BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Human Papillomavirus G-Quadruplexes

Katarína Tlučková,[†] Maja Marušič,[‡] Petra Tóthová,[†] Lubos Bauer,[†] Primož Šket,[‡,§] Janez Plavec,[‡,§,⊥] and Viktor Viglasky*[,†]

[†]Department of Biochemistry, Institute of Chemistry, Faculty of Sciences, P. J. Safarik University, 04001 Kosice, Slovakia

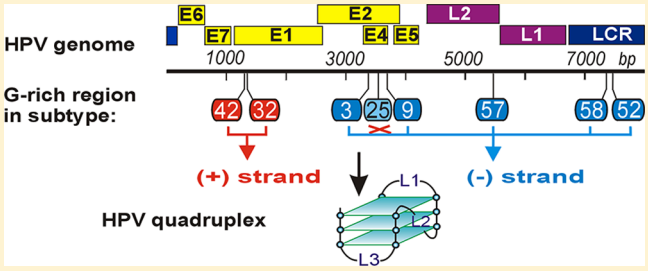
[‡]Slovenian NMR Center, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia

[§]EN-FIST Center of Excellence, SI-1000 Ljubljana, Slovenia

[⊥]Faculty of Chemistry and Chemical Technology, University of Ljubljana, SI-1000 Ljubljana, Slovenia

Supporting Information

ABSTRACT: Infection with human papillomaviruses (HPVs) is one of the most common sexually transmitted infections and can lead to development of head and neck, skin, and anogenital cancer, including cervical cancer, which represents one of the world's most significant health problems. In this study, we analyze G-rich regions in all known HPV genomes in order to evaluate their potential to fold into G-quadruplex structure. Interestingly, G-rich loci fulfilling the criteria for G-quadruplex formation were found in only 8 types of HPV. Nevertheless, viral G-quadruplexes in 7 sequences derived directly from HPVs are confirmed here for the first time. G-rich regions with the capacity to form G-quadruplexes are located in the LCR, L2, E1, and E4 regions of the HPV genome; therefore we assume that regulation processes in viruses could be affected by G-quadruplex formation. Our results represent a starting point for the design of specific ligands with viral G-quadruplex motifs and suggest novel methods for the control of viral replication and transcription.



A remarkable number of nucleic acid sequences with the potential to adopt non-B secondary structures are found within the regulation regions of various genes. Nucleic acid sequences rich in guanine (G) are capable of forming four-stranded structures called G-quadruplexes. These structures are stabilized by eight Hoogsteen hydrogen bonds between G-quartet-forming bases, by coordinated metal ions in the central cavity of the G-quadruplex structure, and also through the stacking of coplanar arrangements of aromatic moieties. There is currently tremendous interest in increasing our understanding of G-quadruplex formation from G-rich sequences. Sequences that have the capacity to form G-quadruplexes are derived from natural sources or can also be artificially designed. Regardless of the origin of these structural motifs, the structures play a significant role in various cellular processes and also show high potential for application in nanotechnology, e.g., quadruplex HIV and thrombin binding aptamers.

G-quadruplexes can also be derived from telomeric repeats. The formation of telomeric G-quadruplexes has been shown to decrease telomerase activity,[1] which is responsible for maintaining the length of telomeres; this enzyme is re-expressed in the vast majority of cancers.[2] Recent years have also seen a growing interest in quadruplex-forming sequences that occur in locations other than at the telomere ends. The formation of G-quadruplexes in the promoters of several genes has been shown to affect their expression and indicates that G-quadruplexes can act as transcriptional regulators,[3] either up-regulating gene expression as in the human insulin gene[4] or, more frequently, down-regulating it as in human *c-myc*,[5,6] human *c-kit*,[7] and many other (onco)genes.[8] Therefore, a considerable effort has been made to find small-molecule ligands that could facilitate the formation and stabilization of G-quadruplex structures.[9] The role of G-quadruplex formation as an important regulating system for gene expression could also be expected for less well studied genomes of bacteria, yeast, and others,[10,11] yet research devoted to G-quadruplex has so far been limited in the field of viral genomes, despite the advantages of their small size and often naturally occurring double-stranded circular (episomal) form. To date only one paper predicts the formation of G-quadruplexes in the Epstein–Barr virus that could disrupt the interaction of EBV nuclear antigen 1 (EBNA1) with RNA; it was found that linking regions of EBNA1 LR1 and LR2 have a strong preference for G-quadruplex RNA and that G-quadruplex RNA-interacting drugs block the functions of EBNA1 that are critical for viral DNA replication and episome maintenance.[12]

The group of viruses that we set out to examine is one of the most widely spread viruses among humans, human papillomaviruses (HPVs). The main focus of the current study is the investigation of G-rich regions in HPV genomes in which stable G-quadruplexes could be formed. HPVs are nonenveloped viruses composed of an episomal DNA of approximately 8

Received: July 8, 2013
Revised: August 30, 2013
Published: September 17, 2013

   dx.doi.org/10.1021/bi400897g | *Biochemistry* 2013, 52, 7207−7216

kilobases in size and show a tropism for squamous epithelium. Different genera establish themselves in different locations; α-HPVs infect mucosal surfaces and skin, while β-, γ-, μ-, and ν-HPVs infect only skin.[13,14] Over 120 types of HPVs have been identified to date, but only a handful of these are harmful. High-risk HPVs cause malignant transformations and precancerous and cancerous changes, while low-risk HPVs cause benign warts and lesions. Mucosal HPVs are therefore classified into high- and low-risk types on the basis of the degree of their association with cervical cancer.[15] Recent years have seen a growing interest in understanding the regulation of HPV gene expression and the structural properties of viral DNA due to the nature of HPV infection and the seriousness of the diseases that it can cause.

The genome of HPV is divided into 3 basic regions: (i) an upstream regulatory region (URR) or long control region (LCR) that contains *cis*-elements that are necessary for the replication and transcription of viral DNA; (ii) the 'early' region, composed of 6 open reading frames (ORFs) labeled E1−E7; and (iii) a 'late' region with two ORFs coding for viral structural proteins L1 and L2, the major and minor capsid proteins, respectively, (Figure 1).[13−15] The E3 region is not highlighted in the figure because this small putative gene has been identified very rarely in *Papillomaviridae*. In addition the gene has not been expressed as a protein and does not appear to serve any function. The E3 gene was identified only in HPV68 type.

