11268–11280 *Nucleic Acids Research, 2017, Vol. 45, No. 19* *Published online 13 September 2017*
doi: 10.1093/nar/gkx823

A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B

Banhi Biswas, Manish Kandpal and Perumal Vivekanandan*

Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi 110016, India

Received March 26, 2017; Revised August 31, 2017; Editorial Decision September 04, 2017; Accepted September 07, 2017

ABSTRACT

HBV genotypes differ in pathogenicity. In addition, genotype-specific differences in the regulation of transcription and virus replication exist in HBV, but the underlying mechanisms are unknown. Here, we show the presence of a G-quadruplex motif in the promoter of the preS2/S gene; this G-quadruplex is highly conserved only in HBV genotype B but not in other HBV genotypes. We demonstrate that this G-quadruplex motif forms a hybrid intramolecular G-quadruplex structure. Interestingly, mutations disrupting the G-quadruplex in HBV genotype B reduced the preS2/S promoter activity, leading to reduced hepatitis B surface antigen (HBsAg) levels. G-quadruplex ligands stabilized the G-quadruplex in genotype B and enhanced the preS2/S promoter activity. Furthermore, mutations disrupting the G-quadruplex in the full-length HBV genotype B constructs were associated with impaired virion secretion. In contrast to typical G-quadruplexes within promoters which are negative regulators of transcription the G-quadruplex in the preS2/S promoter of HBV represents an unconventional positive regulatory element. Our findings highlight (a) G-quadruplex mediated enhancement of transcription and virion secretion in HBV and (b) a yet unknown role for DNA secondary structures in complex genotype-specific regulatory mechanisms in virus genomes.

INTRODUCTION

Hepatitis B virus (HBV) is a small circular DNA virus with a genome length of 3.2 kb. HBV has four overlapping open reading frames that encode for seven proteins. HBV is classified into 10 genotypes (A–J) based on nucleotide variation of 8% or more (1). Differences in geographical distribution, pathogenicity and clinical outcome are well-recognized among HBV genotypes. In addition, differences in replication efficiency and virion secretion are known to exist among HBV genotypes (2).

G-quadruplexes are non-canonical DNA secondary structures that are widely recognized as important regulators of gene expression. G-quadruplexes are formed by sequences containing four stretches of G residues which are separated by any nucleotide residue(s). At least two Gs are present in a stretch which interact with the other G stretches to form G-tetrads and each stretch of Gs is typically separated by one to seven nucleotide residues which form the loops (3). Previously, G-quadruplexes were thought to be primarily located in the telomeres but studies in the last decade have highlighted that majority of these secondary DNA structures lie outside the telomeric regions (4). Extensive genome-wide studies and *in vitro* functional analysis have shown regulatory roles for G-quadruplexes (5–9). While G-quadruplex-mediated increase in promoter activity has been reported (10–12), in general, G-quadruplexes in the promoter region are repressors of gene expression (6,7,13).

Virus genomes of HIV-1 (14), HPV (15), SV40 (16), EBV (17), KSHV (18), HSV-1 (19), HCV (20) and Zika virus (21) have been shown to possess G-quadruplex structures. Regulatory roles for G-quadruplexes have been demonstrated for HIV-1, EBV, KSHV, HSV-1 and HCV (14,17–20,22). Studies on virus genomes have demonstrated (a) transcriptional repression by G-quadruplexes in HIV-1 and herpesvirus promoters (14,23) (b) translational repression by a G-quadruplex in HCV and EBV mRNA (17,20) and (c) regulation of latency and episomal persistence by G-quadruplexes in the KSHV genome (18). The functional significance of disrupting G-quadruplexes in complete virus genomes and its implications on the biology of the virus remains poorly studied.

Efficient replication of HBV is dependent on precise transcription of HBV RNAs. Each of the four HBV transcripts (3.5 kb, 2.4 kb, 2.1 kb and 0.7 kb) is regulated by its own promoter. HBV promoters may or may not have a classical TATA box motif required for initiation of transcription (24). The existence of genotype-specific transcriptional regulation has been demonstrated across HBV geno-

*To whom correspondence should be addressed. Tel: +91 11 26597532; Fax: +91 11 26582037; Email: vperumal@bioschool.iitd.ac.in

types (25,26), although the underlying mechanisms remain unknown. The regulatory regions of HBV promoters are known to mimic host gene promoters (27). G-quadruplexes are well-established transcriptional regulators in the human genome. We therefore sought to investigate G-quadruplex motifs in the HBV genome.

We mapped a highly conserved G-quadruplex to the preS2/S promoter in HBV genotype B. The small size of the HBV genome and the absence of other three tetrad G-quadruplexes in the HBV genome provided us with a unique opportunity to investigate the role of this DNA secondary structure at the whole genome level. We found that this G-quadruplex regulates HBV surface antigen (HBsAg) levels, thus affecting virion secretion. To the best of our knowledge, this is the first study to (a) demonstrate the role of a DNA secondary structure in HBV replication (b) investigate the role of a DNA G-quadruplex with mutagenesis at the whole-genome level and (c) demonstrate genotype-specific mechanisms in the regulation of HBV transcripts.

MATERIALS AND METHODS

Analysis of full-length HBV sequences for PQS

All available full-length HBV sequences (n = 5472) were retrieved from the HBV database (https://hbvdb.ibcp.fr) in November 2016 and were analysed for the presence of putative quadruplex sequences (PQS). This includes 781 genotype A, 1449 genotype B, 1829 genotype C, 873 genotype D, 250 genotype E, 226 genotype F, 38 genotype G and 26 genotype H full-length sequences. The availability of very few ($\leq$ 2) full-length HBV sequences for genotypes I and J preclude meaningful analysis and hence they were not included for analysis. Accession numbers of all sequences studied are given in Supplementary File S1. Putative quadruplex sequences (PQS) were mapped in the full-length HBV genomes using Quadparser, a widely used C program (28). The following parameters were used for mapping PQS: (a) no. of G-tetrads = 3 (b) no. of G-stretches = 4 and (c) loop length = 1–7 nucleotides.

PQS conservation analysis

A total of 7899 partial HBV sequences (genotype A: 1349; genotype B: 2275; genotype C: 2361; genotype D: 1249; genotype E: 308; genotype F: 246; genotype G: 80; genotype H: 31) from the preS1 CDS region that overlaps the preS2/S promoter and contains the PQS motif (identified in the full-length sequences) were analysed to assess conservation. The sequences were downloaded from the HBV database (https://hbvdb.ibcp.fr/HBVdb/). Accession numbers of all sequences studied are given in Supplementary File S2. A PQS was considered to be conserved in a sequence when all the G's involved in forming the G-tetrad of the G-quadruplex were intact. Although variations in the loop region may alter the stability and conformation of the G-quadruplex (29–31), these variations do not disrupt the G-quadruplex formation (21). Therefore, variation within the loop sequences were not taken into account for conservation analysis.

CD spectroscopy and ligand binding studies

Synthetic oligonucleotides were purchased from Integrated DNA technologies (IDT). For CD spectroscopy 10 μM of oligonucleotides were added to a buffer containing 10 mM sodium cacodylate and 100 mM KCl. The samples were heated to 95°C and slowly cooled to room temperature before measuring the spectra. Each sample was scanned at a wavelength range of 220–320 nm with a scanning speed of 50 nm/min and a bandwidth of 0.5 nm at 20°C. CD melting of G-quadruplex forming oligonucleotides was performed with and without the G-quadruplex binding ligands BRACO19 and pyridostatin (PDS). The ellipticity change was observed at a fixed wavelength with a temperature ramp rate of 0.5°C/min. Tm was calculated using first derivative method. The sequence of the oligonucleotides used are listed in Supplementary File S3, Table S1.

NMR spectroscopy

Standard ^{1}H NMR was performed using Bruker *Avance* III spectrometer at 500 MHz field strength. Oligonucleotides at a concentration of 0.3 mM were prepared in 20 mM phosphate buffer (pH 7) containing 100 mM KCl and 10% D_2O. Samples were heated to 95°C and slowly cooled down to room temperature before measuring the spectra. Spectra were taken at a temperature of 20°C.

Native gel electrophoresis

Oligonucleotides were prepared at a concentration of 1 μM in Tris EDTA buffer containing 100 mM KCl. The samples were heated to 95°C and slowly cooled down to room temperature for loading. Native polyacrylamide gels were prepared in 1× Tris borate EDTA (TBE) buffer. Gels were run in 0.5× TBE containing 50 mM KCl. Denaturing polyacrylamide gels were prepared using 7 M urea as a denaturant.

