Published online 8 April 2015 *Nucleic Acids Research, 2015, Vol. 43, No. 9* ***4733–4745***
doi: 10.1093/nar/gkv296

Unique *C. elegans* telomeric overhang structures reveal the evolutionarily conserved properties of telomeric DNA

Petra Školáková[1,†], Silvie Foldynová-Trantírková[2,3,†], Klára Bednářová[1,2], Radovan Fiala[2], Michaela Vorlíčková[1,2,*] and Lukáš Trantírek[2,*]

[1]Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i., Kralovopolska, 135, 612 65 Brno, Czech Republic, [2]Central European Institute of Technology, Masaryk University, Kamenice 735/5, 625 00 Brno, Czech Republic and [3]Institute of Parasitology, Academy of Sciences of the Czech Republic, Branisovska, 31, 375 05 Ceske Budejovice, Czech Republic

Received January 04, 2015; Revised March 10, 2015; Accepted March 25, 2015

ABSTRACT

There are two basic mechanisms that are associated with the maintenance of the telomere length, which endows cancer cells with unlimited proliferative potential. One mechanism, referred to as alternative lengthening of telomeres (ALT), accounts for approximately 10–15% of all human cancers. Tumours engaged in the ALT pathway are characterised by the presence of the single stranded 5′-C-rich telomeric overhang (C-overhang). This recently identified hallmark of ALT cancers distinguishes them from healthy tissues and renders the C-overhang as a clear target for anticancer therapy. We analysed structures of the 5′-C-rich and 3′-G-rich telomeric overhangs from human and *Caenorhabditis elegans*, the recently established multicellular in vivo model of ALT tumours. We show that the telomeric DNA from *C. elegans* and humans forms fundamentally different secondary structures. The unique structural characteristics of *C. elegans* telomeric DNA that are distinct not only from those of humans but also from those of other multicellular eukaryotes allowed us to identify evolutionarily conserved properties of telomeric DNA. Differences in structural organisation of the telomeric DNA between the *C. elegans* and human impose limitations on the use of the *C. elegans* as an ALT tumour model.

INTRODUCTION

Telomeres are specialised, functional DNA-protein structures found at the ends of all eukaryotic linear chromosomes. Telomeres help to protect the ends of chromosomes from being treated as damaged DNA in need of repair and also provide a means of complete replication of the chromosome. In most eukaryotes, telomeres comprise double-stranded DNA consisting of simple tandem repeat sequences (e.g. d(TTAGGG).(CCCTAA) in vertebrates) terminating in a single-stranded, 150- to 250-nt long G-rich 3′-overhang, herein referred to as the G-overhang (1,2). The G-overhang is essential for the formation of a protective higher-order structure referred to as the T-loop (3). Telomeric DNA in cultured cells has been shown to undergo progressive loss of 50−200 bases following each cell division due to incomplete replication of the lagging strand (4,5). This process ultimately leads to critically short telomeres, which triggers cell cycle arrest (senescence) or cell death (reviewed in 6,7). In most eukaryotic organisms, telomere maintenance relies on the activation of telomerase (8). Appropriate telomerase function in stem and germ cell populations counteracts replication-dependent telomere shortening (reviewed in 9). In contrast, repression of telomerase expression in most somatic tissues limits the proliferative potential of these cells (reviewed in 6,7). Thus, senescence, which is induced by telomere shortening and subsequent DNA damage signalling, is an essential tumour suppression mechanism, emphasised by the finding that repression of telomerase is lost from approximately 85–90% of cancer cells and primary tumours, endowing them with unlimited proliferative potential (10,11).

The inherent characteristic of the mammalian G-overhang is its ability to adopt a non-canonical tetraplex DNA structure, defined as the G-quadruplex, both *in vitro* (reviewed, e.g. in 12) and *in vivo* (13,14). The G-quadruplex structure relies on the formation of planar guanine tetrads marked by a Hoogsteen-type guanine-guanine base-pairing

*To whom correspondence should be addressed. Tel: +420 549496476; Fax: +420 549492556; Email: lukas.trantirek@ceitec.muni.cz
Correspondence may also be addressed to Michaela Vorlíčková. Tel: +420 541 517 188; Fax: +420 541 211 293; Email: mifi@ibp.cz
†These authors contributed equally to the paper as first authors.

pattern. Folding of single-stranded, G-rich telomeric DNA, which serves as a substrate for telomerase, into a four-stranded G-quadruplex has been demonstrated to inhibit telomerase activity (15), as G-quadruplex formation ensures that the 3′-end is inaccessible to hybridisation with the telomerase RNA template, the first essential step in the telomerase catalytic cycle. Several small molecular weight ligands that stabilise telomeric G-quadruplex structures have displayed promising anticancer activity in tumour xenograft models, indicating that the stabilisation of telomeric G-quadruplexes might be applicable to the treatment of a wide range of human cancers (reviewed in, e.g. (16,12).

Although most human cancers rely on the activation of telomerase (8) to maintain the telomere length, a distinct telomerase-independent mechanism is active in a subset (∼10%) of tumours (17). In these tumours, telomeres are maintained using the homologous recombination (HR)-based alternative lengthening of telomeres (ALT) mechanism (17). Tumour cells engaged in this pathway display several distinctive characteristics (18,19) particularly the single–stranded, C-rich telomeric 5′-overhang, herein referred to as a C-overhang (20).

Analogous to mammalian G-rich telomeric DNA, the inherent property of the mammalian complementary C-rich telomeric DNA strand is its ability to form a non-canonical DNA tetraplex structure, defined as the i-motif (21,22). In contrast to the G-quadruplex, which requires the formation of Hoogsteen-type guanine-guanine base pairs, the i-motif structure is based on the formation of C^+.C non-canonical base pairs, presuming the protonation of cytosine bases (23,24). Whereas G-quadruplex formation was demonstrated both *in vitro* and *in vivo*, i-motif formation from telomeric DNA has, thus far, only been observed *in vitro* either under acidic (21,22) or neutral pH conditions in the presence of molecular crowding mimics (25). Whether i-motif formation in mammalian telomeric C-rich DNA is only a peculiarity of its conformational space under specific environmental conditions or is associated with telomere function remains to be elucidated.

In this study, we present the results of comparative structural analyses of DNA constructs emulating both 3′-G-rich and 5′-C-rich single-stranded telomeric DNA corresponding to the G- and C-overhangs, respectively. Particularly, we focus on comparison of structural properties of telomeric DNAs from model nematode *Caenorhabditis elegans* and that of mammals. Thus far, *C. elegans* and mammals (humans and mice) provide the only reported cases of multicellular eukaryotes displaying linear chromosomes containing, next to its 3′-G-rich overhang, a 5′-C-rich single-stranded telomeric overhang (20,26). Importantly, while humans, similar to the majority of eukaryotic organisms, display a slightly alkaline intracellular pH (pHi) shift in the course of cell cycle progression (reviewed in 27), *C. elegans* is a rare example of a eukaryote that exhibits notable acidification (pHi ∼ 6.3) of the intracellular space during its development (28). Comparison of the unique structural characteristics of G- and C-rich telomeric DNA from the nematode *C. elegans* and those from other species allowed us identify the properties of telomeric DNA that are conserved during evolution. Our data suggest that the detected variability in the primary sequence of telomeric DNA repeats among eukaryotic species is not a consequence of a positive selection force during evolution but rather a manifestation of the conservation of telomeric DNA structural characteristics, which allow a limited set of permutations in the telomeric DNA primary sequence.

The implications of the C-overhang structure in the therapeutic intervention of tumours engaged in the ALT pathway and in the recently proposed use of *C. elegans* as a model for ALT (29,30) are discussed.

MATERIALS AND METHODS

DNA constructs

The unlabelled DNA oligonucleotides listed in Table 1 were purchased from IDT (USA). Sequences for the *S. cerevisiae* constructs (Table 1) were derived from study by Shampey et al. (31). Desalted and lyophilised oligonucleotides were dissolved in deionised water, and their concentrations were determined using an UNICAM 5625 UV/Vis spectrometer on the basis of UV absorption at 260nm measured at 90°C and molar absorption coefficients calculated according to Gray et al. (32). The $^{13}C/^{15}N$ fully isotopically labelled analogue of the CE20 sequence (see Table 1), which was extended at its 5′-end with a guanosine residue to restrict the terminal fraying effect at the bottom of the hairpin stem (oligonucleotide CE21 in Table 1), was prepared enzymatically (33) and was purified via ion exchange chromatography (Silantes, Germany). To avoid structural artefacts caused by the ion composition of the buffer (34), the oligonucleotide experiments were performed in buffer mimicking the ion composition of the intracellular environment (25 mM sodium phosphate, 110 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 130 nM $CaCl_2$, 10% D_2O (pH 7.5)), herein referred to as intracellular buffer, at a DNA strand concentration of approximately 0.3 mM (unlabelled oligonucleotides) or 0.5 mM (isotopically labelled oligonucleotide, CE21 - see Table 1). Notably, identical samples were employed for the acquisition of both the CD and NMR spectra as well as for native PAGE.

CD spectroscopy

CD spectra were measured using the Jasco J815 spectrometer in 0.1- and 0.2-mm path length Hellma cells placed in a Peltier holder. CD signals were expressed as the difference in the molar absorption $\Delta\epsilon$ of the left- and right-handed circularly polarised light. The molarity was related to nucleosides. Spectra were acquired at a rate of 100 nm/min and averaged from four measurements. The pH values of the samples were adjusted via the addition of HCl and measured using a Sentron Titan pH meter.