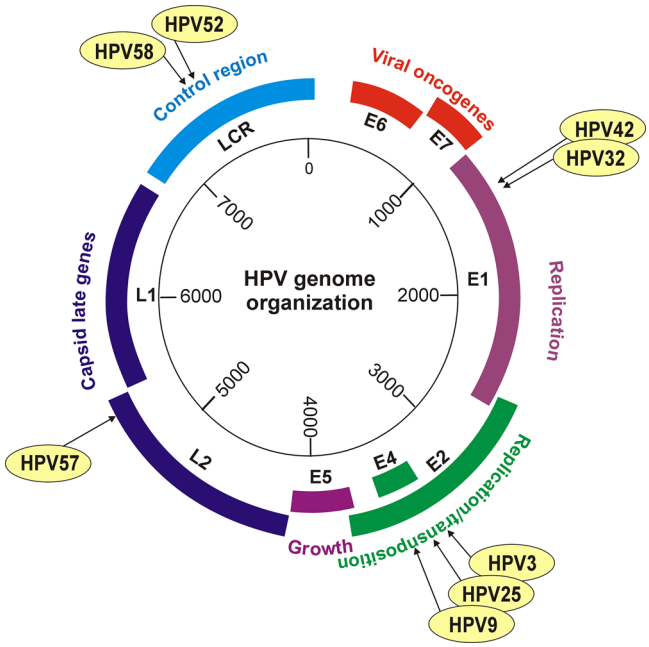


Figure 1. Schematic drawing of the HPV genome organization. The open reading frames encode the five 'early' proteins (E1, E2, E4, E5, E6, and E7) and two 'late' structural proteins (L1 and L2); the numbers indicate approximate positions within the genome. The positions of G-rich regions in appropriate HPVs are also depicted.

Initially all known HPV genomes are investigated in order to identify new potential G-quadruplex-forming sequences. The same searching criteria as used by Burge *et al.* was utilized.[16] Only 8 types of HPVs were found that fulfilled our criteria for the existence of G-rich regions with the potential to form stable G-quadruplexes. Such regions were found in HPV52 and HPV58 in the LCR region, in HPV57 in the coding regions of L2, in HPV32 and HPV42 in E1, and in HPV3, HPV9, and HPV25 in E4/E2 (Table 1 and Figure 1).

HPV58 and HPV52 are α-HPV types that can cause cancer, whereas HPV32, HPV42, HPV3, and HPV57 are low-risk α-subtypes that have been linked to cancer risk. Similarly HPV9 and HPV25 are low-risk β-subtypes. The positions of G-rich oligonucleotides in the LCR, L2, E1, and E4 regions of HPV genome suggest that G-quadruplex formation could alter viral transcription and potentially serve as a basis for novel antiviral therapies. Oligonucleotides with the sequences derived from these DNA regions were studied using spectroscopy (circular dichroism (CD) and UV−vis absorption spectroscopy (UV)), electrophoretic methods (polyacrylamide gel electrophoresis (PAGE) and temperature gradient gel electrophoresis (TGGE)) and ^{1}H nuclear magnetic resonance (NMR). In addition, the stability of HPV G-quadruplexes and their folding patterns were evaluated. One of the main goals of the current study is to ascertain whether oligonucleotide sequences occurring in the HPV regions summarized in Table 2 fold into G-quadruplexes. It is of interest to compare the folding patterns of HPV sequences with the most stable and frequently found G-quadruplexes formed in other promoter regions with respect to strand orientation, loop type, and length in order to find similarities that may indicate the potential biological relevance of a HPV-X G-quadruplex formation (X is 3, 9, 25, 32, 42, 52, 57 and 58).

MATERIAL AND METHODS

All chemicals and reagents were obtained from commercial sources. Acrylamide/bisacrylamide (19:1) solution and ammonium persulfate were purchased from Bio-Rad, *N*,*N*,*N*′,*N*′-Tetramethylethylene-diamine was purchased from Fisher Slovakia. DNA oligomers were obtained from Metabion, Germany (Table 2). PAGE-purified DNA was dissolved in double-distilled water before use. Single-strand concentrations were determined precisely by measuring absorbance (260 nm) at 95 °C using molar extinction coefficients.[17] DNA concentration was determined using UV measurements carried out on a Varian Cary 100 UV−vis spectrophotometer (Amedis, Slovakia). Cells with optical path lengths of 10 mm were used, and the temperature of the cell holder was controlled by an external circulating water bath.

G-Quadruplex Searching Criteria. The sequences of all known human papillomaviruses available in the NCBI Entrez Gene database were analyzed. The search criteria for potential quadruplex sequences were therefore restricted to G_i−L_j−G_i−L_i−G_j−L_i−G_i, where G_i corresponds to a G-tract consisting of at least three G residues ($i \geq 3$) and L_j represents potential loop regions of up to 8 nucleotide units ($j = 1$−8). This sequence of at least four G-tracts with at least three G residues usually leads to formation of G-quadruplexes with three G-quartets that are interrupted by three different loops. The loops, which consist of 1−8 nucleobases,[16] can be formed from any combination of residues; this means that some of the G residues will be located in the loop regions as well. We searched for sequences of both forms: G_i−L_j−G_i−L_i−G_j−L_i−G_i and C-rich corresponding complementary (−) strand C_i−K_j−C_i−K_j−C_i−K_j−C_i. In principle we applied the same algorithm as is utilized by QGRS Mapper, which predicts quadruplex-forming G-rich sequences (QGRS) in nucleotide sequences.[18] The scores of G-quadruplex putative sequences found in HPV genomes were also analyzed by this program.

Table 1. Locations of G-Rich Regions Occurring in Various HPVs

Virus type	[a]Strand Region	[b]Sequence in direction 5'→ 3' and localization	NCBI Reference
HPV-3 (7820 bp)	(-) E4	cacct**GGG**CTT**GGG**T**GGG**CGCTT**GGG**tggta ↑ 3405 ↑ 3385	X74462.1
HPV-9 (7434 bp)	(-) E4	cgacc**GGG**AGT**GGG**AGC**GGG**AAC**GGG**AAC**GGG**ACT**GGG**Acctgg ↑ 3678 ↑ 3645	NC_001596.1
HPV-25 (7713 bp)	(-) E4/E2	gatac**GGG**AGC**GGG**ACT**GGG**ACC**GGG**ACC**GGG**ACC**GGG**acttg ↑ 3569 ↑ 3537	X74471.1
HPV-32 (7961 bp)	(+) E1	agac**GGG**AGTAT**GGG**TAAC**GGGGGGGG**catg ↑ 1321 ↑ 1343	NC_001586.1
HPV-42 (7917 bp)	(+) E1	tacat**GGG**ACTAT**GGG**TAAC**GGGGGGG**cagt ↑ 1300 ↑ 1321	M73236.1
HPV-52 (7942 bp)	(-) LCR	acaca**GGG**TA**GGG**CA**GGGG**ACACA**GGG**TA**GGG**cagga ↑ 7471 ↑ 7445	X74481.1
HPV-57 (7861 bp)	(-) L2	tgcca**GGG**AAA**GGG**TACCTCGA**GGGG**CCGC**GGGG**acatc ↑ 5492 ↑ 5464	X55965.1
HPV-58 (7824 bp)	(-) LCR	aggca**GGG**CA**GGG**TA**GGG**CAATTTA**GGG**acagc ↑ 7384 ↑ 7362	D90400.1

[a]The sense and antisense DNA strands are marked (+) and (−), respectively. [b]Capital letters represent the sequence of oligonucleotides used in our experiments, and lowercase letters represent adjacent at 3′- and 5′-ends. G-runs are underlined.