DMS footprinting

Two 5′ FAM labeled oligonucleotides (FAM-N and FAM-T) were purchased from Integrated DNA Technologies (IDT). FAM-N (5′-FAM-CCAACAAGGTGGGAGTG GGAGCATTCGGGCCAGGG 3′) had an additional 10 bases of the native HBV sequence upstream of the first guanine residue of the PQS and FAM-T (5′-FAM-TTTTTTT TTTGGGAGTGGGAGCATTCGGGCCAGGG 3′) had an extra stretch of ten Ts upstream of the first guanine residue of the PQS. The additional nucleotides were added to lower steric hindrances in formation of a G-quadruplex structure due to FAM labeling. The oligonucleotides prepared in Tris–EDTA (10 mM) with and without KCl were heated to 95°C and slowly cooled down to room temperature allowing formation of G-quadruplexes. Thereafter the oligonucleotides were treated with 0.5% DMS for 2 min and the reactions were stopped by adding 1 μg of calf thymus DNA. The reactions were run in a 20% native polyacrylamide gel, extracted and ethanol precipitated. Reactions were cleaved using 1 M piperidine at 95°C. Cleavage products were dried using speed vac, re-suspended in alkaline sequencing dye and were resolved in a 20% denaturing polyacrylamide gel.

Creation of luciferase constructs

The preS2/S promoter region (–241nt to +35nt with respect to the transcription start site) of HBV genotype B (GTB) containing the G-quadruplex was cloned into a promoterless vector (pGL3 basic, Promega) upstream of the firefly luciferase CDS using primers listed in Supplementary File S3, Table S2. Two point mutations which disrupt the formation of quadruplex structure were introduced in the putative quadruplex sequence of GTB using appropriate primers (Supplementary File S3 Table S2). The presence of the two point mutations was confirmed by nucleotide sequencing. The clone with the wild-type promoter sequence from GTB is referred to as the GTB-luc-WT and the clone with the mutated promoter sequence from GTB is referred to as GTB-luc-MUT.

Creation of HBV surface gene constructs

The 2.1 kb surface mRNA region along with its promoter was amplified using appropriate primers (Supplementary File S3, Table S2) from GTB and cloned in the pGL3 basic vector. Subsequently, the two points (as described for the luciferase constructs) disrupting the G-quadruplex in the preS2/S promoter were introduced. The wild-type GTB and the mutated GTB surface gene constructs are referred to GTB-surface-WT and GTB-surface-MUT respectively.

Creation of wild-type and mutant full-length HBV genome constructs

The full-length HBV genotype B construct (GTB-full-WT) was a kind gift from Prof. Jia-Horng Kao, Graduate Institute of Clinical Medicine, National Taiwan University College of Medicine. The two point mutations were introduced within the PQS in the preS2/S promoter to disrupt the formation of the G-quadruplex structure as done for the promoter region using appropriate primers (Supplementary File S3, Table S2). The full-length wild-type HBV clone (GTB-full-WT) and the respective mutant (GTB-full-MUT) had the same sequence barring the two point mutations introduced in the mutant. The presence of the desired mutations was confirmed by nucleotide sequencing (Supplementary File S3).

Cell culture and transfection

Huh7 cells were maintained in DMEM supplemented with 10% fetal bovine serum and were incubated at 37°C and with 5% CO_2. Wild-type and mutated luciferase reporter constructs were quantitated and equal quantities (mass) were transfected. Cells were harvested for dual reporter luciferase assay after 24 h. For transfection of surface gene constructs, wild-type and mutated plasmids were quantitated and equals quantities (mass) were transfected. Supernatants were collected and cells were harvested after 24 h of transfection. Wild-type and mutated full-length genome constructs were digested with *BspQ1* and gel extracted using Strataprep DNA gel extraction kit. The gel extracted DNA was quantified using real time PCR of the surface gene to transfect equal copy numbers of wild-type and mutated full-length constructs. Cells and supernatants were harvested 3 days after transfection. Transfections were done using lipofectamime 2000 according to manufacturers' protocol.

Quantitation of viral RNAs and cccDNA

Total RNA was extracted using TRIzol reagent (ThermoFisher scientific) according to the manufacturer's protocol. cDNA was synthesized from 1 μg of DNase I treated RNA using iScript cDNA synthesis kit (Biorad). Real time PCR was done using the Faststart essential DNA green master (Roche) for quantifying viral pre-genomic RNA and pre-core RNA using appropriate primers (32). HBV cccDNA was quantified using premix Ex taq master mix (Takara) with a taqman probe and appropriate primers. The list of primers is given in Supplementary File S3, Table S2. Glyceraldehyde 3-phosphate dehydrogenase, a house keeping gene was quantified to serve as an internal control. Standard curves were generated for all real time PCRs to allow absolute quantitation of the target.

Estimation of secreted and intracellular proteins

Intracellular and extracellular hepatitis B virus surface antigen (HBsAg) and secreted hepatitis B virus e antigen (HBeAg) were estimated using commercially available ELISA kits (MONOLISA, Biorad for HBsAg and Dia.Pro Diagnostic Bioprobes for HBeAg). Supernatants and cell lysates were diluted appropriately to ensure absorbance values are at a linear range. The intracellular core antigen levels were assessed using a commercially available ELISA (Cell Biolabs, Inc.). Total intracellular protein content was determined using bicinchoninic acid (BCA) assay kit (ThermoFisher scientific). The intracellular core protein and surface protein levels were normalised to the total protein content.

Luciferase assay

The luciferase reporter constructs (GTB-luc-WT and GTB-luc-MUT) were co-transfected with pRL-TK into Huh7 cells in 20:1 ratio. The pRL-TK, which has a thymidine kinase promoter upstream of the Renilla luciferase coding DNA sequence (CDS) serves as an internal control. For studies with G-quadruplex binding ligands, the ligands were added 2 h after transfection. Cells were harvested 24 h after transfection for luminescence measurement. The luminescence was measured using the dual luciferase reporter assay system (Promega) according to manufacturer's protocol.

Northern blotting

Total cellular RNA was extracted using the TRIzol reagent (ThermoFisher Scientific). Ten micrograms of total RNA was separated using 1.2% of formaldehyde agarose gels and was transferred onto nylon membrane (Roche) overnight. After transfer, RNA was crosslinked to the membrane by baking it at 70°C for 2 h. A DIG-labelled probe corresponding to the HBX ORF was synthesized using DIG Northern starter kit (Roche). The membrane was hybridized with the DIG-labelled probe at 68°C. Actin RNA was assayed similarly and it served as an internal control. Colorimetric

detection was used for visualization of RNA on the membrane. Densitometry analysis of the blot was done using ImajeJ software.

Estimation of secreted virion

Virions secreted in the supernatant were quantitated using an immunocapture-based protocol standardized in our laboratory (33). Briefly, 150 μl of supernatant from cells transfected with GTB-full-WT and GTB-full-MUT were added to HBsAg ELISA plates (MONOLISA, Bio-rad) for immunocapture of hepatitis B virions. DNA was extracted from the captured virions using QIAamp DNA Mini Kit (Qiagen) and virion-associated DNA was quantitated using real time PCR.

Statistical analysis

Data was analysed using the Student's *t*-test. *P* values <0.05 were considered significant.

RESULTS

PQS mapping and conservation in the HBV genome

We mapped PQS in the full-length genomic sequences of HBV ($n = 5472$). We found a single PQS in 92% of HBV genotype B sequences (Figure 1A); this G-quadruplex was present in a very small proportion of sequences from genotypes A (1.2%) and C (0.76%). Other HBV genotypes had no PQS in their genomes (Figure 1A). Since the PQS was present in majority (92%) of all full-length genotype B sequences studied, we sought to investigate the potential role of this DNA secondary structure. This PQS motif is located in the presS2/S promoter (Figure 1B) of HBV genotype B and is 190 bp upstream of the preS2/S transcription start site (Figure 1C). We then analysed the conservation of G residues in this G-quadruplex (i.e. the four stretches of G residues) in all available sequences of the preS2/S promoter ($n = 7899$). Interestingly, the average percentage conservation of each G residue in the G-quadruplex motif in the presS2/S promoter of HBV genotype B was 99.4%; all other genotypes were short of one or more Gs to form the three tetrad quadruplex structure (Figure 1D). Intrigued by the presence and conservation of a PQS motif in the promoter region of the preS2/S gene of HBV genotype B, we went to study its role in the biology of HBV.