Native PAGE

Non-denaturing PAGE was performed in a temperature-controlled submersible apparatus (SE600; Hoefer Scientific) containing circulating buffer. The gels (16%, 29:1 acrylamide:bisacrylamide) were electrophoresed for 19 hours at 25 V and 20°C in the intracellular buffer. The gels were stained with Stains-All (Sigma). Densitometry was performed using a Personal Densitometer (SI, Model 375-A).

Table 1. List of telomeric DNA constructs employed

Organism	Telomeric repeat (5′ → 3′)	Construct	Name
H. sapiens	TAACCC	$CCC(TAACCC)_3$	HS21
		$CCC(TAACCC)_7$	HS45
	GGGTTA	$(GGGTTA)_3GGG$	gHS21
		$(GGGTTA)_7GGG$	gHS45
C. elegans	TAAGCC	$CC(TAAGCC)_3$	CE20
		$GCC(TAAGCC)_3$	CE21
		$CC(TAAGCC)_7$	CE44
	GGCTTA	$(GGCTTA)_3GG$	gCE20
		$(GGCTTA)_3GGC$	gCE21
		$(GGCTTA)_7GG$	gCE44
		$(GGCTTA)_7GGC$	gCE45
S. cerevisiae	$C_{1-3}A$	CCACACCCACACCACACCC	SA-a19
		$CCCAC(ACACCC)_3$	SA-b23
	$TG_{1—3}$	GGGTGTGGTGTGGGTGTG	gSA-a19
		$(GGGTGT)_3GTGGG$	gSA-b23
T. thermophyla	CCCCAA	$CCCC(AACCCC)_3$	TT22

UV absorption melting and thermal difference spectra

Melting curves (5–90°C, 90–5°C and again 5–90°C) were measured using a Varian Cary 4000 spectrophotometer in 1-mm cells at a strand concentration of approximately 0.02 mM. The absorbance was recorded at 260 nm (foldback and i-motif) and at 297 nm (G-quadruplex), and it was considered as the absorbance of the ordered structure.

NMR spectroscopy

The 1D proton spectra of the unlabelled samples were measured at 700 MHz using a Bruker Avance III NMR spectrometer equipped with a triple resonance room temperature probe using the WATERGATE pulse sequence (35), including 1-ms rectangular selective pulses to suppress the water signal. All of the spectra were measured in water solution (90% H_2O/10% D_2O) at 20°C or 1°C unless stated otherwise. All of the NMR experiments of the $^{13}C/^{15}N$-labelled samples were measured at 850 MHz using a Bruker Avance III spectrometer equipped with a triple-resonance cryogenic probe. All spectra of the labelled samples were measured in water solution (90% H_2O/10% D_2O) at 10°C or 1°C. The ^1H-^{13}C 2D correlation spectrum was measured using base-optimised constant-time HSQC (36,37) with gradient selection. The constant time was set to 15 ms, and the proton-carbon evolution time $1/(2J_{HC})$ was set to 2.5 ms. The spectral widths were 9 ppm and 80 ppm in the ^1H and ^{13}C dimensions, respectively; the proton, carbon and nitrogen irradiation frequencies were set at 4.7 ppm, 120 ppm and 157.5 ppm, respectively. The ^{15}N magnetisation was refocused during t_1, and GARP decoupling was applied during acquisition using irradiation fields of 5.56 kHz and 1.47 kHz for ^{13}C and ^{15}N, respectively. All 180° carbon pulses were applied as 20% smoothed Chirp pulses (38) swept over 60 kHz. The NOESY spectrum, utilising a 250-ms mixing time and spectral widths of 21 ppm in both dimensions, were recorded using a 3–9–19 water suppression scheme and a 1-ms rectangular flip-back pulse (39,40) on each fully $^{13}C/^{15}N$-labelled sample. In the directly and indirectly detected dimensions, 2048 and 1200 real points were recorded, respectively. The effect of heteronuclear coupling was removed using refocusing pulses in t_1 evolution and decoupling during the detection period in the same manner as described for the HSQC experiment. The 2D ^1H-^1H version of the HCCH-TOCSY experiment to correlate the adenine H2 and H8 resonances (41,42) was measured using spectral widths of 10 ppm in both dimensions and 60 ms DIPSI-3 mixing using a 5.2-kHz irradiation field set at 150 ppm. GARP decoupling using a 5.56-kHz irradiation field was applied to carbon during acquisition. The HNN-COSY spectra (43,44) were measured with 21-ppm and 116-ppm spectral widths for ^1H and ^{15}N, respectively. The evolution times for the one-bond proton-nitrogen coupling $1/(^2J_{HN})$ and the nitrogen-nitrogen coupling through hydrogen bond $1/(^2J_{NN})$ were set to 5 ms and 30 ms, respectively. The WATERGATE scheme (35) consisting of 1-ms rectangular pulses was used to suppress the water signal, and GARP4 decoupling using a 1.47-kHz field was applied to nitrogen during acquisition.

Pull-down assay

Streptavidin magnetic beads (Promega; Streptavidine Magne-Sphere Paramagnetic Particles; 50 μg/rci) were washed twice with 0.5× SSC and once with 1× binding buffer (10 mM Tris, pH 6.7; 5 mM KCl; 10% glycerol; 1 mM DTT; 0.5 mM $MgCl_2$; 1 mM EDTA; and 1 pill of protease inhibitor cocktail: complete, EDTA free, Roche). Biotinylated HS21 oligo (2.6 μM/rci) was bound to the beads in 1× binding buffer and placed in a roller at room temperature for 30 min. The non-bound DNA was removed by washing (3 times with 1× binding buffer at room temperature for 5 min). The magnetic beads coupled with biotinylated DNA oligos were separated into individual reactions, mixed with an increasing amount of TMPyP4 (Sigma Aldrich) (at a molar ratio of ∼ 0/1; 0.05/1; 1/1 or 20/1 TMPyP4/DNA) and incubated in a roller at 4°C for 1 hour. The oligonucleotide pull-down reactions were performed using recombinant human PCBP2 protein (OriGene Technologies). The mixtures of magnetic beads coupled with biotinylated DNA oligos in the presence of increasing concentrations of TMPyP4 and hPCBP2 (0.3 μg/rci) were incubated in the roller at 4°C for 30 minutes, followed by three washes with 1× binding buffer. The amount of bound protein was analysed via Western blot. The protein samples and molecular weight

standards (Precision Plus Protein Standard, BioRad) were separated via SDS-PAGE in 12% resolving gels and were electro-transferred to PVDF membrane (Trans Blot Turbo mini size PVDF membrane, BioRad). The membrane was subjected to Western blot analysis using polyclonal rabbit antisera against human PCBP2 (OriGene Technologies) and goat anti-rabbit peroxidase-conjugated IgG (Jackson ImmunoResearch) as a secondary antibody. The immunoreactive bands were detected using ECL detection reagents (Clarify Western ECL substrate, BioRad).

RESULTS

Vertebrate C-rich telomeric DNA folds into an i-motif structure in acidic pH

The structural properties of two DNA fragments based on four (HS21) and eight (HS45) repeats of human (vertebrate) telomeric C-rich DNA (see Table 1) were systematically investigated as a function of decreasing pH, which is used herein as a chemical perturbation that facilitates sampling of the DNA conformational space, via NMR, CD and absorption spectroscopy and native polyacrylamide gel electrophoresis (PAGE). As flanking nucleotides (generated during construct design) might influence the structural properties of four-repeat-based telomeric DNA constructs (45,46), the construct based on the eight-repeat sequence was used as an important control. Under slightly alkaline pH (pH $\sim$ 7.3), which corresponds to the physiological pH of metazoan cells, the CD spectra for HS21 and HS45 were virtually identical, and both sequences displayed spectral shape that is typical of unstructured C-rich DNA (Figure 1A). In agreement with the CD spectra, no imino signals were detected in the 1D ^{1}H NMR spectra of HS21 or HS45, indicating the absence of hydrogen bonding-stabilised secondary structure in the investigated constructs (Figure 1B). Altogether, the NMR and CD spectra, along with the lack of cooperative transition upon regulated gradual thermal denaturation of both constructs (data not shown), clearly indicate that, under slightly alkaline conditions, both of the investigated constructs were predominantly unstructured. Upon slight acidification (pH $\sim$ 6.7) - i.e. to levels that mimic the response of the pHi of mammalian cells to mild hyperthermia - the 1D ^{1}H NMR spectra of both constructs displayed a weak signal at approximately 15.3 ppm (data not shown) originating from the proton attached to nitrogen (N3) of cytosine, a hallmark of i-motif formation (47), while the CD spectra were virtually unchanged. Upon lowering pH to 5.5, the CD spectra of both constructs displayed positive bands at 285 nm and 220 nm and a negative peak at 250 nm. In agreement with the NMR data, this pattern was indicative of i-motif formation (48). Formation of the i-motif structure was also detected based on the UV absorption spectra as the pH-dependent absorbance decrease at 260 nm and absorbance increase at 290 nm (Figure 1A, inset). Whereas the decrease in the absorbance at 260 nm reflects formation of structured DNA, the increase in the absorbance at 290 nm is due to protonation of cytosine bases. The migration patterns of HS21 and HS45 via native PAGE at pH $\sim$ 5.5 demonstrated that the respective i-motifs are monomeric (Figure 1C, Supplementary Information - Figure S1). Our observation of intramolecular i-motif formation in the HS21 construct in response to acidic pH is in agreement with previous observation of intramolecular i-motif formation from four-repeat-based constructs of human C-rich telomeric DNA (21,47–49).