Table 2. DNA Oligonucleotides Used in This Study and Apparent Melting Temperatures of Oligonucleotides Originating from HPVs[a]

				T_m [°C]			
abbr	nt	ε_{260}[b]	sequence 5′→ 3′	CD: 265 nm	CD: 295 nm	UV: 295 nm	TGGE
HPV3	21	192.4	GGGCTTGGGTGGGCGCTTGGG	ND	60.5	59.9	58.3
HPV9	34	350.3	GGGAGTGGGAGCGGGAACGGGAACGGGACTGGGA	52.2	56.3	56.0	49.5
HPV9-1	23	234.3	CGGGAACGGGAACGGGACTGGGA	55.1	55.2	ND	54.2
HPV9-2	24	249.0	GTGGGAGCGGGAACGGGAACGGGA	56.1	59.0	57.9	57.0
HPV9-3	24	243.9	CCGGGAGTGGGAGCGGGAACGGGA	48.1	ND	ND	49.9
HPV25	33	312.1	GGGAGCGGGACTGGGACCGGGACCGGGACCGGG	66.8	ND	66.1	40.2/67.3[c]
HPV25-1	28	278.4	GGGAGCGGGACTGGGACCGGGACCGGGA	64.7	63.2	67.3	35.1/61.0[c]
HPV25-2	22	220.9	GGGAGCGGGACTGGGACCGGGA	61.7	ND	ND	60.0
HPV25-3	22	218.2	GGGACTGGGACCGGGACCGGGA	57.2	ND	55.6	51.6
HPV32	23	236.4	GGGAGTATGGGTAACGGGGGGGG	ND	59.9	58.9	60.5
HPV42	22	223.8	GGGACTATGGGTAACGGGGGGG	ND	52.4	51.7	51.8
HPV52	27	280.2	GGGTAGGGCAGGGGACACAGGGTAGGG	59.6	58.5	55.1	57.2
HPV52-1	22	227.2	GGGCAGGGGACACAGGGTAGGG	62.3	60.3	59.7	58.2
HPV52-2	23	235.7	GGGTAGGGCAGGGGACACAGGGT	55.6	ND	ND	54.3
HPV-57	29	285.8	GGGAAAGGGTACCTCGAGGGGCCGCGGGG	63.5	66.5	66.2	64.6
HPV-58	23	234.4	GGGCAGGGTAGGGCAATTTAGGG	ND	50.8	49.7	48.6
HTR	21	215.0	GGGTTAGGGTTAGGGTTAGGG	ND	63.4	63.6	63.7
AC9	18	178.0	ACACACACACACACACAC				
AC18	36	354.4	ACACACACACACACACACACACACACACACACACAC				

[a]Modified Britton−Robinson buffer containing 50 mM KCl was used. The melting transition temperature (T_m) of the individual G-quadruplexes was determined from the first derivative of the curve fitted to the melting curves generated by monitoring at 265 and 295 nm. The standard deviation of T_m is ±0.5 °C for all spectral measurements and ±1.5 °C for TGGE. ND = not determined. [b]Millimolar extinction coefficient at 260 nm in (mM^{-1} cm^{-1}). [c]Biphasic transition was detected.

Circular Dichroism Spectroscopy. CD spectra were recorded on a Jasco J-810 spectropolarimeter (Easton, MD, USA) equipped with a PTC-423L temperature controller using a quartz cell of 1-mm optical path length in a reaction volume of 150 μL. All other parameters and conditions were the same as those which were described previously.[17] The modified Britton−Robinson buffer was used in all spectral experiments where TRIS was used instead of KOH (NaOH): 25 mM H_3PO_4, 25 mM boric acid, and 5 mM acetic acid, supplemented by 2.5−50 mM KCl; pH was adjusted by TRIS to a final value of 7.0. The main advantage of this buffer is that it allows the concentration of potassium ions to be modified without any change in pH, and in addition this buffer is more suitable for electrophoresis.

TDS and UV/CD Melting Curves. CD melting profiles were collected at 293 and 265 nm. The thermal stability of quadruplex-forming sequences was also measured by recording the UV absorbance at 293 nm as a function of temperature, using a method similar to that previously published.[17] The temperature ranged from 10 to 100 °C, and the heating rate

dx.doi.org/10.1021/bi400897g | Biochemistry 2013, 52, 7207−7216

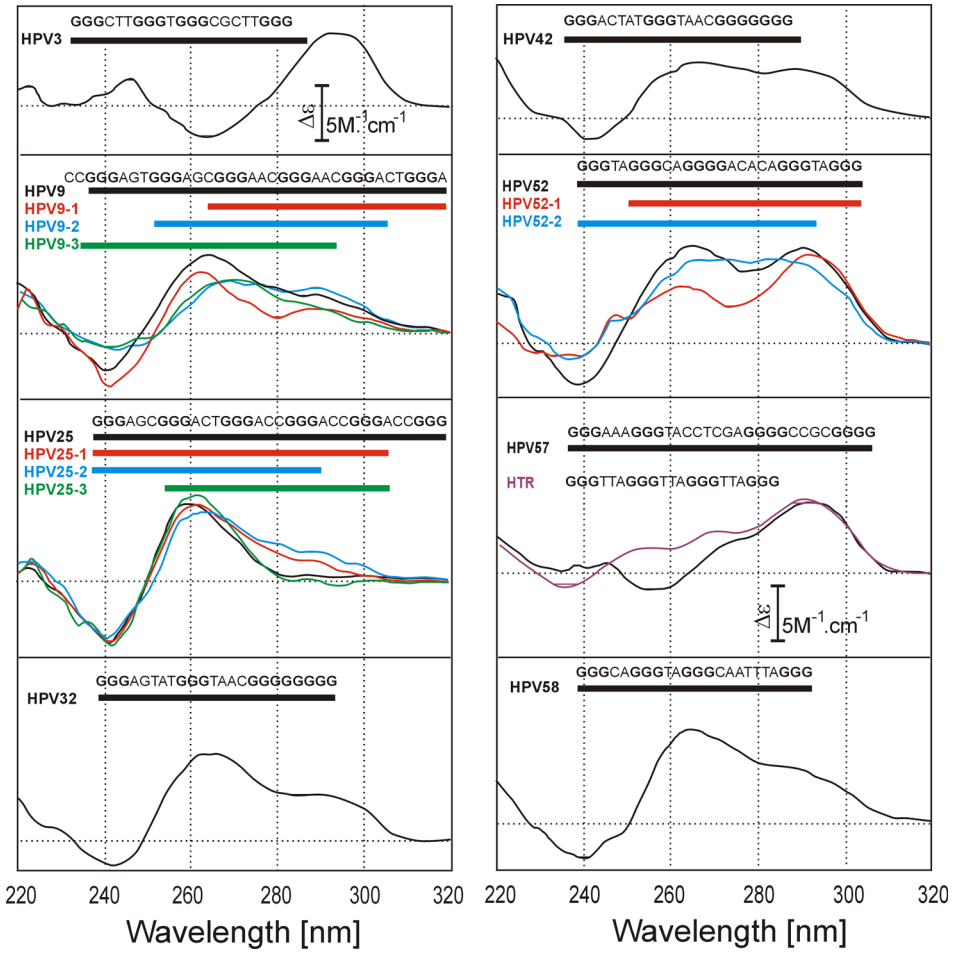


Figure 2. Electronic CD spectra of HPV oligomers in modified 25 mM Britton–Robinson buffer (pH 7.0) in the presence of 50 mM KCl. Each panel depicts the entire G-rich sequence and its shorter derivatives found in original sequences occurring in a given type of HPV. The corresponding UV, TDS, and CD melting curves obtained at 265 and 293 nm are depicted in Figures S1 and S4 (see Supporting Information).

was 0.25 °C per min. Thermal difference spectra (TDS) were obtained by subtraction of UV spectra measured at 10 and 95 °C, respectively.[19] The melting temperature (T_m) was defined as the temperature of the midtransition point. T_m was estimated from the peak value of the first derivative of the fitted curve.