The PQS motif present in the preS2/S promoter forms a G-quadruplex structure

CD spectroscopy is widely used for confirming the formation of G-quadruplex structures (34). A CD spectra with a positive peak at around 260 nm and a negative peak at around 240 nm is indicative of formation of parallel G-quadruplex. Whereas, a positive peak around 290 nm and a negative peak around 260 nm indicates the formation of an anti-parallel G-quadruplex. A positive peak at 260 and 290 nm along with a negative peak at 240 nm suggests hybrid type of G-quadruplex formation. The oligonucleotide of the PQS (GTB-WT-PQS) mapped at the preS2/S promoter region of the HBV genome formed a hybrid G-quadruplex

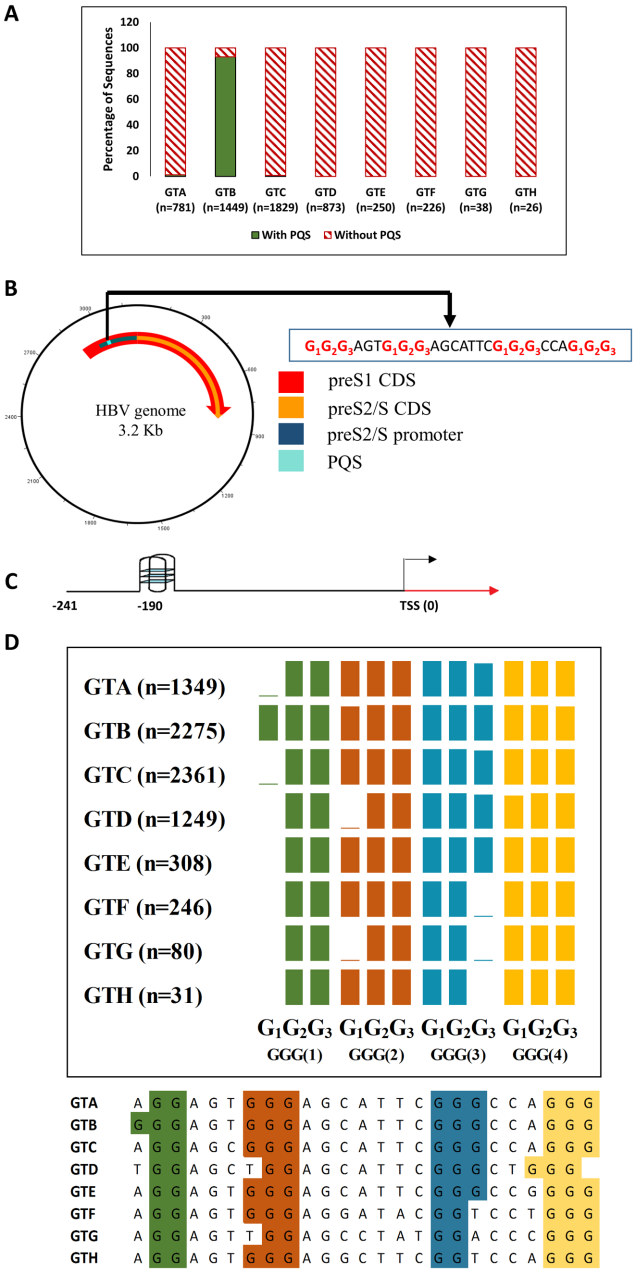

Figure 1. (**A**) Bar graph represents percentage of full-length sequences with and without a putative quadruplex sequence (PQS) among HBV genotypes (GTs). The number of full length sequences analysed is given in parentheses. (**B**) Schematic representation of the position of the putative G-quadruplex motif (blue) in the promoter region of pres2/S gene. The red line with an arrow head represents preS1 CDS and the yellow line within the red line with an arrow head represent preS2/S CDS. All lines are drawn to scale using DNA plotter. (**C**) A schematic showing the position of the G-quadruplex motif with respect to the transcription start site (TSS). (**D**) Schematic representation showing conservation of Gs which are involved in the formation of the G-tetrads within the G-quadruplex structure. Each coloured box represents conservation of a G residue in the PQS sequence within each genotype (GT). The height of the box represents percentage of sequences having a G residue at that particular position (in aligned sequences). For example, if the height is half of the maximum height, 50% of the sequences have a G residue at that position. Representative sequences of the PQS region from each genotype are given below. The number of sequences analysed is given in parentheses.

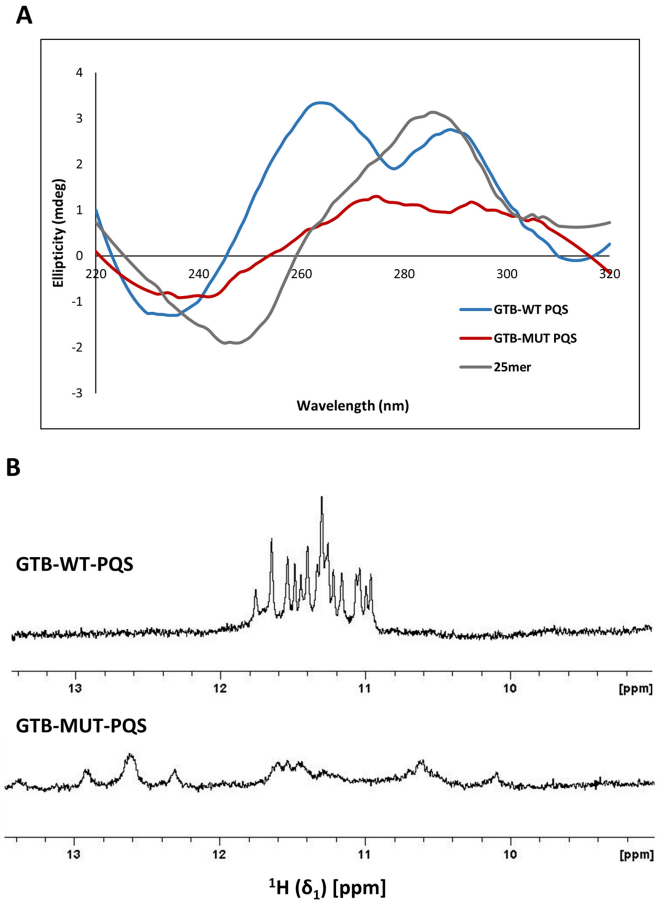

Figure 2. (**A**) CD spectra showing the formation of a G-quadruplex structure in genotype B wild-type PQS (GTB-WT-PQS) oligonucleotide. The mutated oligonucleotide (GTB-MUT-PQS) does not form a G-quadruplex structure. A ssDNA control (25mer) incapable of forming a G-quadruplex is taken as a negative control. (**B**) ^{1}H NMR spectra of genotype B wild type (GTB-WT-PQS) and mutated (GTB-MUT-PQS) PQS oligonucleotide. Please see Supplementary File S3 Table S1 for sequence of the oligonucleotides.

(Figure 2A) which can be either a mixture of parallel or anti-parallel folded species in a solution or a single hybrid species with three parallel and one anti-parallel strand. We then mutated the central G in the second and third stretches of Gs of the PQS motif to disrupt the formation of a G-quadruplex. As expected this mutated PQS oligonucleotide (GTB-MUT-PQS) did not show CD spectral signatures corresponding to G-quadruplex formation. The positive peak in the CD spectra shifted from 265 to 276 nm (Figure 2A). CD spectroscopy of PQS oligonucleotides from other genotypes demonstrated formation of G-quadruplex structure in genotype A and genotype E (Supplementary File S3, Figure S1). Although PQS oligonucleotides from genotype A and E showed CD spectral features corresponding to G-quadruplex structure formation, their melting temperatures as demonstrated by CD melt curves were considerably lower as compared to PQS oligonucleotide from genotype B (Supplemenatary File S3, Figure S2, Table S3).

In addition to CD spectroscopy, we used NMR spectroscopy to confirm if the predicted PQS forms a G-quadruplex structure. In ^{1}H NMR, signals pertaining to the formation of Hoogsteen base pairing appear in the chemical shift (δ) between 10.6 and 12 ppm; whereas signals for Watson-Crick pairing appear in the chemical shift between 12.5 and 13.7 ppm. As expected, peaks in the range assigned for Hoogsteen base paring was seen for the wild-type PQS oligonucleotide (GTB-WT-PQS) but not for the mutated PQS oligonucleotide (GTB-MUT-PQS) (Figure 2B). Both CD spectroscopy and NMR spectroscopy confirmed formation of G-quadruplex structure in the wild-type PQS oligonucleotide.

G-quadruplex structures can be both intramolecular and intermolecular in nature depending on the stoichiometry of the strands involved in forming the structure. We believe that functional G-quadruplexes in the promoter region are likely to be intramolecular in nature. To determine the molecularity of the PQS motif present in the preS2/S promoter of HBV genotype B we did native polyacrylamide gel electrophoresis (PAGE). Intramolecular G-quadruplexes are known to migrate faster owing to their compact structure (35–37). The wild-type PQS oligonucleotide (GTB-WT-PQS; 25 bases) migrated faster than the 25mer ssDNA oligonucleotide control (38) (Figure 3A), suggesting that the wild-type PQS oligonucleotide forms an intramolecular G-quadruplex structure within the preS2/S promoter of HBV genotype B. The migration of the mutated PQS oligonucleotide (GTB-MUT-PQS; 25 bases) was comparable to that of the 25mer ssDNA oligonucleotide control in the native PAGE (Figure 3A), suggesting that mutations disrupted the formation of an intramolecular G-quadruplex. When the oligonucleotides were run in a denaturing polyacrylamide gel containing 7 M urea, all the three oligonucleotides migrated with the same pace as secondary structures were denatured (Figure 3B).

