *Paracoccus denitrificans*. *Biochem. J.* 474, 1769–1787
89. Waller, Z.A. *et al.* (2016) Control of bacterial nitrate assimilation by stabilization of G-quadruplex DNA. *Chem. Commun. (Camb)* 52, 13511–13514
90. Hershman, S.G. *et al.* (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 36, 144–156
91. Benne, R. (1994) RNA editing in trypanosomes. *Eur. J. Biochem.* 221, 9–23
92. Leeder, W.M. *et al.* (2016) The RNA chaperone activity of the *Trypanosoma brucei* editosome raises the dynamic of bound pre-mRNAs. *Sci. Rep.* 6, 19309
93. Leeder, W.M. *et al.* (2016) Multiple G-quartet structures in pre-edited mRNAs suggest evolutionary driving force for RNA editing in trypanosomes. *Sci. Rep.* 6, 29810
94. Bhartiya, D. *et al.* (2016) Genome-wide regulatory dynamics of G-quadruplexes in human malaria parasite *Plasmodium falciparum*. *Genomics* 108, 224–231
95. White, R.E. *et al.* (2007) Mutagenesis of the herpesvirus saimiri terminal repeat region reveals important elements for virus production. *J. Virol.* 81, 6765–6770
96. Dheekollu, J. *et al.* (2013) Timeless-dependent DNA replication-coupled recombination promotes Kaposi's sarcoma-associated herpesvirus episome maintenance and terminal repeat stability. *J. Virol.* 87, 3699–3709
97. Rehm, C. *et al.* (2015) Investigation of a quadruplex-forming repeat sequence highly enriched in *Xanthomonas* and *Nostoc* sp. *PLoS One* 10, e0144275

98. Glazko, V.I. and Kosovsky, G. Yu (2013) Structure of genes coding the envelope proteins of the avian influenza a virus and bovine leucosis virus. *Russian Agric. Sci.* 39, 511–515

99. Day, H.A. *et al.* (2014) i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* 22, 4407–4418

100. Zerrati, M. *et al.* (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* 10, 631–637