Electrophoresis/TGGE. Native polyacrylamide gel electrophoresis (PAGE) was performed in a temperature-controlled vertical electrophoretic apparatus (Z375039-1EA; Sigma-Aldrich, San Francisco, CA). Gel concentration was 14% (19:1 monomer to bis ratio, Applichem, Darmstadt). Approximately 2 μg was loaded onto 14 × 16 × 0.1 cm gels. Electrophoreses were performed at 8 and 37 °C for 4 h at 126 V (~8 V cm^{-1}). The TGGE equipment used has been described in detail previously.[20] DNA oligomers were visualized with Stains-all immediately after the electrophoresis, and the electrophoretic record was photographed on a white pad with a Nikon D3100 camera.

NMR Experiments. Standard ^{1}H NMR spectra in $H_2O/^2H_2O$ were collected on Agilent Technologies VNMRS and DD2 600 MHz NMR spectrometers. Experiments were performed at 25 °C. Spectra of HPV3, HPV25, HPV25-1, HPV25-2, and HPV25-3 were also acquired at higher temperatures at a range of 35–70 °C. Oligonucleotide strand concentration was in the range from 0.3 to 0.4 mM. 1D ^{1}H NMR spectra were recorded at 0, 5, 10, 25, and 50 mM concentrations of potassium ions. After reaching 50 mM potassium concentration samples were heated to 95 °C and slowly cooled overnight to room temperature. ^{1}H NMR spectra of the samples were recorded after the annealing procedure.

RESULTS AND DISCUSSION

CD spectra, PAGE data, CD–UV thermal melting profiles, and TGGE melting bands of HPV sequences summarized in Table 2 were recorded in a Britton–Robinson buffer containing 50 mM KCl. The presence of potassium ions has a significant effect on G-quadruplex stability, and 50 mM concentration of potassium is usually sufficient for G-quadruplex stabilization.[8] It also permits us to display entire pre- and post-melting regimes of the quadruplex transition obtained through analysis with TGGE and spectral methods, which in turn allows the application of curve fitting analysis.[21−23] However, we also analyzed HPV sequences at different concentrations of potassium and sodium ions ranging from 2.5 to 100 mM KCl and from 25 to 100 mM NaCl (data not shown). The results obtained can be extrapolated to a higher and lower potassium concentration in solution.[17]

CD and UV Spectra Analysis. CD spectra can be used to determine whether the folding conformations of G-quadruplexes are parallel or antiparallel, although this method of interpreting CD data in terms of G-quadruplex topology is still a matter of controversy.[24−26] Nevertheless, for many applications, circular dichroism in combination with another

dx.doi.org/10.1021/bi400897g | Biochemistry 2013, 52, 7207−7216

method has been sufficient to discriminate a certain quadruplex fold from other architectures.[27] It is generally accepted that antiparallel quadruplexes exhibit positive CD signals at ~295 nm, with a negative signal at 260 nm, but that the so-called (3 + 1) conformer also exhibits a shoulder at 265−270 nm.[28] Parallel G-quadruplex structures exhibit a positive band at ~265 nm and a negative peak at 240 nm. This spectral feature is mainly attributed to specific guanine stacking in G-quadruplex structures, because unfolded DNA oligomers do not display these spectral signatures. However, if this positive peak shows a maximum of less than 265 nm, this spectrum may not necessarily indicate the presence of a G-quadruplex. It has been shown that the 'frayed' G-wires can exhibit a positive peak below 265 nm.[29]

Figure 2 shows the CD spectra of HPV oligonucleotides in a buffer containing 50 mM KCl. The profile of the HPV3 spectrum corresponds to an antiparallel arrangement of strands and the profile of the HPV57 and HPV52-1 spectra match the (3 + 1) arrangement of strand orientation that is recorded for HTR (human telomeric repeats) in solutions containing potassium. All other oligomers show positive peaks at ~265 and 295 nm, although HPV9, HPV9-2, HPV9-3, HPV32, HPV42 and HPV58 exhibit a more dominant peak at ~265 nm. Oligomers HPV25, HPV25-1, HPV25-2, and HPV25-3 exhibit a positive peak below 265 nm. TDS studies were performed to verify the formation of G-quadruplex structures[19] (see also Figure S1 in Supporting Information). These results indicate the existence of G-quadruplexes for all studied oligomers under the conditions used with the exception of four of them, HPV25, HPV25-1, HPV25-2, and HPV25-3, all of which show negative and positive peaks at ~294 and ~270 nm, respectively. Despite the fact that the negative signal for the HPV9 series is not clear, the NMR results support the formation of a G-quadruplex structure; see down.

The molar dichroic signal ($\Delta\varepsilon$) falls in the range of 6−10 M^{-1} cm^{-1} for all studied HPV oligomers, and we therefore assume that G-wire-like and/or higher ordered structures are not preferentially formed. Similar values of $\Delta\varepsilon$ were obtained for HTR.[17]

Electrophoresis. Electrophoretic separation can provide valuable information about the molecularity of G-quadruplexes and the presence of multimeric conformers (Figure 3). Electrophoresis was performed in the presence of 50 mM KCl at 8 and 37 °C. DNA oligomers $d(AC)_9$, $d(AC)_{18}$, and $d(G_3TTA)_3G_3$ were used as molecular standards. Under these conditions, HTR forms intramolecular G-quadruplex(es).[17] The lengths of HPV3, HPV32, and HTR oligomers are identical, but the folded G-quadruplexes exhibit different topologies due to their different levels of mobility (Figure 3). HPV3 may form both intramolecular and intermolecular structures, which are indicated by the faster and more slowly moving bands, respectively. The formation of the structure represented by the slower band is strongly dependent on temperature, and this band disappears when the electrophoresis is performed at 37 °C (Figure 3B).