DMS footprinting helps in fine mapping of the guanine (G) residues involved in formation of the G-quadruplex structure. G residues involved in formation of a G-quadruplex structure are inaccessible for methylation by DMS and are hence not cleaved by piperidine. DMS footprinting analysis of both FAM-N and FAM-T oligonucleotides indicate the formation of intramolecular G-quadruplex structures. All the four stretches of G-residues speculated to be involved in intramolecular G-quadruplex formation were protected from methylation whereas G-residues in the loop regions were sensitive to DMS methylation (Figure 3C and D). The formation of intramolecular G-quadruplex in the FAM-N oligonucleotide was more evident at a higher KCl concentration (600mM) than at a lower KCl concentration (100mM) (Figure 3C). The formation of intramolecular G-quadruplex in the FAM-T oligonucleotide was comparable at both the KCl concentrations (Figure 3D). At 600 mM, KCl concentration more number of molecules of the FAM-N oligonucleotide formed a G-quadruplex structure compared to that in 100 mM KCl as demonstrated by CD spectroscopy (data not shown). This analysis convincingly showed formation of an intramolecular G-quadruplex structure. Based on CD spectroscopy, NMR spectroscopy, native PAGE and DMS footprinting analysis we conclude that the PQS from the promoter re-

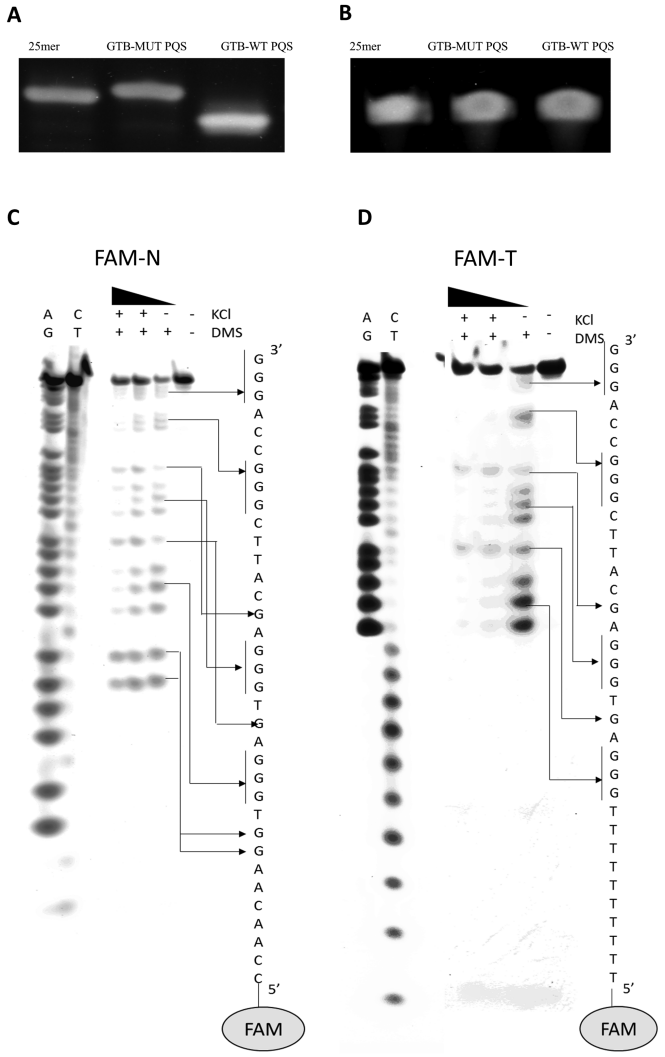

Figure 3. (**A**) Native polyacrylamide gel electrophoresis suggests the formation of intramolecular G-quadruplex structure as indicated by faster migration of genotype B wild-type PQS (GTB-WT-PQS) oligonucleotide compared to the mutated (GTB-MUT-PQS) oligonucleotide and the ss-DNA control (25mer). (**B**) Denaturing polyacrylamide gel electrophoresis showing similar migration rate for the wild-type (GTB-WT-PQS) and the mutated (GTB-MUT-PQS) oligonucleotide in the presence of a denaturant (7 M urea). DMS footprinting of 5′ FAM labelled PQS oligonucleotides namely (**C**) FAM-N (contains 10 additional bases of the native HBV sequence upstream of the the PQS in the preS2/S promoter) and (**D**) FAM-T (contains 10 additional stretch of 10 thymine (T) residues upstream of the PQS in the preS2/S promoter). The KCl concentrations used were 600, 100 and 0 mM (no KCl). AG and TC represents sequencing cleavage reactions specific to purine and pyrimidine. The vertical bars next to the sequence of the oligonucleotide represents the G residues protected from DMS methylation due to formation of G-quadruplex structure. Please see Supplementary File S3, Table S1 for sequence of the oligonucleotides.

gion of preS2/S gene in HBV genotype B forms a hybrid type intramolecular G-quadruplex structure.

The G-quadruplex motif in the preS2/S promoter enhances transcription in luciferase reporter constructs

We then investigated the role of the G-quadruplex within the preS2/S promoter of HBV genotype B using luciferase reporter constructs. Briefly, the preS2/S promoter with the intact G-quadruplex motif from HBV genotype B (wild-type) was cloned upstream of a luciferase gene to measure relative luminescence. We then introduced two mutations in the G-quadruplex motif (the two mutations are identical to that described earlier for GTB-MUT-PQS) to disrupt intramolecular G-quadruplex formation. The wild-type promoter with intact G-quadruplex (GTB-luc-WT) was associated with significantly higher promoter activity as compared to that of the mutant promoter (GTB-luc-MUT) (Figure 4A).

Stabilization of the preS2/S promoter G-quadruplex enhances promoter activity

To extend our investigation on this G-quadruplex motif in the preS2/S promoter, we used G-quadruplex binding ligands. BRACO19 and PDS have been shown to specifically bind and stabilize G-quadruplex structures (39–41). CD melting studies indicate that the GTB-WT-PQS oligonucleotide is stabilized by the G-quadruplex binding ligands BRACO19 and PDS (Figure 4B and C). We then repeated the luciferase reporter assays for GTB-luc-WT and GTB-luc-MUT in the presence of BRACO19 and PDS. Interestingly, the addition of G-quadruplex ligands selectively enhanced the promoter activity of the GTB-luc-WT but not that of the GTB-luc-MUT (Figure 4D and E). The selective enhancement of the GTB-luc-WT promoter activity by the G-quadruplex stabilizing ligands further supports a regulatory role of the G-quadruplex within the genotype B preS2/S promoter. Importantly, this finding vindicates that the differences in the promoter activity between the GTB-luc-WT and the GTB-luc-MUT are linked to differences in DNA secondary structure formation rather than to primary sequence changes. Taken together, these findings convincingly demonstrate that the G-quadruplex motif present in the genotype B preS2/S promoter acts as a positive regulator of transcription.

G-quadruplex modulates transcription of preS2/S gene

To assess the effect of the G-quadruplex motif directly on the transcription of preS2/S gene we cloned the surface gene along with its promoter region into pGL3 basic vector. We made a preS2/S gene construct where the gene is under the regulation of its own promoter sequence and has its own poly A signal. This construct would allow us to determine preS2/S transcript levels using real time PCR. Two point mutations were introduced in the PQS motif as previously stated. The wild-type genotype B surface construct (GTB-surface-WT) was associated with increased preS2/S transcripts as compared to the mutated genotype B surface construct (GTB-surface-MUT) (Figure 5A). The GTB-surface-WT construct was also associated with higher levels of secreted HBsAg than the GTB-surface-MUT construct (Figure 5B).

The G-quadruplex motif selectively regulates transcription of preS2/S gene from full-length HBV constructs

HBV genotype B full-length wild-type (GTB-full-WT) and mutant constructs (GTB-full-MUT) were created to assess

