Biological Structures

DOI: 10.1002/anie.200903858

Consecutive Formation of G-Quadruplexes in Human Telomeric-Overhang DNA: A Protective Capping Structure for Telomere Ends**

Yan Xu, Takumi Ishizuka, Kaori Kurabayashi, and Makoto Komiyama**

Telomere structures are essential for genome stability.[1] Telomere dysfunction can cause cell senescence, death, or genomic instability.[1] Human telomeric DNA consists of a duplex region composed of TTAGGG repeats, ending in a 100–200 nucleotide (nt) G-rich single-stranded overhang.[1] The telomeric overhang is recognized as a critical component of telomere structures, required for telomere end protection and extension.[1] It is also particularly attractive as a target for anticancer therapeutic development.[2] The detailed structural features and mechanisms of telomeric overhang that are responsible for protecting the chromosome ends are still insufficiently understood, although several alternative structures have been proposed for human telomeric DNA, such as G-quadruplex or T-loop.[3,4]

The T-loop and G-quadruplex structural models with the same human telomere sequence seem to leave many questions unanswered. Some important experimental findings are not adequately explained by these models. Among those findings are the following: 1) The T-loop structure, which is stabilized by the 3′-overhang strand encroaching into the double-stranded region, cannot explain the inhibition of telomerase activity by a G-quadruplex stabilizing molecule.[2] 2) Previous studies showed that only one TTAGGG repeat (6 nt) of 3′-overhang is sufficient for T-loop formation.[5] However, a recent study showed that, in most senescent cells, telomeric overhangs are still long, often averaging 50–100 nt.[6] This finding raises the question of why that DNA length, seemingly sufficient for T-loop formation, does not provide a protective structure for prevention of senescence in cells. For the G-quadruplex structure, most previous studies focused on individual G-quadruplexes formed by short telomeric DNA (especially 22 nt unimolecular G-quadruplexes).[3] However, it might be more biologically relevant to take into account the real length of the 3′-terminal single-stranded overhang of human telomeric DNA in vivo (100–200 nt) to directly reveal the structural features of long-telomeric-overhang DNA, owing to its similarity in length to that of the intracellular telomere.

Several research groups, including our own, have mentioned the possibility that long-telomeric-overhang DNA may form a higher-order G-quadruplex structure (Figure 1).[3] However, experimental data to directly demonstrate the existence of such a structure have not been reported to date. Presumably, higher-order telomeric DNA structures with relatively large molecular sizes are not amenable to study by traditional methods such as NMR spectroscopy and crystallography.

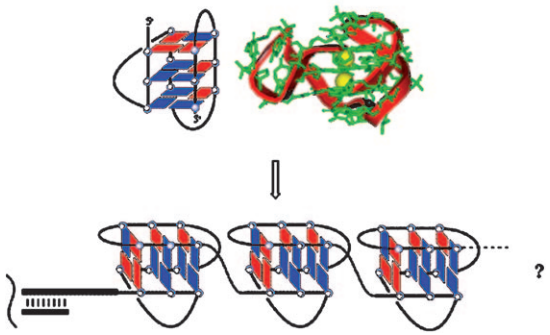

Figure 1. Schematic drawing showing that the mixed-type structure (top) reported in previous studies might fold to form a higher-order telomeric G-quadruplex structure in telomeric-overhang DNA (bottom).

Atomic force microscopy (AFM) is a powerful and widely used imaging technique for visualizing molecular structure at the single-molecule level.[7] Herein, we use a combination of AFM, fluorescent resonance energy transfer (FRET), and circular dichroism (CD) experiments to investigate the structural characteristics of long-telomeric-overhang DNA. We found that telomeric-overhang DNA forms a higher-order DNA structure containing consecutive G-quadruplexes, and that this structure protects DNA double-strand ends from being recognized as double-strand breaks and directs against nuclease hydrolysis. This structure thus addresses the puzzling questions left unanswered by the T-loop or individual G-quadruplex structural models. These results provide valuable information to promote understanding of the structure and function of human single-stranded telomeric-overhang DNA.

To gain structural information on human single-stranded telomeric-overhang DNA, we visualized the elongated telomere sequence by atomic force microscopy (AFM). For this purpose, we designed several DNA samples. DNA 1 and 2 are shown in Figure 2a. DNA 1 has a 96 nt human telomere

[*] Prof. Y. Xu, T. Ishizuka, K. Kurabayashi, Prof. M. Komiyama
Research Center for Advanced Science and Technology
The University of Tokyo, Komaba
Meguro-ku, Tokyo 153-8904 (Japan)
Fax: (+81) 3-5254-5201
E-mail: xuyan@mkomi.rcast.u-tokyo.ac.jp
komiyama@mkomi.rcast.u-tokyo.ac.jp

[**] This work was partially supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, Sports, Culture, and Technology (Japan) (18001001). Support by the Global COE Program for Chemistry Innovation through Cooperation of Science and Engineering is also acknowledged.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200903858.