The mobility of intramolecular structures HPV3, HPV9-1, HPV32, HPV42, and HPV58 are very similar, but they are slower than HTR (Figure 3A). On the other hand, the mobilities of HPV52-1 and HPV52-2 are even faster than that of the shorter HTR. Although the HPV52 oligomer is about 8 bases longer, its mobility is still comparable, which suggests that the folded structure is highly compact. The lengths of HPV9 and HPV57 are 34 and 29 nucleotides, respectively; their

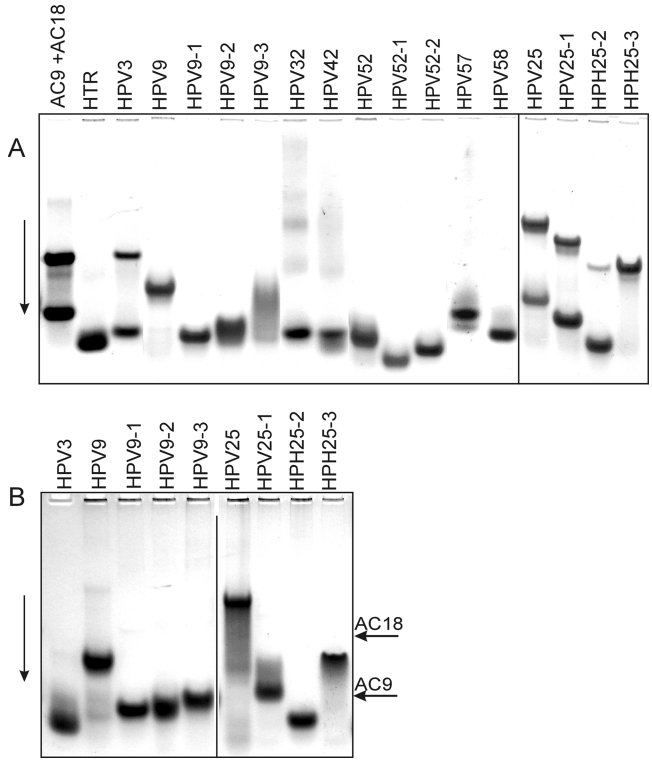


Figure 3. HPV oligomers resolved by gel electrophoresis and visualized by Stains-all staining. Molecular standards (mix of $d(AC)_9$ and $d(AC)_{18}$) and HTR are used. Electrophoretic separation was performed in a 14% polyacrylamide gel at 10 °C in 25 mM Britton−Robinson buffer (pH 7.0) and 50 mM KCl at 8 °C in panel A and 37 °C in panel B. Prior to being used, the DNA sample was heated in the same buffer for 5 min at ~98 °C and slowly cooled to room temperature within 50 min.

mobilities are again higher than those of the unfolded molecular standards to which they correspond in size.

Similarly, HPV9-1, HPV9-2, and HPV9-3 move faster than the unfolded controls that are equivalent in size and therefore adopt a more compact structure. Interestingly, HPV9-3 is less smeared at 37 °C, which suggests that the formation of less stable conformers is eliminated at higher temperatures (Figure 3B). The oligomers derived from HPV25 form also intermolecular assemblies that move significantly slower than intramolecular structures. HPV25, HPV25-1, and HPV25-2 all form intramolecular structural forms. The bands corresponding to intermolecular structures of HPV25-1 and HPV25-2 are not observed at higher temperatures (Figure 3B).

However, as was mentioned above, TDS and CD maxima below 265 nm do not correspond to the formation of G-quadruplexes. A negative signal at ~295 nm was not observed in TDS of HPV25 derivatives (Supplementary Figure S1). These oligomers are G-rich but contain also many cytosine (C) residues allowing Watson−Crick (WC) hydrogen pairing between proximal C and G nucleotides. Therefore, the formation of other non-G-quadruplex structural motifs is more likely. For example, one calculated dimeric structure consisting of only WC pairing to which we applied Zuker's algorithm is depicted in Supplementary Figures S2.[30] This structure could correspond to electrophoretic mobility of HPV25-3.

NMR Results. NMR experiments on HPV oligonucleotides were performed between 5 and 50 mM K^+ ion concentrations.

dx.doi.org/10.1021/bi400897g | Biochemistry 2013, 52, 7207−7216

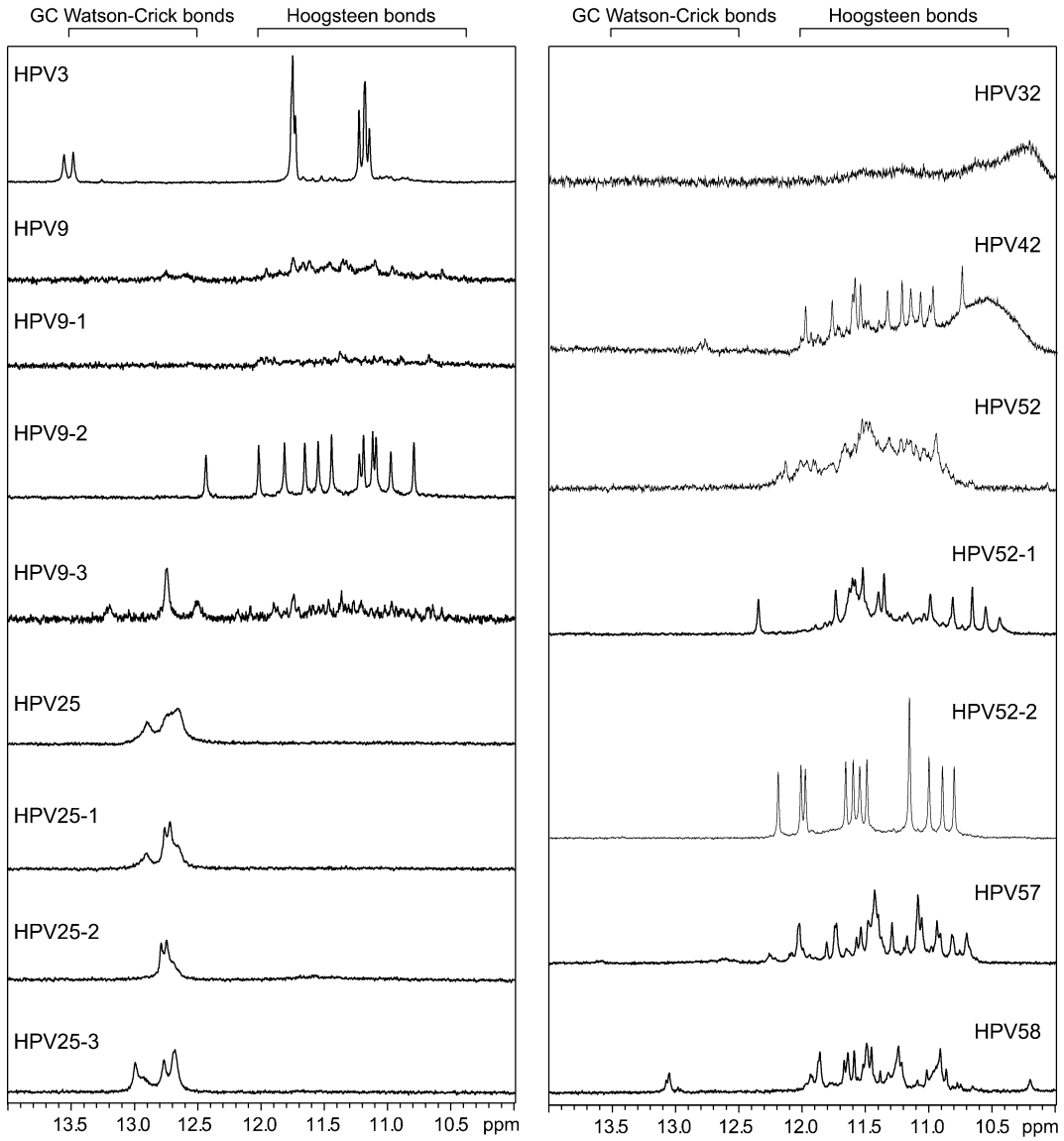


Figure 4. Imino regions of the ^{1}H NMR spectra of the HPV oligonucleotides at 0.3−0.4 mM oligonucleotide concentration per strand. Spectra were collected after annealing in phosphate buffer (pH = 6.8) containing 50 mM KCl at 25 °C.