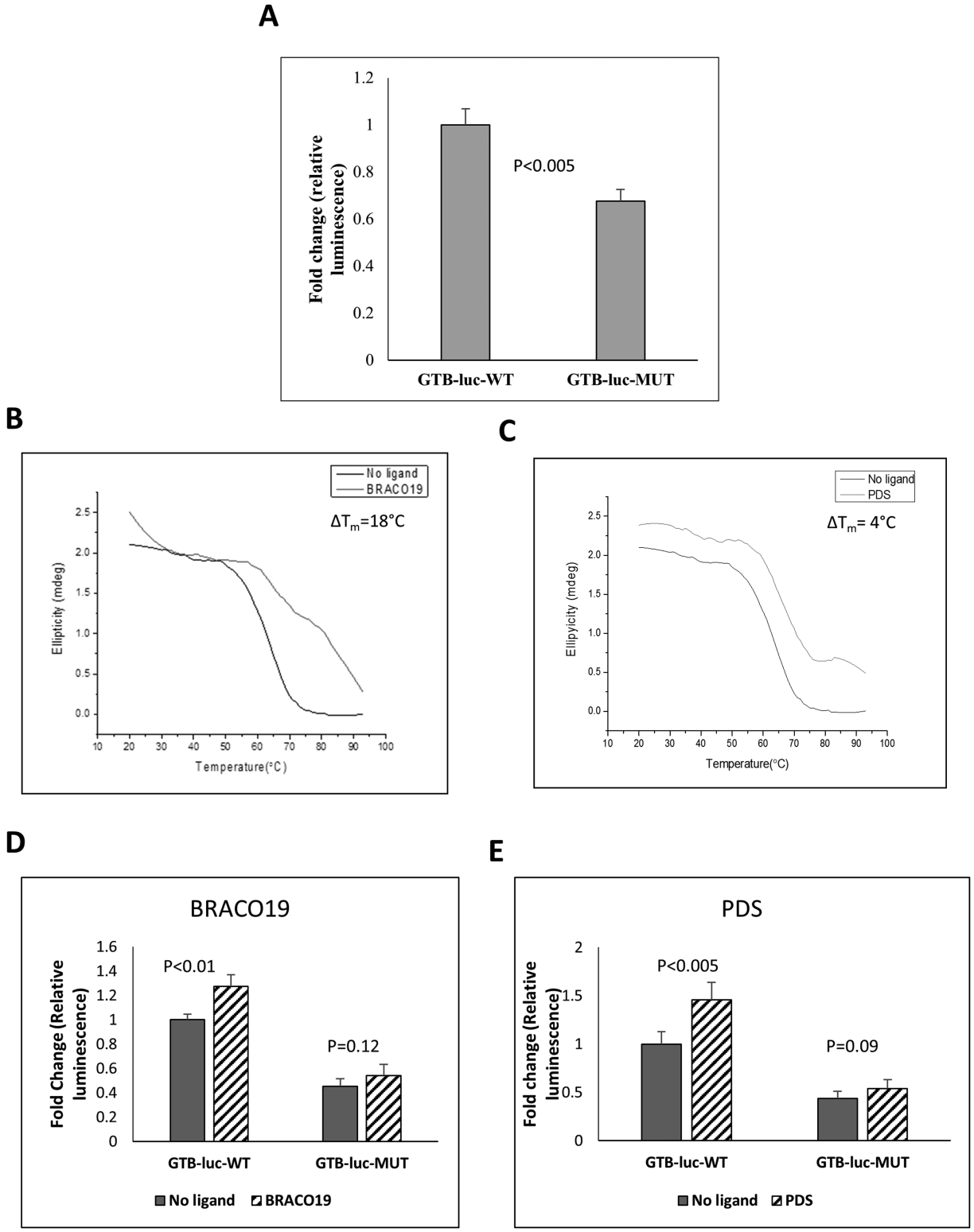

Figure 4. (**A**) Bar graphs indicate increased promoter activity of the genotype B wild-type preS2/S promoter (GTB-luc-WT) as compared to the mutated genotype B preS2/S promoter (GTB-luc-MUT) as measured using luciferase reporter constructs. The luminescence from pRL-TK, which has a thymidine kinase promoter upstream of the Renilla luciferase CDS served as an internal control for normalization. CD melting showing stabilization of the wild-type PQS (GTB-WT-PQS) oligonucleotide in the presence of (**B**) BRACO19 (10 μM) and (**C**) pyridostatin (PDS) (10 μM). $\Delta T_m = T_m$ (with ligand) – T_m (without ligand). Melting was monitored at a wavelength of 292 nm. (**D**) Bar graph showing enhanced promoter activity of the genotype B wild-type preS2/S promoter (GTB-luc-WT) in the presence of BRACO19 as measured using luciferase reporter constructs (**E**) Bar graph showing enhanced promoter activity of the genotype B wild-type preS2/S promoter (GTB-luc-WT) in the presence of 10 μM pyridostatin (PDS) as measured using luciferase reporter constructs. Data are presented as mean ± SD (n = 4 replicates).

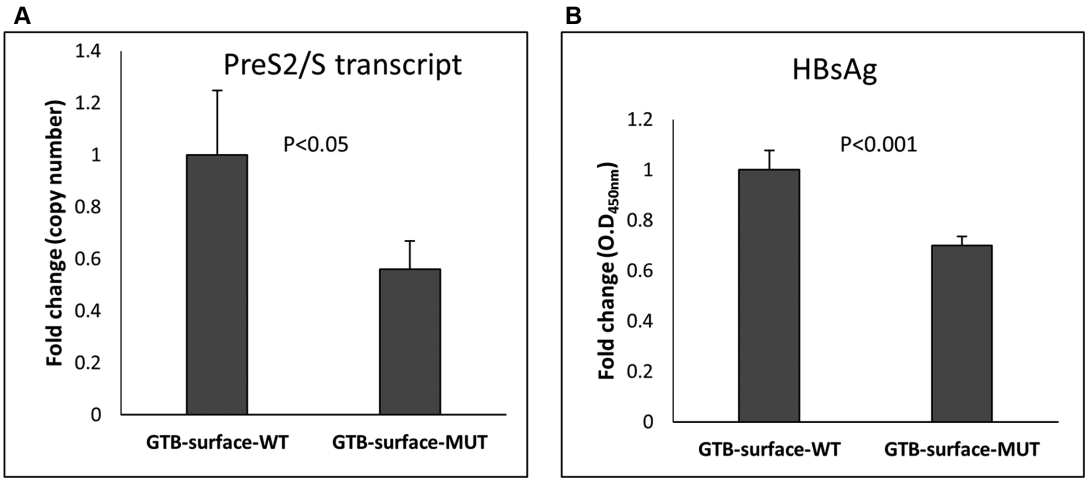

Figure 5. (**A**) Bar graph indicates increased expression of the surface transcripts from the wild-type surface constructs (complete preS2/S gene with G-quadruplex containing preS2/S promoter; GTB-surface-WT) as compared to the mutant surface constructs (complete preS2/S gene with preS2/S promoter with mutations disrupting the G-quadruplex; GTB-surface-MUT). The surface transcripts were quantitated using real-time PCR. (**B**) Bar graph indicates increased expression of the hepatitis B surface antigen (HBsAg) from the wild-type surface constructs (GTB-surface-WT) as compared to the mutant surface constructs (GTB-surface-MUT). The HBsAg levels were measured from the supernatant of cells transfected with GTB-surface-WT or GTB-surface-MUT using ELISA. Data are presented as mean ± SD ($n = 4$ replicates)

the role of this G-quadruplex in the replication of HBV. The HBV cccDNA levels in cells were comparable between GTB-full-WT and GTB-full-MUT (Figure 6A). Northern blot analysis indicates higher levels of preS2/S transcripts in the wild-type construct (GTB-full-WT) compared to the mutant construct (GTB-full-MUT) (Figure 6B and C), while other HBV transcripts including precore/pregenomic RNA and the preS1 transcripts were comparable between the two constructs (Figure 6B and D). In addition to northern blot, comparable levels of pregenomic (pg) RNA (Figure 6E) and precore (pc) RNA (Figure 6F) were observed between the GTB-full-WT and the GTB-full-MUT constructs in RT qPCR analysis. This finding clearly indicates that the G-quadruplex motif in the preS2/S promoter selectively regulates the preS2/S transcript without having an effect on other HBV transcripts.

We then assessed HBV protein levels in the wild-type and mutated full-length constructs. First, we assessed the intracellular and secreted HBsAg levels that correspond to the preS2/S transcript. As expected, the GTB-full-WT construct was associated with significantly higher levels of both intracellular and secreted HBsAg levels (Figure 7A and B). The hepatitis B core protein (HBcAg) levels and the hepatitis B 'e' antigen (HBeAg) levels were comparable between the wild-type (GTB-full-WT) and the mutant constructs (GTB-full-MUT) (Figure 7C and D); this finding confirms that the G-quadruplex in the preS2/S promoter selectively enhances HBsAg levels by increasing the preS2/S transcript.

Disruption of the G-quadruplex motif leads to a significant reduction in virion secretion

It is well known that a significant fraction of the HBV envelope is contributed by the major surface protein (HBsAg) produced from the preS2/S transcript (42). Since the disruption of the G-quadruplex was associated with significant reduction in HBsAg levels, we were intrigued to assess its effect on virion secretion. Interestingly, virion secretion from the cells transfected with the GTB-full-WT was almost 5-fold higher as compared to that from the cells transfected with the GTB-full-MUT (Figure 8). Therefore, the G-quadruplex-mediated increase in preS2/S transcripts leads to higher HBsAg levels, which in turn results in increased virion secretion.

DISCUSSION

In this study, we report the presence of a G-quadruplex motif in the HBV genome and demonstrate a regulatory role for this DNA secondary structure in transcription and virion secretion. Interestingly, we found only a single three tetrad G-quadruplex motif in the entire HBV genome; this motif is highly conserved (92%) in HBV genotype B and is rarely present in genotypes A and C. There are no three tetrad G-quadruplex motifs in any of the full-length sequences analysed from other HBV genotypes. The presence of RNA secondary structures in HBV pregenomic RNA and their role in HBV replication is well known (43). Nonetheless, to the best of our knowledge, the existence of G-quadruplex motifs or DNA secondary structures in the HBV genome has not been reported. The G-quadruplex motif we identified in this study mapped to the preS2/S promoter of the HBV genotype B. The preS2/S promoter regulates HBsAg, a key envelope protein of HBV. The preS2/S promoter has an unusual TATA box and is regulated by both liver-specific and ubiquitous transcription factor (24). Our CD spectroscopy, NMR spectroscopy, native PAGE electrophoresis and DMS footprinting results indicate that the PQS in the preS2/S promoter forms a hybrid type intramolecular G-quadruplex. The presence and conservation of this G-quadruplex motif in just one of the ten HBV genotypes suggests a potential role for this DNA secondary structure.