C. elegans C-rich telomeric DNA forms a foldback structure in acidic pH

Two DNA fragments based on four (CE20) and eight (CE44) *C. elegans* telomeric repeats (see Table 1) were used for structural investigations of *C. elegans* (nematode) telomeric C-rich DNA. Similar to the results of the human constructs at mildly alkaline pH ($\sim$ 7.3), the CD spectra of the CE20 and CE44 fragments were virtually identical and their shape and non-cooperative melting corresponded to unstructured DNA (Figure 1D). In contrast to the human telomeric DNA, the CD and UV absorption spectra of both constructs did nearly not change with the gradual decrease in pH from 7.3 to 5.5 (Figure 1D). However, the NMR spectra displayed profound pH-dependent changes. In line with the CD results, the 1D ^{1}H NMR spectra of both constructs were devoid of imino signals under mildly alkaline conditions (Figure 1E). However, weak signals in the imino region of the 1D ^{1}H NMR spectrum between 12.5 and 13.5 ppm, clearly absent at pH $\sim$ 7.3, were detected upon mild acidification (pH $\sim$ 6.7) of the CE44 construct (data not shown). These signals became notably pronounced upon further acidification to pH $\sim$ 6.3 or pH $\sim$ 5.5. At these pH values, a similar pattern of imino signals was detected in the 1D ^{1}H NMR spectra of both the CE20 and CE44 constructs (Figure 1E). The dispersions in the chemical shifts of the imino-protons in the 1D ^{1}H NMR spectra of the CE20 and CE44 constructs (Figure 1E) suggested that both constructs adopted a similar secondary structure that was stabilised by canonical Watson-Crick (W-C) base pairs. The electrophoretic migration of both CE20 and CE44 indicated that the formed structures were monomeric (Figure 1F, Supplementary Information - Figure S1). Taken together, these data indicated that acidic pH promoted the formation of a foldback structure that is stabilised by W-C base pairs. Formation of the foldback structure provides an explanation for the virtual absence of changes in CD spectra of the CE20 and CE44 upon sample acidification (Figure 1D), since the B-type DNA structures generally displays only very small CD changes (50).

However, the data did not provide an explanation for the observed pH-dependency of the *C. elegans* C-rich telomeric DNA folding: no sign of cytosine(s) protonation, which is responsible for the pH dependence of i-motif formation in human C-rich telomeric DNA, was detected in the NMR, CD or absorption spectra. To identify the origin of the pH-dependent formation of the secondary structure of *C. elegans* C-rich telomeric DNA, we analysed the structural properties of a ^{13}C/^{15}N isotopically labelled construct (CE21) (Table 1) in detail using high-resolution NMR spectroscopy methods. (Note: The CE21 construct is an analogue of the CE20 sequence, which was extended at its 5′-end with a guanosine residue to restrict the terminal frying effect at the bottom of the hairpin stem. The response of the CE21 and CE20 to pH is identical (Supplementary Information - Figure S2)). In agreement with the properties

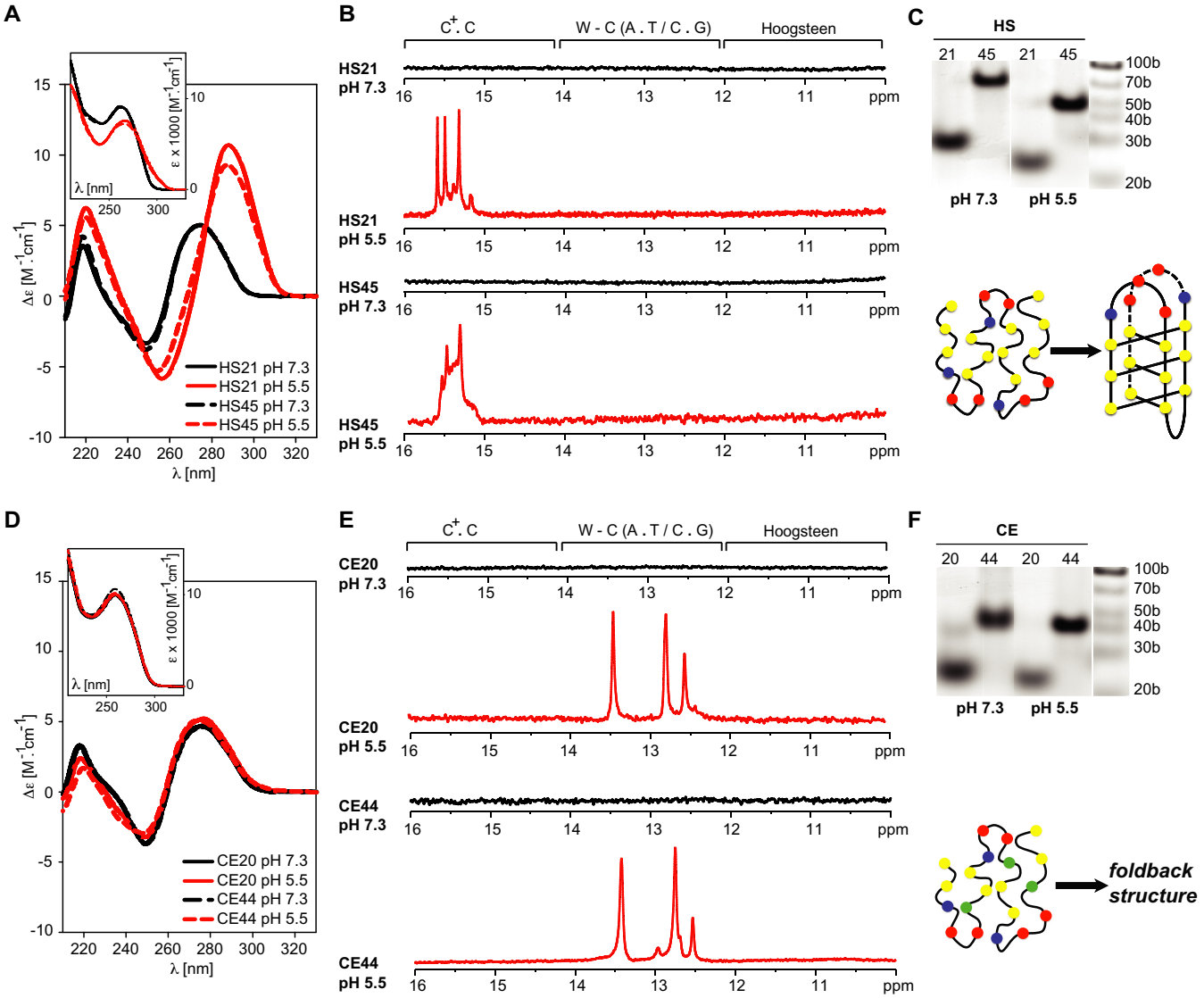

Figure 1. C-rich telomeric DNA from *C. elegans* and humans forms fundamentally different structures. LEFT: CD spectra of the DNA constructs based on human (**A**) and *C. elegans* (**D**) C-rich telomeric DNA repeats as a function of pH and the number of telomeric repeats. The spectra were acquired at 20°C. The insets show the pH-induced changes in the UV absorption spectra of the constructs. CENTRE: Imino regions of the 1D ^{1}H NMR spectra of the DNA constructs based on human (**B**) and *C. elegans* (**E**) C-rich telomeric DNA repeats as a function of pH and the number of telomeric repeats. The spectra were acquired at 20°C. Regions of the spectra typical of imino protons involved in C.C^{+}, Watson-Crick (W-C) and Hoogsteen base-pairing are indicated. RIGHT: Non-denaturing PAGE of the DNA constructs based on human (**C**) and *C. elegans* (**F**) C-rich telomeric DNA repeats as a function of pH and the number of telomeric repeats.

of the imino region of the 1D ^{1}H NMR spectrum, six pairs of cross peaks, two and four pairs at characteristic nitrogen frequencies for W-C A.T and W-C G.C base pairs, respectively, were detected in the 2D HNN-COSY spectra at pH 5.5 (Supplementary Information - Figure S3). The broad peak at ∼ 10.5 ppm, apparent in the 1D ^{1}H NMR spectrum acquired at low temperature, did not display any correlation in the HNN-COSY spectrum (Supplementary Information - Figure S3). These findings indicated that the CE21 secondary structure was stabilised by two A.T and four G/C W-C base pairs and contained one unpaired, imino-proton bearing residue. This unpaired residue was designated as thymine based on the cross peak in the 2D ^{1}H-^{1}H NOESY spectrum between the imino proton resonating at ∼ 10.5 ppm and a signal from the proton resonating at ∼1.2 ppm, a frequency that is typical of a thymine methyl group (data not shown). The comparison of the 2D ^{1}H-^{13}C HSQC spectra acquired at alkaline (pH = 7.3) and acidic pH (5.5) provided an explanation of the pH-dependent formation of the foldback motif. At alkaline pH, the aromatic region of the HSQC spectrum was characterised by a low number of (mostly overlapped) cross peaks resonating at frequencies typical of unstructured DNA (Figure 2A). At acidic pH, the aromatic region of the HSQC spectrum displayed dispersed resonances, indicative of a well-defined secondary structure (Figure 2A). Notably, four anomaly-shifted cross peaks, two at ∼142/7.9 ppm and two at ∼150/6.7 ppm in the ^{13}C/^{1}H dimension, correlating to the four C2/H2 peaks