1D ^{1}H NMR spectra of HPV oligonucleotides after annealing at 50 mM potassium ion concentration are shown in Figure 4. Signals of guanine imino protons in Hoogsteen hydrogen bond alignments appear between ca. δ 10.6 and 12.0 ppm.[31] On the other hand, signals indicating formation of WC base pairs appear at higher chemical shifts between ca. δ 12.5 and 13.7 ppm (Figure 4). Although the concentration of DNA in samples used in NMR is at least 10-fold higher than that used in PAGE and CD measurements, ^{1}H NMR offers valuable complementary information concerning the presence of Hoogsteen and WC base pairs in folded structures.

HPV3 exhibits three and two strong signals between δ 11.1 and 11.3 and δ 11.7 and 11.8 ppm, respectively, after addition of potassium ions. The two signals at δ 13.4 and 13.6 ppm indicate formation of WC base pairs. However, the imino signals corresponding to GC base pairs completely disappear at 50 °C, while signals characteristic of Hoogsteen-bonded imino protons of G residues persist till 65 °C (Supplementary Figure S3). We believe that disappearance of WC signals does not indicate unfolding of dimeric form but merely that these two base pairs are less stable and perhaps their imino protons exchange faster.

The group of HPV9 oligonucleotides displays only small changes in the imino region upon titration of K^{+} ions, whereas subsequent annealing procedure promotes formation of G-quadruplex structures (Figure 4). The ^{1}H NMR spectrum profile of HPV9 indicates that almost no change took place upon titration with K^{+} ions. Only weak signals were observed after annealing between δ 10.5 and 12.0 and δ 12.5 and 12.9 ppm. Similar observations were also made for HPV9-1, where no changes were recorded upon addition of K^{+} ions, but very weak peaks appeared between δ 10.5 and 12.0 ppm after annealing (Figure 4). For HPV9-2, signals of weak intensity appeared at 25 mM K^{+} ion concentration between δ 10.5 and 12.0 ppm, and their intensity did not increase substantially at 50 mM K^{+} ion concentration. However, a very nice and sharp set of signals was observed after annealing of HPV9-2, showing that annealing promotes formation of a single G-quadruplex structure (Figure 4). HPV9-3 displayed a weak signal in the imino region at δ 12.7 ppm before addition of K^{+} ions. This

dx.doi.org/10.1021/bi400897g | *Biochemistry* 2013, 52, 7207−7216

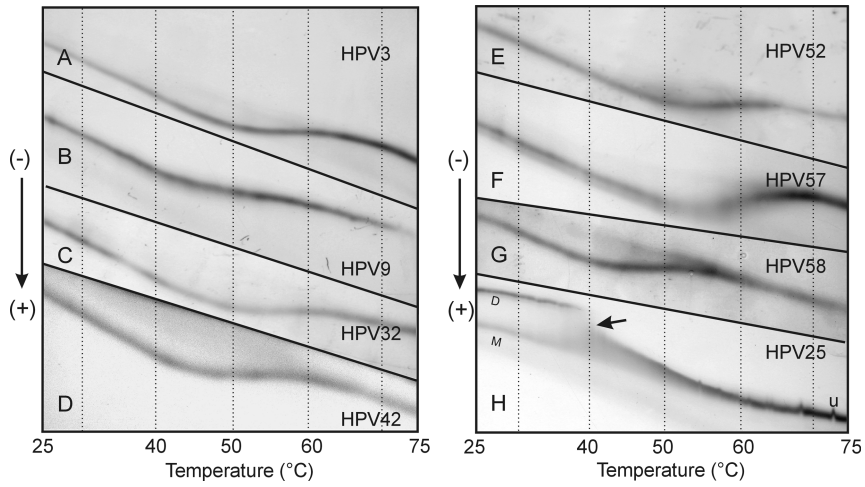


Figure 5. Representative TGGE records of the HPV oligomers in Britton–Robinson buffer in 50 mM KCl: HPV3 (A), HPV9 (B), HPV32 (C), HPV42 (D), HPV52 (E), HPV57 (F), HPV58 (G), and HPV25 (H). Noncontinuous transition of dimer to monomer is labeled by an arrow in (H).

signal increased in intensity at 5 mM K^+ ion concentration, and two additional signals appeared at δ 12.5 and 13.2 ppm, indicating formation of a structure stabilized with WC hydrogen bonds. These three signals could still be observed after annealing, but very weak signals appeared also between δ 10.5 and 12.3 ppm.

Oligonucleotides HPV25 are both G- and C-rich, which enables the formation of stable (including hairpin-like) structures. All samples in this group display reasonably strong signals in the region of δ 12.5–13.0 ppm before addition of potassium ions, a trait that is characteristic of WC base-paired structures. There was no substantial change in these signals upon titration with potassium ions and even after annealing, which indicates that the structures formed by these oligonucleotides are very stable and may prevent the formation of G-quadruplexes. Signals in the imino region of 1D ^{1}H NMR spectra characteristic of GC base pairs in WC geometry completely disappeared only at temperatures higher than 55 °C for HPV25, 50 °C for HPV25-1, 35 °C for HPV25-2, and 55 °C for HPV25-3. No other signals could be observed in the imino region at high temperatures, which suggests that increasing the temperature leads to unfolding of the secondary structures but not the formation of secondary structures with Hoogsteen hydrogen-bonded G-quartets (see also Supplementary Figure S2).

HPV32 exhibits eight sequential guanine residues in the third G-tract, which potentially enables the formation of several G-quadruplex structures. However, it shows strong tendency to form aggregates as indicated by the broad NMR signal between δ 9.9 and 10.9 ppm that is observed already before addition of K^+ ions (Figure 4). Upon titration with K^+ ions additional broad signals appeared in the region between δ 11.0 and 12.0 ppm that decreased in intensity after annealing, showing that albeit oligonucleotide can form G-quadruplex structures, the majority of the molecules formed aggregates. Their formation was supported by the appearance of smeared bands in PAGE gels (Figure 3).

HPV42 has seven sequential guanine residues in the last G-tract, which enables the formation of several G-quadruplex structures as indicated by the presence of the signals between δ 10.9 and 12.2 ppm. In addition, two signals between δ 12.9 and 13.1 ppm correspond to WC base pairing (Figure 4).

Similarly as HPV9, HPV52 comprises several G-runs, which enable them to fold into several topologically different structures as indicated by several signals of low intensities in the region between δ 10.8 and 12.2 ppm. HPV52-1 forms several G-quadruplex structures upon titration with potassium ions and spectral profile is not changed after annealing. HPV52-2 forms a structure consisting of Hoogsteen base-pairs. Before as well as after annealing we observed a very nice and sharp set of 12 signals between δ 10.9 and 12.3 ppm (Figure 4).

For HPV57, imino signals were observed at δ 13.6 ppm and between δ 12.5 and 13.0 ppm. Additional peaks appeared at 50 mM K^+ ion concentration in the region of δ 10.7–11.9 ppm. After annealing, the signals between δ 10.7 and 12.0 ppm became sharper and their intensity increased, while the signals at δ 13.6 ppm and between δ 12.5 and 13.0 ppm almost disappeared, indicating that annealing promoted formation of Hoogsteen and not WC hydrogen-bonded secondary structures.