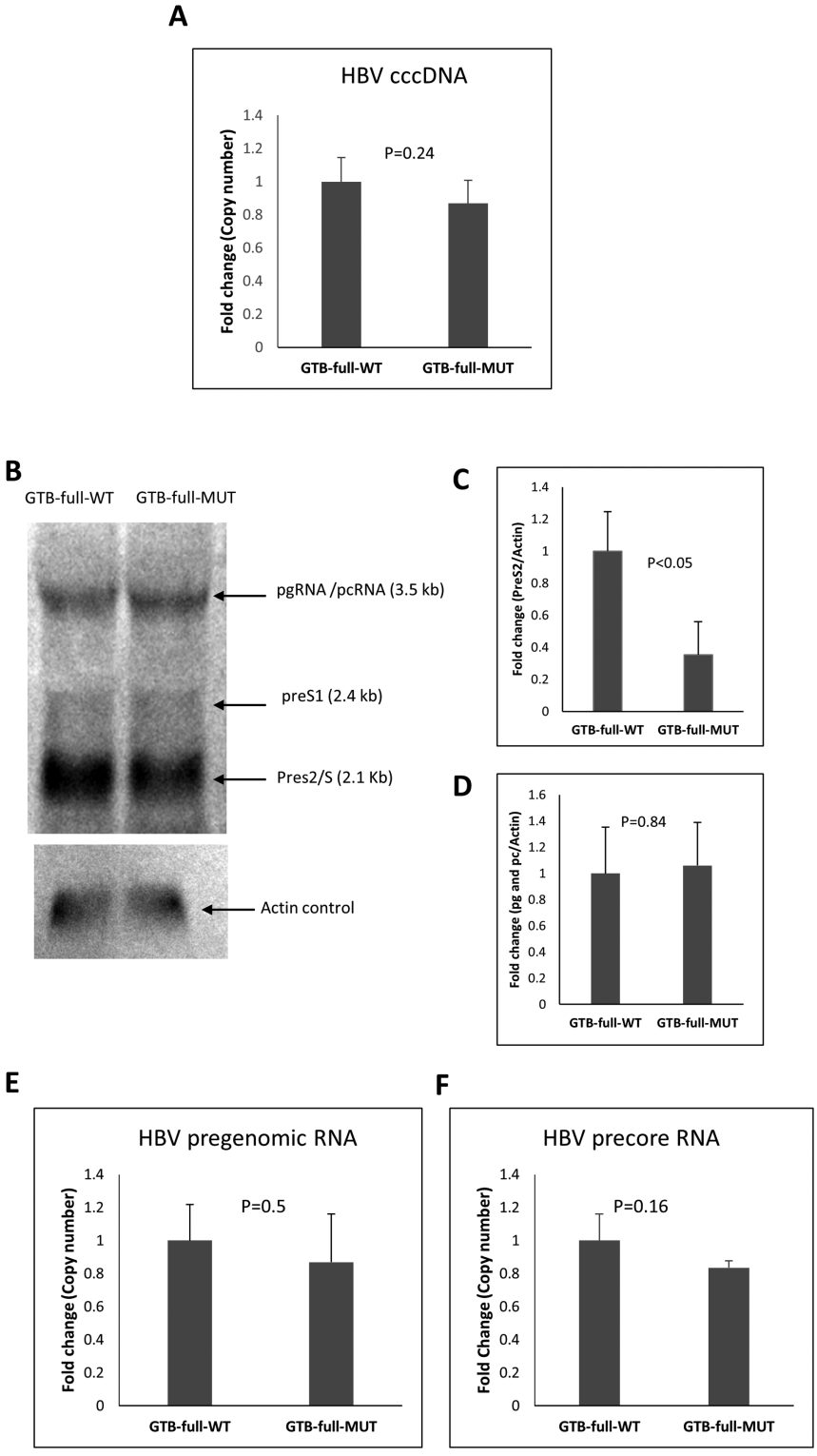

Figure 6. (**A**) The HBV covalently closed circular (cccDNA) levels are comparable in Huh7 cells transfected with the full-length genotype B wild-type construct (GTB-full-WT; the preS2/S promoter of this construct has an intact G-quadruplex motif) and the full-length genotype B mutant construct (GTB-full-MUT; the preS2/S promoter is mutated to disrupt the G-quadruplex motif). The HBV cccDNA levels were measured using real-time PCR. Data are presented as mean ± SD ($n = 4$ replicates). (**B**) Northern blot analysis of HBV RNA from Huh7 cells transfected with GTB-full-WT and GTB-full-MUT constructs. The 3.5 kb HBV transcripts and the 2.4 kb HBV transcripts are comparable between GTB-full-WT and GTB-full-MUT. The GTB-full-MUT is associated with lower levels of the preS2/S transcript (2.1 kb) as compared to the GTB-full-WT. Actin mRNA is used as an internal control for northern blot analysis. Bar graphs showing densitometry analysis of (**C**) the 3.5 kb HBV transcripts (HBV pg RNA and pc RNA) and (**D**) the 2.1 kb transcript in the northern blot. Data are presented as mean ± SD ($n = 3$ replicates). Relative expression levels of (**E**) HBV pregenomic RNA and (**F**) HBV precore (pc RNA) using real-time quantitative PCR. Data are presented as mean ± SD ($n = 4$ replicates)

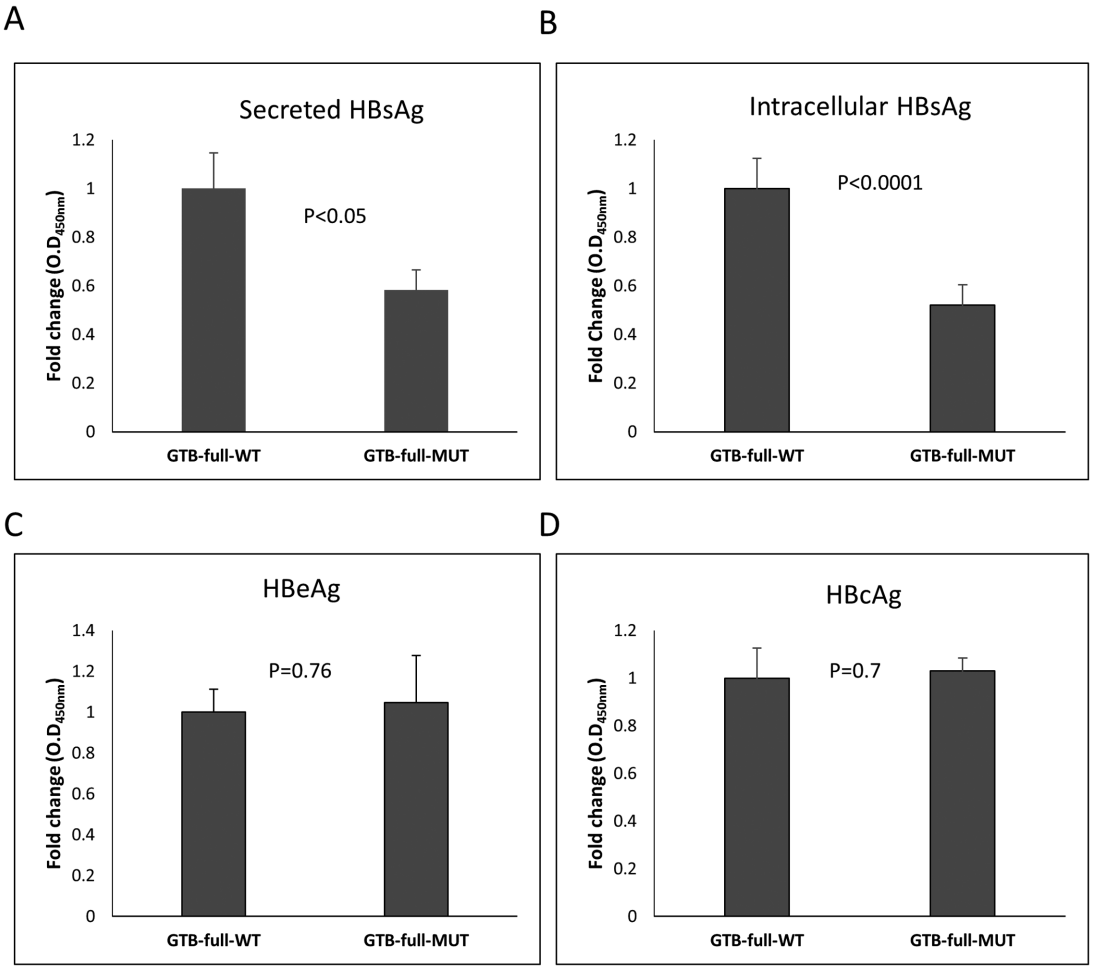

Figure 7. (**A**) The full-length genotype B wild-type construct (GTB-full-WT; the preS2/S promoter of this construct has an intact G-quadruplex motif) is associated with higher levels of secreted hepatitis B surface antigen(HBsAg) as compared to full-length genotype B mutant construct (GTB-full-MUT; the preS2/S promoter is mutated to disrupt the G-quadruplex motif). (**B**) The GTB-full-WT construct was associated with higher levels of intracellular HBsAg as compared to the GTB-full-MUT construct. HBsAg levels were measured using an ELISA. (**C**) The hepatitis B 'e' antigen (HBeAg) levels measured using an ELISA in the supernatant of Huh7 cells transfected with the GTB-full-WT constructs and GTB-full-MUT constructs were comparable. (**D**) Hepatitis B core antigen (HBcAg) levels measured using an ELISA in the cell lysates of Huh7 cells transfected with the GTB-full-WT constructs and GTB-full-MUT constructs were comparable. The intracellular HBsAg and HBcAg levels were normalized to the total protein content in the respective cell lysate. Data are presented as mean ± SD ($n = 3$ replicates)