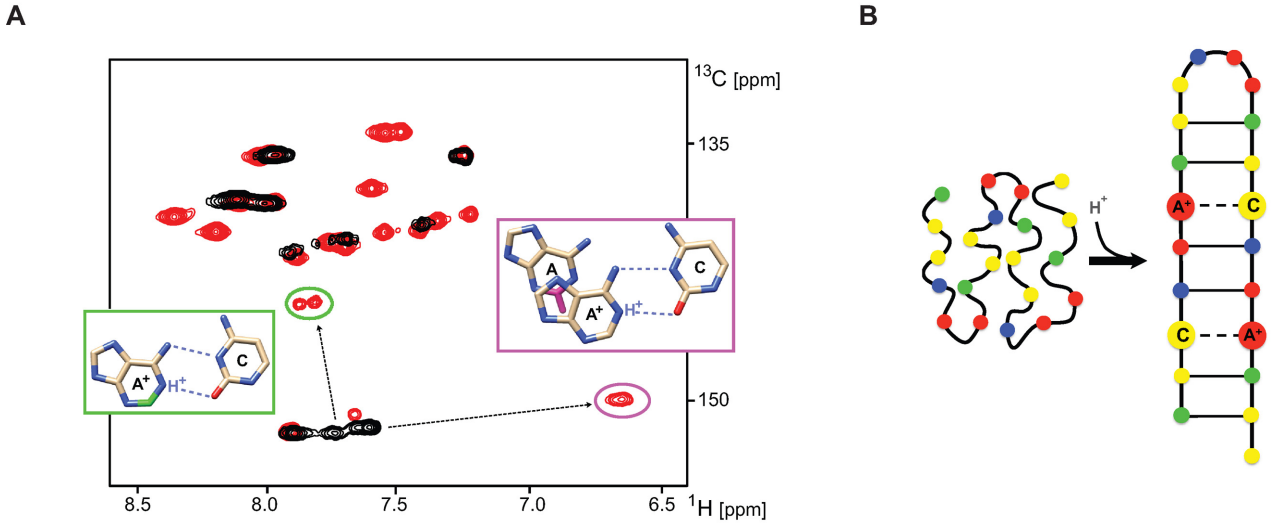

Figure 2. Formation of asymmetrical wobble A^+.C base-pair is responsible for pH dependent formation of foldback motif within *C. elegans* C-rich telomeric DNA. (**A**) Aromatic regions of the 2D ^1H-^{13}C HSQC spectra of the CE21 construct recorded at 10°C and pH = 7.3 (black) and at pH = 5.5 (red). Anomalously shifted cross-peaks (C2/H2) from (two) protonated adenine bases are highlighted in green ellipse. Two anomalously shifted cross-peaks (C2/H2) due to 5′-stacking of non-protonated adenine to the protonated adenine base are highlighted in purple ellipse. Nuclei affected by the adenosine protonation are highlighted in the figure insets: C2-H2 adjoining the protonated N1 of adenosine (green) and C2-H2 exposed to the ring current the from 5′-protonated adenosine (purple). (**B**) Schematic representation of the low pH-induced foldback motif adopted by C-rich telomeric DNA from *C. elegans*.

of adenines, were detected (Figure 2A). These anomalous shifts in the C2/H2 peaks, two at ∼ 142/7.8 ppm and two at ∼ 150/6.7 ppm, are unique spectral markers of protonated adenine at N1 and the 5′-non-protonated adenine stacked on the protonated adenine, respectively (51,52). The network of inter- and intra-residue NOEs indicated that the adenines were stabilised in the protonated form by forming asymmetric A^+.C wobble-type base pairs (52). Taken together, the NMR data at pH = 5.5 indicated that the formation of asymmetrical A^+.C non-canonical wobble base-pair(s) is responsible for the pH-dependent formation of the foldback structure. The primary structural characteristics of the CE21 construct at acidic pH that are consistent with all spectral data are schematically represented in Figure 2B.

The secondary structures within C-overhang appear to mirror the secondary structures within the G-overhang

Previous studies using short constructs based on four repeats of human telomeric DNA have demonstrated that both C- and G-rich strands can adopt intramolecular tetraplex structures (47). As shown by others for extended human telomeric G-rich sequences (53–56) (Figure 3A and B) and as demonstrated by our CD and NMR spectra (Figure 1), this property is preserved even for extended constructs of C-rich human telomeric DNA. In addition to vertebrate telomeric DNA, the simultaneous formation of tetraplex structures in both C- and G-rich strands was observed for telomeric DNA from evolutionarily distal eukaryotic species, such as plants, insects, or unicellular eukaryotes (22,57). However, whereas the *C. elegans* (nematode) G-rich telomeric DNA was previously shown to preferentially fold into a G-quadruplex structure (57), we find that the C-rich strand of the telomeric *C. elegans* DNA adopts duplex-based motif. While these observations on one hand support idea of the G-quadruplex in the G-rich strand as an evolutionarily conserved property of telomeric DNA, they argue against evolutionary conservation of the i-motif in the C-rich strand of telomeric DNA.

Considering the importance of these implications and the fact that the flanking nucleotides (particularly those at the 3′-end) might substantially influence the structural properties of four-repeat-based telomeric DNA constructs (45–46,54), we re-examined the structural properties of G-rich telomeric DNA from *C. elegans* to exclude possible interpretational bias caused by the use of a short construct of nematode G-rich telomeric DNA in the study by Tran et al. (57). Four constructs were employed to re-investigate the structural properties of telomeric G-rich DNA from *C. elegans*: (i) the original four-repeat-based construct used in the study by Tran et al. (57), referred here to as the minimal construct (gCE20—Table 1), (ii) the minimal construct containing additional cytosine at the 3′-end to assess the influence of the 3′-end flanking nucleotide (gCE21—Table 1) and (iii) two extended telomeric DNA constructs based on eight-telomeric G-rich repeats with (gCE45) or without (gCE44) a 3′-terminal cytosine (Table 1). The CD and NMR spectra of all of these constructs are shown in Figure 3C and D, respectively. The gCE20 construct displayed CD spectrum typical of an anti-parallel G-quadruplex, including positive and negative bands at 295 nm and 265 nm, respectively (Figure 3C). This spectrum, as well as its interpretation, is fully consistent with those reported by Tran et al. (57). Also the absorption thermal difference (TD) spectrum of the gCE20 (Figure 3C, inset) evidenced the G-quadruplex formation (58). However, the CD and TD spectra of the gCE20 construct were notably distinct from the TD and CD spectra of all of the extended constructs (gCE21, gCE44 and gCE45). The shape of CD spectra for the gCE21, gCE44 and gCE45 corresponded to either unstructured DNA or the B-type

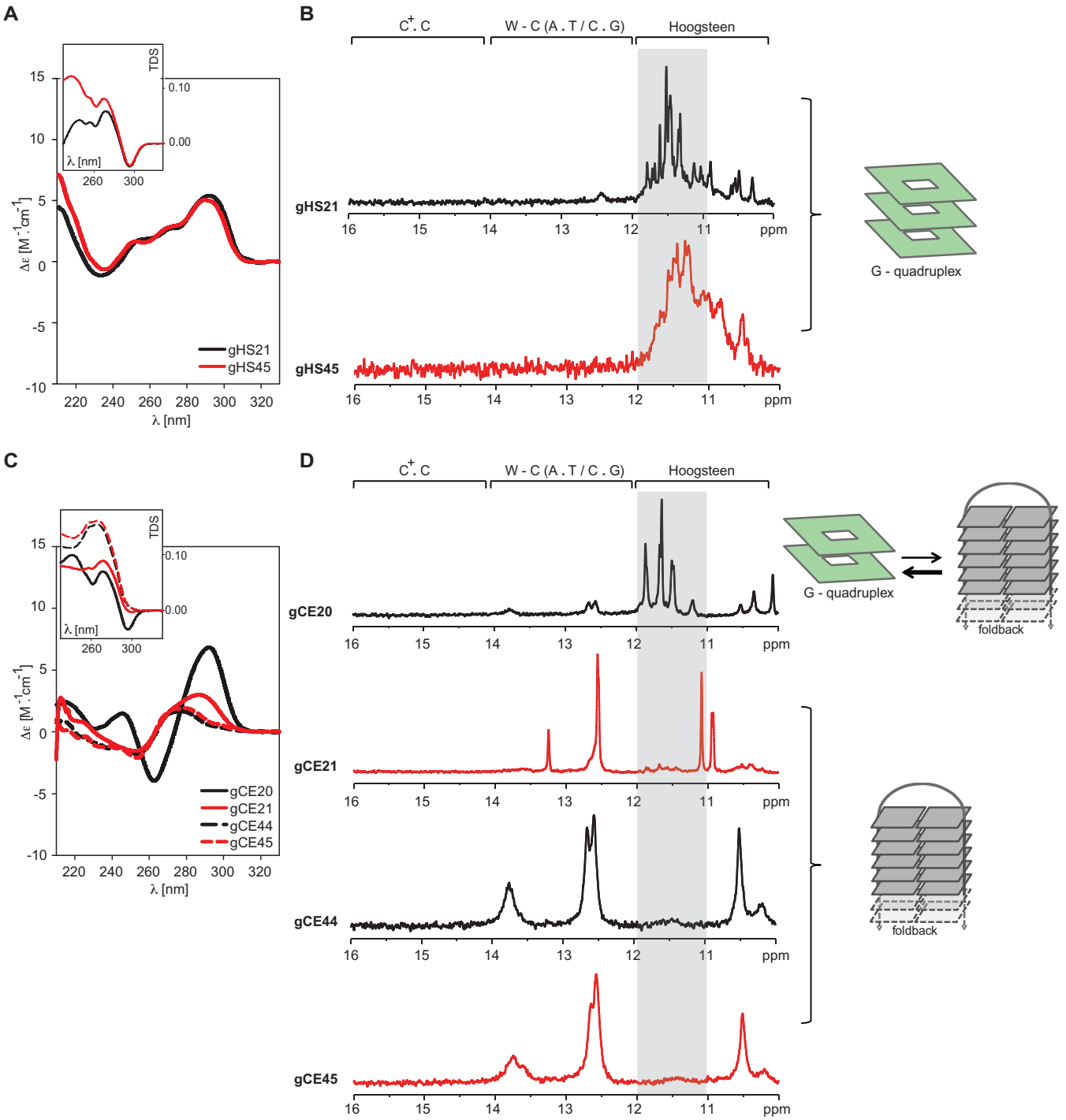

Figure 3. G-rich telomeric DNA from C. elegans and humans forms fundamentally different structures. LEFT: CD spectra of the DNA constructs based on human (**A**) and *C. elegans* (**C**) G-rich telomeric DNA repeats as a function of the number of telomeric repeats and/or the presence of the 3′-flanking residue. The insets display the thermal difference spectra (TDS) of each construct (58). CENTRE: Imino regions of the 1D ^{1}H NMR spectra of the DNA constructs based on human (**B**) and *C. elegans* (**D**) G-rich telomeric DNA repeats showing the effect of the number of telomeric repeats and the presence of the 3′-flanking residue (*C. elegans*). RIGHT: Structural interpretation of the spectra.