Sample HPV58 folded into several G-quadruplex structures upon titration with K^+ ions. It was also possible to identify signals corresponding to a GC base-pair in WC geometry (signal at δ 13.1 ppm). No significant changes were observed after annealing of the sample (Figure 4).

CD and UV Melting Curves Analysis. Curve analysis typically assumes a two-state equilibrium between the folded and unfolded forms; this results in an S-shaped dependence of the melting process. However, the presence of polymorphic quadruplex structures usually leads to a melting profile that deviates from this dependence. Non-S shapes of the melting curve have been reported by many authors.[32,33] Melting curves were obtained at 265 and 295 nm with UV absorption and CD spectroscopies, and individual melting curves obtained with CD are shown in Supplementary Figure S4. Slight slopes in pre- and post-melting regimes may arise from intrinsic physical phenomena, such as the intrinsic temperature dependence of absorbance changes resulting from solvent expansion.[22] When the ellipticity maximum was achieved at 265 and/or 295 nm, the CD melting curves were obtained at the same wavelength. UV melting curves were obtained only at 295 nm. The apparent melting temperatures are reported in Table 2.

All quadruplexes are relatively stable; their melting transitions fall in the range from 51 to 65 °C in the presence of 50 mM KCl, but melting temperatures are approximately

dx.doi.org/10.1021/bi400897g | *Biochemistry* 2013, 52, 7207–7216

15−20 °C lower when sodium salt is used instead of potassium (data not shown). All intramolecular G-quadruplex structures studied to date are in agreement with these findings. However, this phenomenon was not observed for HPV25 derivatives; their melting temperatures and melting profiles were not significantly affected in NaCl. Despite the fact that HPV25 derivatives do not fold into G-quadruplexes, the melting temperatures of these WC structures are among the highest of those studied.

Although the second loop-forming region of the HPV57 sequence comprises 8 bases, this sequence forms the most stable HPV-quadruplex. HPV58 and HPV9-1 form the least stable structures.

TGGE of HPV Quadruplexes. TGGE is used to record the dependence of mobility of any DNA sample on temperature, and in addition, different coexisting DNA conformers and their melting curves can be evaluated separately thanks to the main attribute of electrophoresis, the ability to separate different macromolecules.[20] TGGE usually distinguishes between different conformers and formations of intermediates; it offers an objective melting profile over a linear temperature gradient. A linear increase in mobility with increase of temperature does not correspond to any structural change in the molecule, but every nonlinear declination corresponds to a change in the structure. In cases in which two conformers have the same mobilities, their bands are superimposed within a certain range of temperatures, but they usually melt at different temperatures and show different melting profiles. Therefore, we would expect that this type of band would be split within the range of temperatures where the unfolding process begins. However, no such effect was observed for HPV oligomers (Figure 5).The concentration of 50 mM KCl is suitable for recording the entire pre- and post-melting regimes of the quadruplex transition within the temperature range used in TGGE, which is limited in experiments to 10−80 °C. The selected results depicted in Figure 5 clearly confirm a sigmoidal melting profile for all HPV oligonucleotides with the exception of the HPV25-derived sequences.

Electrophoretic melting profiles were analyzed using the fitting procedure that has been described in detail previously.[20] The melting temperatures obtained with TGGE are depicted in Table 2. An analogous procedure cannot be directly applied to noncontinuous melting profiles observed for HPV25 derivatives. The dissociation of bimolecular structure was also observed for these oligomers (marked as D in Figure 5H), but as this process is driven by a significantly slower kinetics, a noncontinuous transition was observed.[20] After the dissociation of the intermolecular structure to an intramolecular structure standard unfolding was observed at 67 °C, but this transition was detectable only with CD spectroscopy. Truncated forms of HPV25 are less stable; in comparison with HPV9 derivatives, the opposite effect was observed. However, the same electrophoretic patterns were observed in the presence of sodium with only a slight change in melting temperatures. This effect was also observed in spectroscopy measurements.

Our results demonstrate that not all of the G-rich HPV-derived sequences studied here show a tendency to form G-quadruplexes. Despite the fact that the electrophoretic separation of G-quadruplexes was performed in a polyacrylamide gel that partially induces mild crowding conditions, the thermal stability studies performed with spectral and electrophoretic methods are in agreement over the findings. It is known that oligonucleotide concentration has a strong effect on the topology and stability of G-quadruplexes.[35] With the exception of the NMR experiments, the oligonucleotide concentration used in CD, UV spectroscopy, and electrophoresis was in the same range. Antiparallel quadruplexes typically display a positive CD signal at ∼295 nm, while parallel quadruplexes display a positive signal at ∼265 nm. However, the sequences of HPV25 are an excellent example of why this property is not sufficient for the confirmation of G-quadruplex folding. Electrophoresis in a nondenaturing condition, NMR, and CD spectra confirm that all other sequences spontaneously form intramolecular G-quadruplexes and some of them also form less stable intermolecular conformers in solution in the presence of 50 mM K^+ ion concentration. The disappearance of an intermolecular band in TGGE at temperatures higher than 40 °C is direct evidence of this, as was seen in the cases of HPV3 and HPV58. These observations were also confirmed by the combined staining procedure of electrophoretic gels. The first staining used with Stains-all for HPV58 can be seen in Figure 5G, and the second staining was performed using a silver staining procedure (see Supplementary Figure S5). The second staining is more sensitive; the smaller population of conformers scarcely visible after the first staining can be seen clearly.[36] However, our study has revealed even more details about the arrangement of the identified intermolecular structure. Its disappearance without 'continuous' transition at higher temperatures is a result of the slow kinetics of transition, but it also shows that this structure is a dimer in which individual subunits of G-quadruplexes coexist; the intramolecular conformers coexist across the entire temperature range. However, the temperature of dissociation also depends on the DNA concentration.[35] This is why we have observed different temperatures with electrophoresis and NMR for this process. This dimeric conformer is completely different from those with interlocked and G-wire -like arrangements, because the unfolding of the conformer is also observed at significantly higher temperatures with CD spectroscopy.

CONCLUSIONS

It is important to mention that G-rich regions are highly conserved in the genome within one HPV type.[37] However, among different types of HPV we observe only poor sequential homology of the studied HPV regions. For the studied types of HPV there are known sequences of several genomic variants except for HPV types 3, 9, and 25. For example, all known genomes of HPV52 and HPV58 isolates in NCBI databases contain the same G-rich sequence in their LCR region; the genome analysis covered 25 and 70 isolates of HPV52 and HPV58, respectively. G-rich sequences located in E1 gene in 24 and 8 variants of HPV42 and HPV3, respectively, perfectly match the sequence of the studied oligonucleotides. The G-rich L2 region in three different HPV57 isolates is the same too.