Our findings using luciferase reporter constructs and HBV surface constructs convincingly demonstrate that the G-quadruplex within the preS2/S promoter of HBV genotype B positively regulates transcription. Mutations disrupting the G-quadruplex motif were associated with impaired promoter activity. In addition, an increase in the promoter activity after stabilization of the G-quadruplex motif with ligands supports our findings on G-quadruplex-mediated enhancement of the preS2/S promoter activity. Theoretically, it has been suggested that the G-quadruplexes in promoter regions may either mask or serve as binding sites for positive or negative transcriptional regulators (44,45). However, most studies analysing the regulatory role of G-quadruplexes in promoter regions found that G-quadruplexes are negative regulators of transcription (6,7,13). There are a few studies that have identified G-quadruplexes in promoter regions as enhancers of transcription (10–12). G-quadruplexes in HIV-1 LTR promoters were identified as transcriptional repressors (14). Similarly, studies on herpesvirus promoters indicate a negative regulatory role of G-quadruplexes (23). Our findings that the G-quadruplex within the preS2/S promoter enhances promoter activity highlight how viruses may mimic host regulatory elements and exploit the host transcriptional machinery to their advantage. In addition, the G-quadruplex we identified in the preS2/S promoter of HBV genotype B sheds light on how G-quadruplexes in virus promoters enhance transcription.

Differences among HBV genotypes in transmission efficiencies, ability to cause persistent infection, pathogenesis and response to antiviral therapy are well-documented (1,46). In addition, HBV genotype-specific differences in virus replication and HBsAg expression (1,47) have been reported; however, the underlying mechanisms remain elusive. Our findings on G-quadruplex-mediated regulation of HBsAg in HBV genotype B provides a novel perspective on

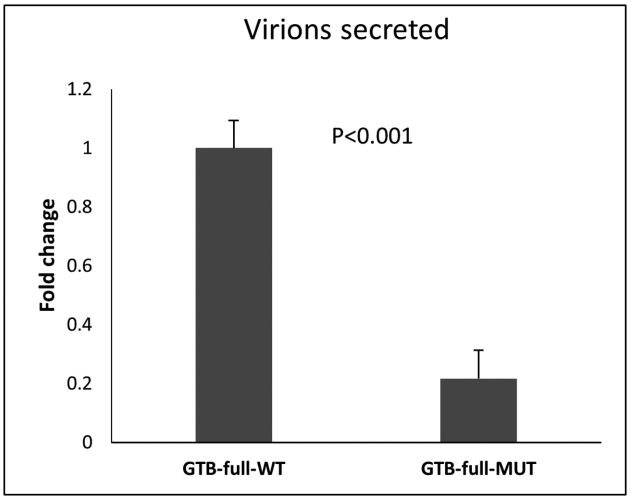

Figure 8. Virion secretion from Huh7 cells transfected with full-length HBV constructs. The full-length genotype B wild-type construct (GTB-full-WT; the preS2/S promoter of this construct has an intact G-quadruplex motif) is associated with significantly higher levels of hepatitis B virion secretion as compared to full-length genotype B mutant construct (GTB-full-MUT; the preS2/S promoter is mutated to disrupt the G-quadruplex motif). Virions were estimated from the supernatants using an immunocapture-based method described in the methods section. Data are presented as mean ± SD ($n = 3$ replicates)

the mechanisms underlying genotype-specific differences in the regulation of HBV transcripts.

The presence of multiple G-quadruplexes in virus genomes and large genome sizes (19,23) make it challenging for mutational studies at the whole genome level. Therefore, previous reports on G-quadruplexes in DNA viruses are based on disruption of G-quadruplex motifs in promoter constructs or modulation of promoter activity by G-quadruplex ligands. The small size of the HBV genome and the presence of a single three tetrad G-quadruplex motif in the entire HBV genome allowed us to study the role of G-quadruplex disrupting mutations at the whole genome level. Our findings with the full-length HBV constructs demonstrate that the G-quadruplex within the preS2/S region of HBV genotype B selectively enhances the preS2/S transcript resulting in increased HBsAg synthesis. This G-quadruplex motif did not influence other HBV transcripts / HBV proteins. Importantly, the disruption of the G-quadruplex leads to a 5-fold reduction in virion secretion. The G-quadruplex-mediated enhancement of virion secretion is more pronounced than the increase in HBsAg levels. This is in keeping with previous studies that have shown that even subtle differences in HBsAg levels may lead to major differences in virion secretion (48,49).

The dependence of HBV genotype B on the G-quadruplex motif to enhance virion secretion explains the conservation of this DNA secondary structure. We believe that other HBV genotypes may have alternative, G-quadruplex independent mechanisms that regulate virion secretion. For example, the transcription factor COUP-TF1 represses transcription from HBV genotype A, while enhancing transcription from genotype D (25). A point mutation that is more common in one genotype than another genotype may adversely affect HBV replication efficiency (2). Differences in evolutionary rates (50) and differences in quasi-species evolution patterns (51) exist among HBV genotypes but the underlying mechanisms remain unclear. In RNA viruses, RNA secondary structures have been reported as evolutionary constraints (52,53). In our study, the presence and the conservation of a G-quadruplex motif in the preS2/S promoter of HBV genotype B also highlights how DNA secondary structures critical for virus replication may act as constraints on virus evolution.

In sum, we have identified a highly conserved DNA G-quadruplex motif within the preS2/S promoter of the HBV genotype B that selectively enhances the expression of a key envelope protein (HBsAg) resulting in increased virion secretion. This work also highlights how G-quadruplexes in virus genomes may contribute to complex genotype-specific regulatory mechanisms.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Prof. Jia-Horng Kao for providing us the genotype B wild-type clone. We would also like to thank Dr Neel S Bhavesh, ICGEB New Delhi for helping us with NMR spectroscopy. Banhi Biswas is a recipient of DBT-JRF fellowship from DBT India.

FUNDING

Intramural funding. Funding for open access charge: Intramural research grants.
Conflict of interest statement. None declared.

REFERENCES

1. Sunbul,M. (2014) Hepatitis B virus genotypes: global distribution and clinical importance. *World J. Gastroenterol. WJG*, **20**, 5427–5434.
2. Qin,Y., Tang,X., Garcia,T., Hussain,M., Zhang,J., Lok,A., Wands,J., Li,J. and Tong,S. (2011) Hepatitis B virus genotype C isolates with wild-type core promoter sequence replicate less efficiently than genotype B isolates but possess higher virion secretion capacity. *J. Virol.*, **85**, 10167–10177.
3. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
4. Biffi,G., Tannahill,D., McCafferty,J. and Balasubramanian,S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–186.
5. Huppert,J.L. and Balasubramanian,S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, **35**, 406–413.
6. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (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.
7. 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.
8. Dexheimer,T.S., Sun,D. and Hurley,L.H. (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.