structure, but not a G-quadruplex (Figure 3C). Additionally, the corresponding TD spectra lacked the 295 nm negative band characteristic of the G-quadruplex structure (Figure 3C, inset). In agreement with the interpretation of the CD/TD spectra, the imino region of the 1D ^{1}H NMR spectrum of the gCE20 construct displayed characteristic signals in the 'Hoogsteen' region (11–12 ppm) (Figure 3D), indicating the formation of the G-quadruplex structure. In addition to the signals in the 'Hoogsteen region', weak signals in the 'W-C' region (12–14 ppm) were observed (Figure 3D). These signals can be attributed to minor population of 'hairpin' structure noted also in the original study by Tran et al. (57). However, the signals indicating Hoogsteen-type base pairing were clearly absent from the imino regions of the 1D ^{1}H NMR spectra of all of the extended constructs (Figure 3D). Instead, the spectra of the extended constructs (gCE44, gCE45) displayed three dominant signals in the 'W-C' region between 12 and 14 ppm. The signal at ∼ 12.8 ppm was indicative of the G.C W-C base pair(s) formation and the signals between 13.2 and 13.8 ppm indicated the formation of A.T W-C base pairs. In conclusion, our data suggest that the G-quadruplex formation in nematode telomeric G-rich DNA observed by Tran et al. (57) is restricted to the gCE20 construct. Extension of the original *C. elegans* construct (gCE20) remodelled the conformation space of the *C. elegans* telomeric DNA towards formation of the foldback structure that was stabilised by W-C base pairs.

It needs to be emphasised that the telomeric DNA from both human and *C. elegans* is based on regular telomeric repeat sequences. However, telomeric DNA in many species, predominantly from the class of *Saccharomycetes* but including some plants, slime moulds and fungi, is characterised by irregular telomeric repeats (59,60). To investigate whether secondary structures within the C-rich telomeric DNA based on irregular repeats mirror the secondary structures within the corresponding G-rich telomeric DNA, we probed structural properties of four distinct constructs from the G- (gSA-a19 and gSA-b23) and C-rich (SA-a19 and SA-b23) telomeric DNA from the *Saccharomyces cerevisiae* (Table 1) using NMR spectroscopy. At mildly alkaline pH (pH ∼ 7.3), the imino regions of the 1D ^{1}H NMR spectra of both gSA-a19 and gSA-b23 constructs displayed characteristic signals in the 'Hoogsteen' region (11–12 ppm) (Figure 4A), indicating the formation of the G-quadruplex structure. At this pH, no imino signals were detected in the 1D ^{1}H NMR spectra of both SA-a19 and SA-b23, indicating the absence of hydrogen bonding-stabilised secondary structure in the investigated constructs (Figure 4B). Yet, upon sample acidification (pH ∼ 5.5), the 1D ^{1}H NMR spectra of both constructs displayed characteristic signals in the 'i-motif' region (15–16 ppm). Altogether, our data indicate that the telomeric DNA from *S. cerevisiae* can form stable secondary structures of the same type (tetraplex) in the G- as well as in the C-rich strand. Similar to the results for the human and the *C. elegans* C-rich constructs, formation of the secondary structure within the *S. cerevisie* C-rich strand is pH dependent.

The UV melting experiments

The stabilities of the CE21 and CE45 were assessed via thermal denaturation followed by monitoring of absorbance changes at 260 nm at pH ∼ 5.5. The thermodynamic stabilities of corresponding G-rich constructs, namely gCE21 and gCE45, were assessed via thermal denaturation followed by monitoring of absorbance changes at 260 nm at pH ∼ 7.3 (Figure 5). The melting temperatures for individual constructs estimated from thermal denaturation experiments were 43.7°C, 45.6°C, 50.3°C and 52.2°C for the CE21, CE45, gCE21 and gCE45, respectively.

DISCUSSION

Identified structural properties of both *C. elegans* and human telomeric DNA are summarised in Figure 6. Their comparison with the properties of telomeric DNA from other eukaryotic species (Figure 7) provides interesting and important insights into the evolution of telomeric DNA.

In the past two decades, it has become apparent that the primary sequence of the minisatellite repeat at the functional telomere does not define the major phylogenetic groups (61). On the one hand, the human type of telomeric repeat (GGGTTA.TAACCC) is found to be scattered across all major phyla, ranging from primitive moulds to invertebrates and mammals. On the other hand, many closely related species carry distinct primary telomeric repeat sequences, and several examples of species carrying mixed arrays of telomeric repeats exist (62,63). Although it is clear that more than one type of minisatellite repeat might fulfil the function of telomeric DNA, the factors that control and drive the variability of telomeric sequences during evolution are unknown (64). Thus far, cumulative evidence has established that the primary telomeric sequence is not the primary determinant of telomere function; however, there must be some universal epigenetic hallmarks of telomeric DNA.

Previous findings regarding the formation of the i-motif and the G-quadruplex in the C- and G-rich strands from a broad range of phylogenetically distal multicellular eukaryotes have implicated the i-motif and G-quadruplex as evolutionarily conserved characteristics of telomeric DNA (65). However, the phylogenetic conservation of the i-motif, implying its functional importance, has been controversial. One argument states that acidic pH required for the i-motif formation in vitro appears generally incompatible with the pH of intracellular environment. However, it needs to be considered that other general physical mechanisms such as molecular crowding or confinement, both related to chromatine compaction state (66,67), can induce the intramolecular i-motif structure at physiological pH (68). Another (phylogenetic) argument states that the apparent conservation of the i-motif is a coincidental characteristic of the conformational space of C-rich DNA that plays no functional role and is only a consequence of the propagation of sequence information from the complementary evolutionarily conserved G-rich (and G-quadruplex-forming) strand.

Our findings of the foldback motif formation within the *C. elegans* telomeric C-rich DNA clearly argue against the phylogenetic conservation of the i-motif. However, on

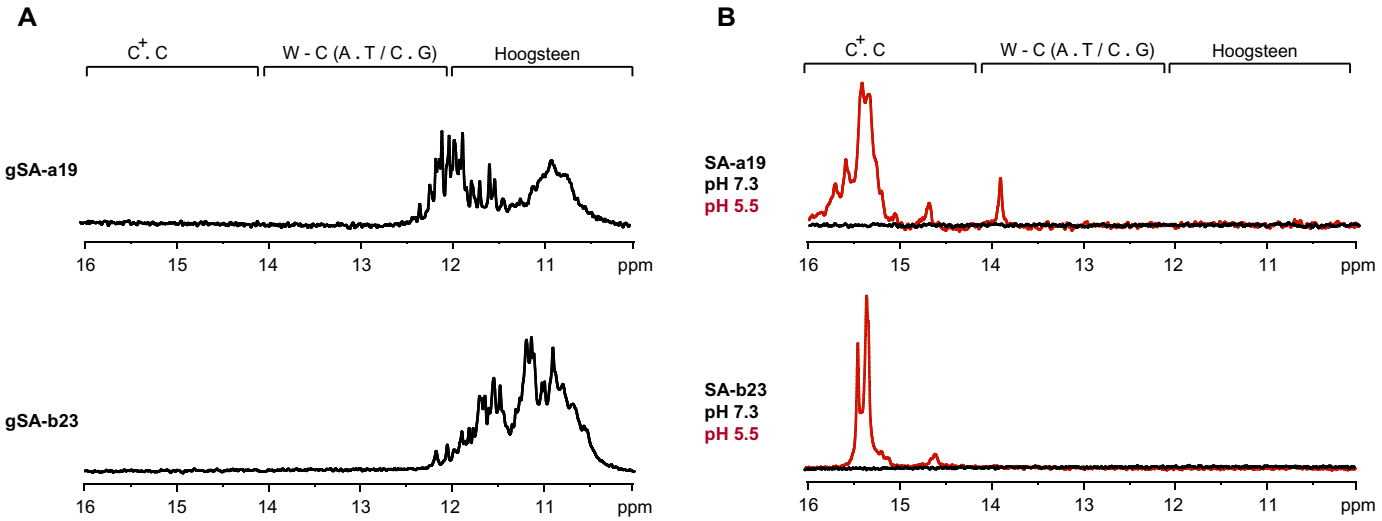

Figure 4. (**A**) Imino regions of the 1D ^{1}H NMR spectra of the DNA constructs based on *S. cerevisiae* G-rich telomeric DNA repeats (gSA-a19 and gSA-b23) at pH 7.3 (**B**) Imino regions of the 1D ^{1}H NMR spectra of the DNA constructs based on *S. cerevisiae* C-rich telomeric DNA repeats (SA-a19 and SA-b23) at pH 7.3 (black) and pH 5.5 (red). The spectra were acquired at 700 MHz at 20°C.