We have shown that the genome of some HPVs consist of G-rich sequences that can form stable G-quadruplexes. Their stabilities are comparable with those of other known G-quadruplexes, but their topology has not yet been described in detail. This study shows that sequences consisting of at least four G-runs interrupted by three nucleotides may not fold into G-quadruplexes. If they contain a critical amount of cytosine residues, they prefer alternative WC base-paired conformations; this is particularly true for HPV25 derivatives. Interestingly, HPV9 and HPV52 genomic DNA contains G-rich regions where six and five G-runs, respectively, were found. However, any G-quadruplex consumes only four G-runs; therefore in

sequences consisting of more than four G-runs, more than one arrangement is feasible. We cannot forget that the sequences that are not directly associated with G-quadruplex motif, the protruding sequence, can also influence G-quadruplex folding. Such effect was observed for telomeric repeats by many authors. In this study we are focusing mainly on the HPV sequences with minimal protrusions.

■ ASSOCIATED CONTENT

Supporting Information

Additional figures as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +421 55 2341262. Fax: +421 55 6222124. E-mail: viktor.viglasky@upjs.sk.

Funding

This work was supported by the Slovak Research and Development Agency under contract No. APVV-0280-11 and APVV-0134-11, Slovenian Research Agency (Grants P1-0242 and J1-4020), EAST-NMR FP7 project (contract 228461), Bio-NMR FP7 project (contract 261863), European Cooperation in Science and Technology (COST MP0802), and Slovak Grant Agency (1/0153/09).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank G. Cowper and L. Sieber for critical reading of the manuscript.

■ REFERENCES

(1) Fletcher, T. M., Sun, D., Salazar, M., and Hurley, L. H. (1998) Effect of DNA secondary structure on human telomerase activity. *Biochemistry 37*, 5536−5541.

(2) Mergny, J. L., Riou, J. F., Mailliet, P., Teulade-Fichou, M. P., and Gilson, E. (2002) Natural and pharmacological regulation of telomerase. *Nucleic Acids Res. 30*, 839−865.

(3) 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.

(4) Catasti, P., Chen, X., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) Structure-function correlations of the insulin-linked polymorphic region. *J. Mol. Biol. 264*, 534−545.

(5) 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.

(6) Brown, R. V., Danford, F. L., Gokhale, V., Hurley, L. H., and Brooks, T. A. (2011) Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex. *J. Biol. Chem. 286*, 41018−41027.

(7) Fernando, H., Reszka, A. P., Huppert, J., Ladame, S., Rankin, S., Venkitaraman, A. R., Neidle, S., and Balasubramanian, S. (2006) A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry 45*, 7854−7860.

(8) Qin, Y., and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie 90*, 1149−1171.

(9) Neidle, S., and Parkinson, G. H. (2002) Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Discovery 1*, 383−393.

(10) Rawal, P., Kummarasetti, V. B. R., Ravindran, J., Kumar, N., Halder, K., Sharma, R., Mukerji, M., Das, S. K., and Chowdhury, S. (2006) Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation. *Genome Res. 16*, 644−655.

(11) Hershman, S. G., Chen, Q., Lee, J. Y., Kozak, M. L., Yue, P., Wang, L., and Johnson, F. B. (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae. *Nucleic Acids Res. 36*, 144−156.

(12) Norseen, J., Johnson, F. B., and Lieberman, P. M. (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.

(13) Myers, G., Halpern, H., Baker, C., McBride, A., Doorbar, J., and Wheeler, C. (1996) Human Papillomaviruses 1996: A compilation and analysis of nucleic acid and amino acid sequences. *Theoretical Biology and Biophysics*, Los Alamos National Laboratory, Los Alamos, NM.

(14) de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., and zur Hausen, H. (2004) Classification of papillomaviruses. *Virology 324*, 17−27.

(15) Grm, H. S., Bergant, M., and Banks, L. (2009) Human papillomavirus infection, cancer & therapy. *Indian J. Med. Res. 130*, 277−285.

(16) 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.

(17) Víglaský, V., Bauer, L., and Tlucková, K. (2010) Structural features of intra- and intermolecular G-quadruplexes derived from telomeric repeats. *Biochemistry 49*, 2110−2120.

(18) Kikin, O., D'Antonio, L., and Bagga, P. S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res. 34* (Web Server issue), W676−682.

(19) Mergny, J. L., Li, J., Lacroix, L., Amrane, S., and Chaires, J. B. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res. 33*, e138.

(20) Víglaský, V., Antalík, M., Bagel'ová, J., Tomori, Z., and Podhradský, D. (2000) Heat-induced conformational transition of cytochrome c observed by temperature gradient gel electrophoresis at acidic pH. *Electrophoresis 21*, 850−858.

(21) Smargiasso, N., Rosu, F., Hsia, W., Colson, P., Baker, E. S., Bowers, M. T., DePauw, E., and Gabelica, V. (2008) G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation. *J. Am. Chem. Soc. 130*, 10208−10216.

(22) Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligonucleotides. *Nucleic Acids Res. 31*, 2097−2107.

(23) Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett. 435*, 74−78.

(24) Gray, D. M., Wen, J. D., Gray, C. W., Repges, R., Repges, C., Raabe, G., and Fleischhauer, J. (2008) Measured and calculated CD spectra of G-quartets stacked with the same or opposite polarities. *Chirality 20*, 431−440.

(25) Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., and Spada, G. P. (2010) A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem. 8*, 2683−2692.

(26) Vorlíčková, M., Kejnovská, I., Bednářová, K., Renčiuk, D., and Kypr, J. (2012) Circular dichroism spectroscopy of DNA: from duplexes to quadruplexes. *Chirality 24*, 691−698.

(27) Karsisiotis, A. I., Hessari, N. M., Novellino, E., Spada, G. P., Randazzo, A., and Webba da Silva, M. (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem., Int. Ed. 50*, 10645−10648.

(28) Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006) Structure of the human telomere in K+ solution: An intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc. 128*, 9963−9970.

(29) Abu-Ghazalah, R. M., Irizar, J., Helmy, A. S., and Macgregor, R. B., Jr. (2010) A study of the interactions that stabilize DNA frayed wires. *Biophys. Chem. 147*, 123−129.

(30) Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406−3415.
(31) Webba da Silva, M. (2007) NMR methods for studying quadruplex nucleic acids. *Methods 43*, 264−277.
(32) Lane, A. N., Chaires, J. B., Gray, R. D., and Trent, J. O. (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res. 36*, 5482−5515.
(33) Chaires, J. B. (2009) Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *FEBS J. 277*, 1098−1106.
(34) Wallimann, P., Kennedy, R. J., Miller, J. S., Shalongo, W., and Kemp, D. S. (2003) Dual wavelength parametric test of two-state models for circular dichroism spectra of helical polypeptides: anomalous dichroic properties of alanine-rich peptides. *J. Am. Chem. Soc. 125*, 1203−1220.
(35) Palacký, J., Vorlíčková, M., Kejnovská, I., and Mojzeš, P. (2013) Polymorphism of human telomeric quadruplex structure controlled by DNA concentration: a Raman study. *Nucleic Acids Res. 41*, 1005−1016.
(36) Bassam, B. J., and Gresshoff, P. M. (2007) Silver staining DNA in polyacrylamide gels. *Nat. Protoc. 2*, 2649−2654.
(37) Chen, Z., Schiffman, M., Herrero, R., Desalle, R., Anastos, K., Segondy, M., Sahasrabuddhe, V. V, Gravitt, P. E, Hsing, A. W., and Burk, R. D. (2011) Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One 6*, e20183.