9. Sun,D., Guo,K., Rusche,J.J. and Hurley,L.H. (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.
10. Fernando,H., Sewitz,S., Darot,J., Tavare,S., Huppert,J.L. and Balasubramanian,S. (2009) Genome-wide analysis of a G-quadruplex-specific single-chain antibody that regulates gene expression. *Nucleic Acids Res.*, **37**, 6716–6722.
11. Lam,E.Y.N., Beraldi,D., Tannahill,D. and Balasubramanian,S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.*, **4**, 1796.
12. Wei,P.-C., Wang,Z.-F., Lo,W.-T., Su,M.-I., Shew,J.-Y., Chang,T.-C. and Lee,W.-H. (2013) A cis-element with mixed G-quadruplex structure of NPGPx promoter is essential for nucleolin-mediated transactivation on non-targeting siRNA stress. *Nucleic Acids Res.*, **41**, 1533–1543.
13. Palumbo,S.L., Memmott,R.M., Uribe,D.J., Krotova-Khan,Y., Hurley,L.H. and Ebbinghaus,S.W. (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.
14. Perrone,R., Nadai,M., Frasson,I., Poe,J.A., Butovskaya,E., Smithgall,T.E., Palumbo,M., Palù,G. and Richter,S.N. (2013) A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J. Med. Chem.*, **56**, 6521–6530.
15. Tlučková,K., Marušič,M., Tóthová,P., Bauer,L., Šket,P., Plavec,J. and Viglasky,V. (2013) Human papillomavirus G-quadruplexes. *Biochemistry*, **52**, 7207–7216.
16. Tuesuwan,B., Kern,J.T., Thomas,P.W., Rodriguez,M., Li,J., David,W.M. and Kerwin,S.M. (2008) Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNA-interactive agents. *Biochemistry*, **47**, 1896–1909.
17. Murat,P., Zhong,J., Lekieffre,L., Cowieson,N.P., Clancy,J.L., Preiss,T., Balasubramanian,S., Khanna,R. and Tellam,J. (2014) G-quadruplexes regulate Epstein-Barr virus–encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.*, **10**, 358–364.
18. Madireddy,A., Purushothaman,P., Loosbroock,C.P., Robertson,E.S., Schildkraut,C.L. and Verma,S.C. (2016) G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Res.*, **44**, 3675–3694.
19. Artusi,S., Nadai,M., Perrone,R., Biasolo,M.A., Palù,G., Flamand,L., Calistri,A. and Richter,S.N. (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.
20. Wang,S.-R., Min,Y.-Q., Wang,J.-Q., Liu,C.-X., Fu,B.-S., Wu,F., Wu,L.-Y., Qiao,Z.-X., Song,Y.-Y. and Xu,G.-H. (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.
21. Fleming,A.M., Ding,Y., Alenko,A. and Burrows,C.J. (2016) Zika virus genomic RNA possesses conserved G-quadruplexes characteristic of the flaviviridae family. *ACS Infect. Dis.*, **2**, 674–681.
22. Artusi,S., Perrone,R., Lago,S., Raffa,P., Di Iorio,E., Palù,G. and Richter,S.N. (2016) Visualization of DNA G-quadruplexes in herpes simplex virus 1-infected cells. *Nucleic Acids Res.*, **44**, 10343–10353.
23. Biswas,B., Kandpal,M., Jauhari,U.K. and Vivekanandan,P. (2016) Genome-wide analysis of G-quadruplexes in herpesvirus genomes. *BMC Genomics*, **17**, 949.
24. Moolla,N., Kew,M. and Arbuthnot,P. (2002) Regulatory elements of hepatitis B virus transcription. *J. Viral Hepat.*, **9**, 323–331.
25. Fischer,S.F., Schmidt,K., Fiedler,N., Glebe,D., Schüttler,C., Sun,J., Gerlich,W.H., Repp,R. and Schaefer,S. (2006) Genotype-dependent activation or repression of HBV enhancer II by transcription factor COUP-TF1. *World J. Gastroenterol. WJG*, **12**, 6054–6058.
26. Qin,Y., Zhou,X., Jia,H., Chen,C., Zhao,W., Zhang,J. and Tong,S. (2016) Stronger enhancer II/core promoter activities of hepatitis B virus isolates of B2 subgenotype than those of C2 subgenotype. *Sci. Rep.*, **6**, 30374.
27. Kropp,K.A., Angulo,A. and Ghazal,P. (2014) Viral enhancer mimicry of host innate-immune promoters. *PLoS Pathog*, **10**, e1003804.
28. Huppert,J.L. and Balasubramanian,S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
29. Babinský,M., Fiala,R., Kejnovská,I., Bednářová,K., Marek,R., Sagi,J., Sklenář,V. and Vorlíčková,M. (2014) Loss of loop adenines alters human telomere d [AG3 (TTAG3) 3] quadruplex folding. *Nucleic Acids Res.*, **42**, 14031–14041.
30. Guedin,A., Alberti,P. and Mergny,J.-L. (2009) Stability of intramolecular quadruplexes: sequence effects in the central loop. *Nucleic Acids Res.*, **37**, 5559–5567.
31. Sattin,G., Artese,A., Nadai,M., Costa,G., Parrotta,L., Alcaro,S., Palumbo,M. and Richter,S.N. (2013) Conformation and stability of intramolecular telomeric G-quadruplexes: sequence effects in the loops. *PLoS One*, **8**, e84113.
32. Laras,A., Koskinas,J., Dimou,E., Kostamena,A. and Hadziyannis,S.J. (2006) Intrahepatic levels and replicative activity of covalently closed circular hepatitis B virus DNA in chronically infected patients. *Hepatology*, **44**, 694–702.
33. Samal,J., Kandpal,M. and Vivekanandan,P. (2015) A simple and rapid method for the quantitation of secreted hepatitis B virions in cell culture models. *Indian J. Med. Microbiol.*, **33**, 290.
34. Paramasivan,S., Rujan,I. and Bolton,P.H. (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods*, **43**, 324–331.
35. Henderson,E., Hardin,C.C., Walk,S.K., Tinoco,I. and Blackburn,E.H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine· guanine base pairs. *Cell*, **51**, 899–908.
36. Williamson,J.R., Raghuraman,M.K. and Cech,T.R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, **59**, 871–880.
37. Qin,Y., Rezler,E.M., Gokhale,V., Sun,D. and Hurley,L.H. (2007) Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. *Nucleic Acids Res.*, **35**, 7698–7713.
38. Víglaský,V., Bauer,L. and Tlučková,K. (2010) Structural features of intra-and intermolecular G-quadruplexes derived from telomeric repeats. *Biochemistry*, **49**, 2110–2120.
39. Read,M., Harrison,R.J., Romagnoli,B., Tanious,F.A., Gowan,S.H., Reszka,A.P., Wilson,W.D., Kelland,L.R. and Neidle,S. (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 4844–4849.
40. White,E.W., Tanious,F., Ismail,M.A., Reszka,A.P., Neidle,S., Boykin,D.W. and Wilson,W.D. (2007) Structure-specific recognition of quadruplex DNA by organic cations: influence of shape, substituents and charge. *Biophys. Chem.*, **126**, 140–153.
41. Koirala,D., Dhakal,S., Ashbridge,B., Sannohe,Y., Rodriguez,R., Sugiyama,H., Balasubramanian,S. and Mao,H. (2011) A single-molecule platform for investigation of interactions between G-quadruplexes and small-molecule ligands. *Nat. Chem.*, **3**, 782–787.
42. Ueda,K., Tsurimoto,T. and Matsubara,K. (1991) Three envelope proteins of hepatitis B virus: large S, middle S, and major S proteins needed for the formation of Dane particles. *J. Virol.*, **65**, 3521–3529.
43. Kidd-Ljunggren,K., Zuker,M., Hofacker,I.L. and Kidd,A.H. (2000) The hepatitis B virus pregenome: prediction of RNA structure and implications for the emergence of deletions. *Intervirology*, **43**, 154–164.
44. Bochman,M.L., Paeschke,K. and Zakian,V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, **13**, 770–780.
45. Kendrick,S. and Hurley,L.H. (2010) The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements. *Pure Appl. Chem.*, **82**, 1609–1621.
46. Vivekanandan,P. and Singh,O.V. (2010) Molecular methods in the diagnosis and management of chronic hepatitis B. *Expert Rev. Mol. Diagn.*, **10**, 921–935.
47. Kao,J.-H. (2011) Molecular epidemiology of hepatitis B virus. *Korean J. Intern. Med.*, **26**, 255–261.
48. Ramanan,V., Shlomai,A., Cox,D.B.T., Schwartz,R.E., Michailidis,E., Bhatta,A., Scott,D.A., Zhang,F., Rice,C.M. and Bhatia,S.N. (2015) CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. *Sci. Rep.*, **5**, 10833.
49. Khan,N., Guarnieri,M., Ahn,S.H., Li,J., Zhou,Y., Bang,G., Kim,K.-H., Wands,J.R. and Tong,S. (2004) Modulation of hepatitis

B virus secretion by naturally occurring mutations in the S gene. *J. Virol.*, **78**, 3262–3270.
50. Zehender,G., De Maddalena,C., Giambelli,C., Milazzo,L., Schiavini,M., Bruno,R., Tanzi,E. and Galli,M. (2008) Different evolutionary rates and epidemic growth of hepatitis B virus genotypes A and D. *Virology*, **380**, 84–90.
51. Wu,Q., Xu,C., Li,J., Li,L., Yan,G., Yue,L., Zeng,Y., Huang,H., Deng,G. and Wang,Y. (2015) Evolution and mutations of hepatitis B virus quasispecies in genotype B and C during vertical transmission. *J. Med. Virol.*
52. Simmonds,P. and Smith,D.B. (1999) Structural constraints on RNA virus evolution. *J. Virol.*, **73**, 5787–5794.
53. Fraile,A. and García-Arenal,F. (1991) Secondary structure as a constraint on the evolution of a plant viral satellite RNA. *J. Mol. Biol.*, **221**, 1065–1069.