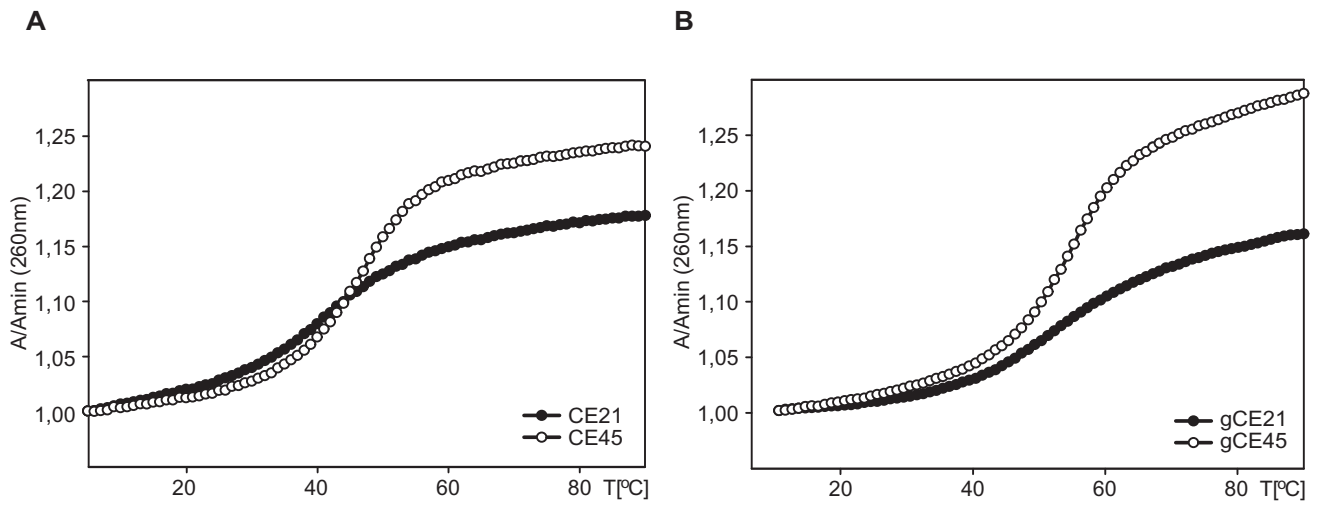

Figure 5. Thermal melting curves recorded at 260 nm and at 0.02 mM oligonucleotide strand concentration: (**A**) CE21 and CE45; (**B**) gCE21 and gCE45. Thermal melting curves for the C- and G-rich sequences were acquired at pH = 5.5 and pH = 7.3, respectively. The individual constructs are labelled according Table 1.

the other hand, our findings identify a microenvironment-dependent secondary structure formation, regardless of an actual structural type, as a plausible conserved property of the telomeric C-rich strand indicating its functional role independent of the G-rich counterpart.

Analogously, the detection of the foldback motif in the G-rich strand of *C. elegans* telomeric DNA refute the G-quadruplex formation as an evolutionarily conserved property of the G-rich strand. Our data rather suggest that the ability to form thermodynamically stable secondary structure, regardless of an actual structural type, is the property of the G-rich strand of telomeric DNA that is conserved in the course of evolution.

Noteworthy, we observe that the melting temperatures of the short, CE21 (43.7°C) and gCE21 (50.3°C), and the corresponding extended, CE44 (45.6°C) and gCE44 (52.2°C), constructs from *C. elegans* telomeric DNA are remarkably similar with differences being generally less than 2°C. (Figure 5). This similarity indicates that both G- and C-rich telomeric DNA might consist of four-repeat based foldback motifs organised in 'bead-on-a-string' like fashion, i.e. in arrangement analogous to that observed in extended human telomeric G-rich DNA (54–55,69), rather than forming long continuous hairpins.

Altogether, our data are indicative that the properties of the telomeric DNA are evolutionarily conserved but not at the primary sequence level. Changes in the telomeric repeat sequence are clearly tolerated during evolution; however, these changes appear to be 'neutral' (70) with respect to the specific structural properties of both the G- and C-rich strand (Figure 7), namely: i) ability to form stable secondary structures of the same type in the G- as well as in

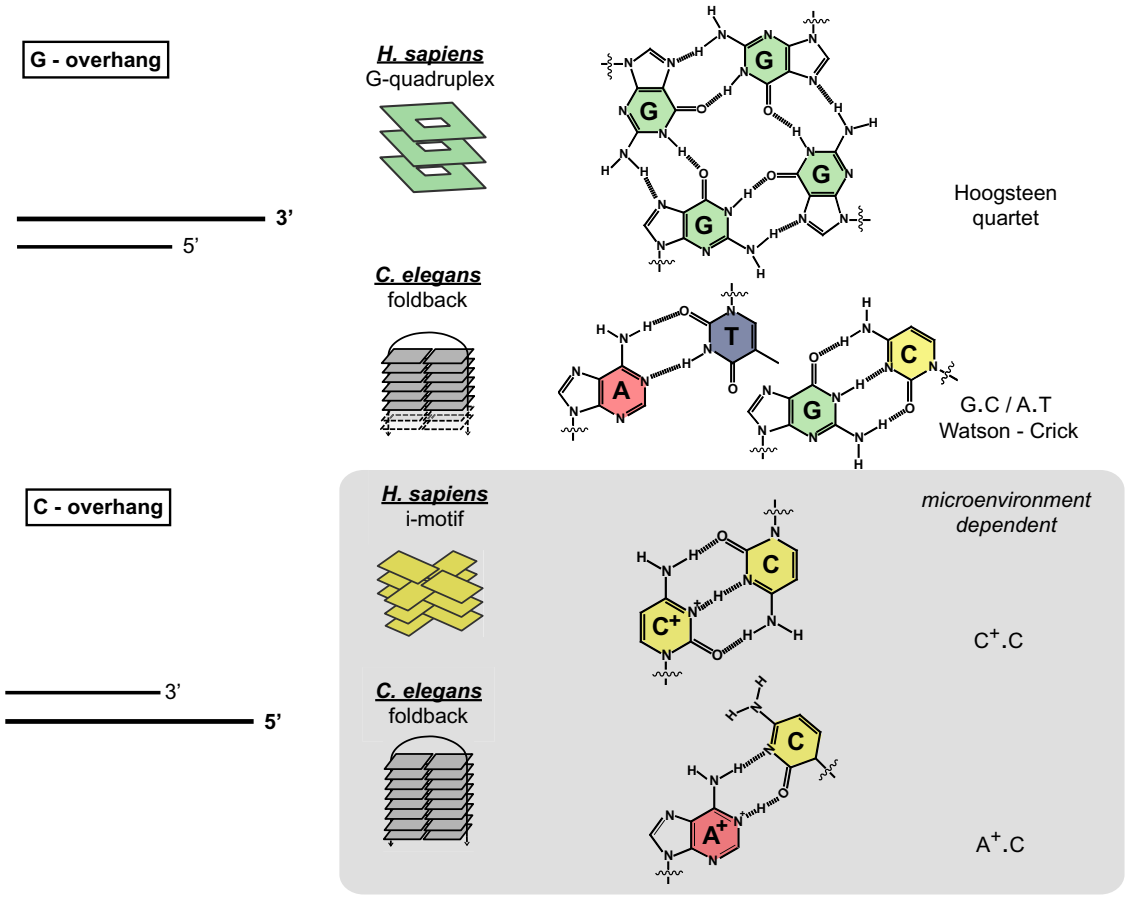

Figure 6. Schematic representations of the structural differences of *C. elegans* and human 5′- and 3′- telomeric C- and G-overhangs. *C. elegans* telomeric repeats from C- and G-rich strands form simple foldback motifs. Whereas the foldback motif formation in the G-rich strand relies on the formation of canonical Watson-Crick (G.C and A.T) base pairs, the formation of the foldback structure in the C-rich strand relies on the environmentally induced formation of non-canonical wobble A^+.C pair(s). Human telomeric repeats in the G-rich strand form a G-quadruplex structure relying on the formation of so-called G-quartet stabilised by a Hoogsteen-type network of hydrogen bonds. Human C-rich telomeric DNA forms an intercalated tetraplex (i-motif), which is stabilised by non-canonical wobble $C.C^+$ base pairs.

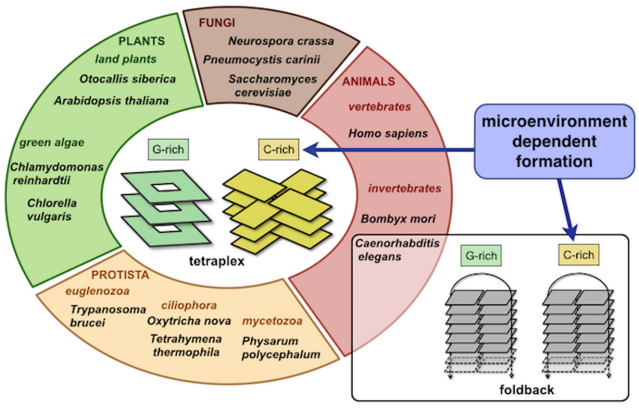

Figure 7. The structural characteristics of the *C. elegans* telomeric DNA are distinct from those of other eukaryotes. Changes in the telomeric repeat primary sequence (Supplementary Information—Table S1) appear to be 'neutral' with respect to the specific structural properties of both the G- and C-rich strand, namely: i) ability to form stable secondary structures of the same type in the G- as well as in the C-rich strand and ii) microenvironment dependent formation of secondary structure within the C-rich strand.

the C-rich strand and ii) microenvironment dependent formation of secondary structure within the C-rich strand.

Importantly, the conservation of a secondary structure formation within the C-rich strand of human telomeric DNA implies the i-motif importance for the telomere function. In this respect, a logical question arises—can targeting of the 5′ C-overhang, a characteristic of ALT tumours, be employed for the selective elimination of tumour cells engaged in the ALT pathway? Although addressing this question is beyond the scope of this study, several lines of indirect evidence provide support for the concept of the i-motif as a negative regulator of telomere maintenance and, thus, the potential for the human telomeric i-motif as a therapeutic target. Recently, Chen et al. (29) showed that treatment of K562 and HeLa cells with single walled carbon nanotube (SWCN), a selective stabiliser of human telomeric i-motif, leads to telomere dysfunction via telomere uncapping marked by the displacement of several telomere binding proteins including the poly-C binding protein (PCBP) that has been identified as a specific telomere C-rich strand binder (71,72). Along the same line, Fujimori et al. (73) showed that treatment of telomerase-negative,

ALT-positive U2OS cells with the a small molecular weight cationic porphyrin TMPyP4, known stabiliser of both human telomeric i-motif (74) and human (telomeric) DNA G-quadruplex (75,76), resulted in inhibition of cell growth and elevated levels of apoptosis supposedly due to telomere dysfunction. Unfortunately, complexity of the reported cellular phenotypes did not allow identification of primary events responsible for the telomere dysfunction. We observed that the TMPyP4 could directly interfere with the binding of the PCBP to human telomeric C-rich DNA in in-vitro assay (Supplementary Information—Figure S4). This observation provides support for explanation by Chen et al. (29) that displacement of PCBP binding/shielding protein might indeed be due to the SWCN induced i-motif formation within the telomeric DNA. However, it needs to be stressed that all available evidence is only indirect, and detailed studies will be needed to validate this concept.

Finally, *C. elegans* mutants that can survive in the absence of a functional telomerase pathway and that fully rely on the ALT mechanism were recently proposed as a multi-cellular model organism for ALT (30,77–78). The telomeric DNA in this model organism appears to display all of the important characteristics underlying ALT-maintained telomeres, including the presence of the 5′-C-rich overhang. Although this organism appears to represent an outstanding tool for elucidating the mechanisms underlying the ALT pathway, our structural data implicating distinct types of secondary structures in both 3′-G-rich and 5′-C-rich overhangs between *C. elegans* and humans emphasises the important limitation of *C. elegans* as an ALT model. This model organism should be used with caution in situations involving the use of ligands targeting telomeric DNA.

It is important to stress that our conclusions presented here were drawn based on the properties of model oligonucleotides in vitro. The conformational properties of telomeric repeats in the context of genomic DNA in vivo might differ from those of isolated oligonucleotides employed in our study.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Czech Science Foundation [13–28310S to L.T. and M.V., P205/12/0466 to M.V.]; CEITEC [CZ.1.05/1.1.00/02.0068]. L.T. was supported by the R&D grant from Instruct, career development grant from the European Organization for Molecular Biology [IG2535] and by the Marie-Curie Re-integration grant. S.F.T. was supported from the SoMoPro II Programme, co-financed by the European Union and the South Moravian Region (Czech Republic). We thank Michaela Krafcikova for help with NMR spectra acquisition. Funding for open access charge: Czech Science Foundation [13-28310S].

Author Contributions: P.S. and K.B. conducted the CD/UV and native PAGE experiments; L.T., R.F. and S.F.T. conducted the NMR experiments; S.F.T. conducted the pull-down assays; P.S., K.B., R.F., S.F.T., L.T. and M.V. designed the experiments; and M.V. and L.T. wrote the manuscript.

Conflict of interest statement. None declared.

REFERENCES

1. Wright,W.E., Tesmer,V.M., Huffman,K.E., Levene,S.D. and Shay,J.W. (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.*, **11**, 2801–2809.
2. Makarov,V.L., Hirose,Y. and Langmore,J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657–666.
3. Griffith,J.D., Comeau,L., Rosenfield,S., Stansel,R.M., Bianchi,A., Moss,H. and de Lange,T. (1999) Mammalian telomeres end in a large duplex loop. *Cell*, **97**, 503–514.
4. Allsopp,R.C., Vaziri,H., Patterson,C., Goldstein,S., Younglai,E.V., Futcher,A.B., Greider,C.W. and Harley,C.B. (1992) Telomere length predicts replicative capacity of human fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 10114–10118.
5. Engelhardt,M. and Martens,U.M. (1998) The implication of telomerase activity and telomere stability for replicative aging and cellular immortality (Review). *Oncol. Rep.*, **5**, 1043–1052.
6. Aubert,G. and Lansdorp,P.M. (2008) Telomeres and aging. *Physiol. Rev.*, **88**, 557–579.
7. Shore,D. and Bianchi,A. (2009) Telomere length regulation: coupling DNA end processing to feedback regulation of telomerase. *EMBO J.*, **28**, 2309–2322.
8. Greider,C.W. and Blackburn,E.H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, **43**, 405–413.
9. Hiyama,E. and Hiyama,K. (2007) Telomere and telomerase in stem cells. *Br. J. Cancer*, **96**, 1020–1024.
10. Kim,N.W., Piatyszek,M.A., Prowse,K.R., Harley,C.B., West,M.D., Ho,P.L., Coviello,G.M., Wright,W.E., Weinrich,S.L. and Shay,J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–2015.
11. Lipps,H.J. and Rhodes,D. (2009) G-quadruplex structures: in vivo evidence and function. *Trends Cell. Biol.*, **19**, 414–422.
12. Phan,A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.*, **277**, 1107–1117.
13. Hansel,R., Foldynova-Trantirkova,S., Lohr,F., Buck,J., Bongartz,E., Bamberg,E., Schwalbe,H., Dotsch,V. and Trantirek,L. (2009) Evaluation of parameters critical for observing nucleic acids inside living Xenopus laevis oocytes by in-cell NMR spectroscopy. *J. Am. Chem. Soc.*, **131**, 15761–15768.
14. 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.
15. Zahler,A.M., Williamson,J.R., Cech,T.R. and Prescott,D.M. (1991) Inhibition of telomerase by G-quartet DNA structures. *Nature*, **350**, 718–720.
16. De Cian,A., Lacroix,L., Douarre,C., Temime-Smaali,N., Trentesaux,C., Riou,J.F. and Mergny,J.L. (2008) Targeting telomeres and telomerase. *Biochimie*, **90**, 131–155.
17. Bryan,T.M., Englezou,A., Dalla-Pozza,L., Dunham,M.A. and Reddel,R.R. (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat. Med.*, **3**, 1271–1274.
18. Ogino,H., Nakabayashi,K., Suzuki,M., Takahashi,E., Fujii,M., Suzuki,T. and Ayusawa,D. (1998) Release of telomeric DNA from chromosomes in immortal human cells lacking telomerase activity. *Biochem. Biophys. Res. Commun.*, **248**, 223–227.
19. Cesare,A.J. and Reddel,R.R. (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat. Rev. Genet.*, **11**, 319–330.
20. Oganesian,L. and Karlseder,J. (2011) Mammalian 5′ C-rich telomeric overhangs are a mark of recombination-dependent telomere maintenance. *Mol. Cell*, **42**, 224–236.
21. Ahmed,S., Kintanar,A. and Henderson,E. (1994) Human telomeric C-strand tetraplexes. *Nat. Struct. Biol.*, **1**, 83–88.
22. Leroy,J.L., Gueron,M., Mergny,J.L. and Helene,C. (1994) Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res.*, **22**, 1600–1606.
23. Gehring,K., Leroy,J.L. and Gueron,M. (1993) A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature*, **363**, 561–565.

24. Lieblein,A.L., Kramer,M., Dreuw,A., Furtig,B. and Schwalbe,H. (2012) The nature of hydrogen bonds in cytidine…H+…cytidine DNA base pairs. *Angew. Chem.* **51**, 4067–4070.
25. Zhou,J., Wei,C., Jia,G., Wang,X., Feng,Z. and Li,C. (2010) Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst.*, **6**, 580–586.
26. Raices,M., Verdun,R.E., Compton,S.A., Haggblom,C.I., Griffith,J.D., Dillin,A. and Karlseder,J. (2008) C. elegans telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell*, **132**, 745–757.
27. Madshus,I.H. (1988) Regulation of intracellular pH in eukaryotic cells. *Biochem. J.*, **250**, 1–8.
28. Wadsworth,W.G. and Riddle,D.L. (1988) Acidic intracellular pH shift during Caenorhabditis elegans larval development. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 8435–8438.
29. Chen,Y., Qu,K., Zhao,C., Wu,L., Ren,J., Wang,J. and Qu,X. (2012) Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.*, **3**, 1074–1088.
30. Lackner,D.H. and Karlseder,J. (2013) C. elegans survivors without telomerase. *Worm*, **2**, e21073.
31. Shampay,J., Szostak,J.W. and Blackburn,E.H. (1984) DNA sequences of telomeres maintained in yeast. *Nature*, **310**, 154–157.
32. Gray,D.M., Hung,S.H. and Johnson,K.H. (1995) Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes. *Methods Enzymol.*, **246**, 19–34.
33. Masse,J.E., Bortmann,P., Dieckmann,T. and Feigon,J. (1998) Simple, efficient protocol for enzymatic synthesis of uniformly 13C, 15N-labeled DNA for heteronuclear NMR studies. *Nucleic Acids Res.*, **26**, 2618–2624.
34. Fiala,R., Spackova,N., Foldynova-Trantirkova,S., Sponer,J., Sklenar,V. and Trantirek,L. (2011) NMR cross-correlated relaxation rates reveal ion coordination sites in DNA. *J. Am. Chem. Soc.*, **133**, 13790–13793.
35. Piotto,M., Saudek,V. and Sklenar,V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, **2**, 661–665.
36. Vuister,G.W. and Bax,A. (1992) Measurement of two-bond JCOH alpha coupling constants in proteins uniformly enriched with ^{13}C. *J. Biomol. NMR*, **2**, 401–405.
37. Santoro,J. and King,G.C. (1992) A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J. Mag. Reson.*, **97**, 202–207.
38. Bohlen,J.M. and Bodenhausen,G. (1993) Experimental aspects of Chirp NMR spectroscopy. *J. Magn. Reson.*, **102**, 293–301.
39. Sklenar,V., Peterson,R.D., Rejante,M.R. and Feigon,J. (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in ^{15}N,^{13}C-labeled RNA oligonucleotides. *J. Biomol. NMR*, **3**, 721–727.
40. Lippens,G., Dhalluin,C. and Wieruszeski,J.M. (1995) Use of a water flip-back pulse in the homonuclear NOESY experiment. *J. Biomol. NMR*, **5**, 327–331.
41. Legault,P., Farmer,B.T., Mueller,L. and Pardi,A. (1994) Through-bond correlation of adenine protons in a C-13-labeled ribozyme. *J. Am. Chem. Soc.*, **116**, 2203–2204.
42. Marino,J.P., Prestegard,J.H. and Crothers,D.M. (1994) Correlation of adenine H2/H8 resonances in uniformly C-13 labeled RNAs by 2D HCCH-TOCSY - A new tool for H-1 assignment. *J. Am. Chem. Soc.*, **116**, 2205–2206.
43. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
44. Pervushin,K., Ono,A., Fernandez,C., Szyperski,T., Kainosho,M. and Wuthrich,K. (1998) NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 14147–14151.
45. Dai,J., Carver,M. and Yang,D. (2008) Polymorphism of human telomeric quadruplex structures. *Biochimie*, **90**, 1172–1183.
46. Vorlickova,M., Tomasko,M., Sagi,A.J., Bednarova,K. and Sagi,J. (2012) 8-oxoguanine in a quadruplex of the human telomere DNA sequence. *FEBS J.*, **279**, 29–39.
47. Phan,A.T. and Mergny,J.L. (2002) Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. *Nucleic Acids Res.*, **30**, 4618–4625.
48. Simonsson,T., Pribylova,M. and Vorlickova,M. (2000) A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Biophys. Res. Commun.*, **278**, 158–166.
49. Phan,A.T. and Leroy,J.L. (2000) Intramolecular i-Motif Structures of Telomeric DNA. *J. Biomol. Struct. Dyn.*, **17**(Suppl. 1), 245–251.
50. Kypr,J., Kejnovska,I., Renciuk,D. and Vorlickova,M. (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.*, **37**, 1713–1725.
51. Legault,P. and Pardi,A. (1994) In-situ probing of adenine protonation in RNA by C-13 NMR. *J. Am. Chem. Soc.*, **116**, 8390–8391.
52. Stefl,R., Oberstrass,F.C., Hood,J.L., Jourdan,M., Zimmermann,M., Skrisovska,L., Maris,C., Peng,L., Hofr,C., Emeson,R.B. *et al.* (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, **143**, 225–237.
53. Azarkh,M., Singh,V., Okle,O., Dietrich,D.R., Hartig,J.S. and Drescher,M. (2012) Intracellular conformations of human telomeric quadruplexes studied by electron paramagnetic resonance spectroscopy. *ChemPhysChem*, **13**, 1444–1447.
54. Hansel,R., Lohr,F., Trantirek,L. and Dotsch,V. (2013) High-resolution insight into G-overhang architecture. *J. Am. Chem. Soc.*, **135**, 2816–2824.
55. Vorlickova,M., Chladkova,J., Kejnovska,I., Fialova,M. and Kypr,J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res.*, **33**, 5851–5860.
56. Petraccone,L., Spink,C., Trent,J.O., Garbett,N.C., Mekmaysy,C.S., Giancola,C. and Chaires,J.B. (2011) Structure and stability of higher-order human telomeric quadruplexes. *J. Am. Chem. Soc.*, **133**, 20951–20961.
57. Tran,P.L., Mergny,J.L. and Alberti,P. (2011) Stability of telomeric G-quadruplexes. *Nucleic Acids Res.*, **39**, 3282–3294.
58. Mergny,J.L., De Cian,A., Ghelab,A., Sacca,B. and Lacroix,L. (2005) Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.*, **33**, 81–94.
59. Cohn,M. (2008) Molecular Diversity of Telomeric Sequences. In: Nosek,J and Tomaska,L (eds). *Origin and Evolution of Telomeres*. Landes Bioscience, pp. 70–82.
60. Gunisova,S., Elboher,E., Nosek,J., Gorkovoy,V., Brown,Y., Lucier,J.F., Laterreur,N., Wellinger,R.J., Tzfati,Y. and Tomaska,L. (2009) Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements. *RNA*, **15**, 546–559.
61. Fajkus,J., Sykorova,E. and Leitch,A.R. (2005) Telomeres in evolution and evolution of telomeres. *Chromosome Res.*, **13**, 469–479.
62. Sykorova,E., Lim,K.Y., Kunicka,Z., Chase,M.W., Bennett,M.D., Fajkus,J. and Leitch,A.R. (2003) Telomere variability in the monocotyledonous plant order Asparagales. *Proc. Biol. Sci.*, **270**, 1893–1904.
63. Rotkova,G., Sklenickova,M., Dvorackova,M., Sykorova,E., Leitch,A.R. and Fajkus,J. (2004) An evolutionary change in telomere sequence motif within the plant section Asparagales had significance for telomere nucleoprotein complexes. *Cytogenet Genome Res.*, **107**, 132–138.
64. Tomaska,L. and Nosek,J. (2009) Telomere heterogeneity: taking advantage of stochastic events. *FEBS Lett.*, **583**, 1067–1071.
65. Garavis,M., Gonzalez,C. and Villasante,A. (2013) On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution. *Genome Biol. Evol.*, **5**, 1142–1150.
66. Walter,A., Chapuis,C., Huet,S. and Ellenberg,J. (2013) Crowded chromatin is not sufficient for heterochromatin formation and not required for its maintenance. *J. Struct. Biol.*, **184**, 445–453.
67. Bancaud,A., Huet,S., Daigle,N., Mozziconacci,J., Beaudouin,J. and Ellenberg,J. (2009) Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J.*, **28**, 3785–3798.
68. Rajendran,A., Nakano,S. and Sugimoto,N. (2010) Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun. (Camb.)*, **46**, 1299–1301.

69. Yu,H.Q., Miyoshi,D. and Sugimoto,N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc.*, **128**, 15461–15468.
70. Gray,M.W., Lukes,J., Archibald,J.M., Keeling,P.J. and Doolittle,W.F. (2010) Cell biology. Irremediable complexity? *Science*, **330**, 920–921.
71. Bandiera,A., Tell,G., Marsich,E., Scaloni,A., Pocsfalvi,G., Akintunde Akindahunsi,A., Cesaratto,L. and Manzini,G. (2003) Cytosine-block telomeric type DNA-binding activity of hnRNP proteins from human cell lines. *Arch. Biochem. Biophys.*, **409**, 305–314.
72. Fenn,S., Du,Z., Lee,J.K., Tjhen,R., Stroud,R.M. and James,T.L. (2007) Crystal structure of the third KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.6 A resolution. *Nucleic Acids Res.*, **35**, 2651–2660.
73. Fujimori,J., Matsuo,T., Shimose,S., Kubo,T., Ishikawa,M., Yasunaga,Y. and Ochi,M. (2011) Antitumor effects of telomerase inhibitor TMPyP4 in osteosarcoma cell lines. *J. Orthop. Res.*, **29**, 1707–1711.
74. Fedoroff,O.Y., Rangan,A., Chemeris,V.V. and Hurley,L.H. (2000) Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry*, **39**, 15083–15090.
75. 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.
76. Wheelhouse,R.T., Sun,D.K., Han,H.Y., Han,F.X.G. and Hurley,L.H. (1998) Cationic porphyrins as telomerase inhibitors: the interaction of tetra-(N-methyl-4-pyridyl)porphine with quadruplex DNA. *J. Am. Chem. Soc.*, **120**, 3261–3262.
77. Lackner,D.H., Raices,M., Maruyama,H., Haggblom,C. and Karlseder,J. (2012) Organismal propagation in the absence of a functional telomerase pathway in *Caenorhabditis elegans*. *EMBO J.*, **31**, 2024–2033.
78. Cheng,C., Shtessel,L., Brady,M.M. and Ahmed,S. (2012) *Caenorhabditis elegans* POT-2 telomere protein represses a mode of alternative lengthening of telomeres with normal telomere lengths. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 7805–7810.