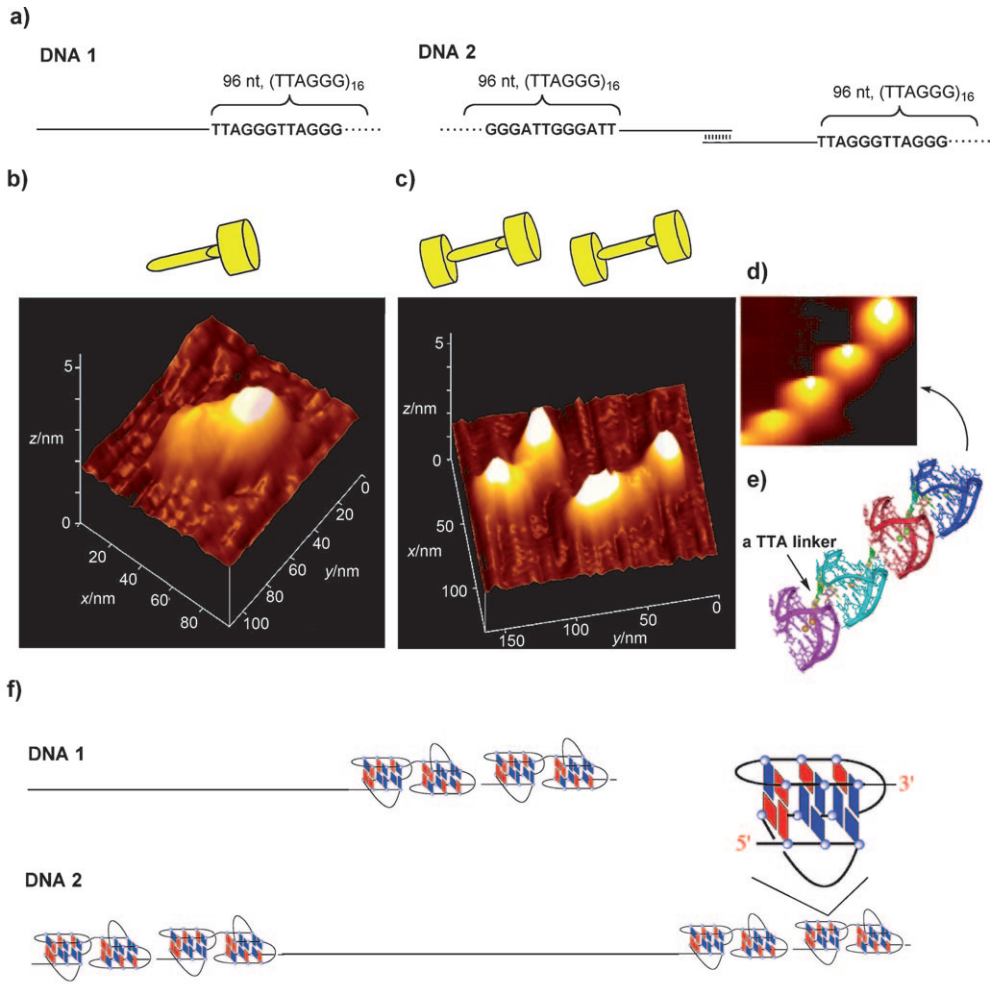

Figure 2. Human single-stranded telomeric-overhang DNA folds into a higher-order G-quadruplex structure by consecutive formation of G-quadruplexes. a) Schematic drawings of DNA models used for AFM visualization. DNA 1 has a TTAGGG repeat sequence (96 nt) at one end and a single-stranded DNA segment (120 nt). DNA 2 has a TTAGGG repeat sequence (96 nt) at both ends and a single-stranded DNA segment (240 nt) linked by a 15 nt duplex. b) AFM image of DNA 1 and schematic drawing of its half-dumbbell structure. c) High-resolution image of DNA 2 and schematic drawing of its dumbbell-shaped structure. d) High-resolution AFM image of higher-order telomeric DNA structures. e) Overall view of a possible model for the higher-order telomeric DNA structure; four G-quadruplex units shown in different colors are linked by added TTA linkers. f) Schematic drawing of higher-order telomeric G-quadruplex structures formed by DNA 1 and DNA 2. The 5′ and 3′ ends of each mixed-type G-quadruplex point in opposite directions.

TTAGGG repeat sequence at one end. DNA 2 has the same sequence but at both ends. The AFM image of DNA 1 clearly shows a half-dumbbell-shaped structure with a globular protrusion at the end (Figure 2b). The height of the end protrusion is four times the height of the single-stranded DNA, suggesting that the end was formed by G-quadruplexes. A higher-resolution AFM image shows the G-quadruplex molecular structure at the protrusion ends. The image clearly shows that the higher-order G-quadruplex consists of blob-shaped protrusions arranged end-to-end and formed by adjacent G-quadruplex units (Figure 2d). AFM measurements show that the superhelix structure has a mountain-shaped profile (heights averaged over 30 images, Figure S1 in the Supporting Information). The peaks and valleys relate directly to the difference in height between a G-quadruplex unit and a single-stranded TTA linker that connects two adjacent G-quadruplex units (Figure 1, Figure 2e). Four blobs, corresponding to four peaks in the height profile, are consistent with the fact that a 96 nt TTAGGG repeat sequence can form four consecutive G-quadruplex units, in which a minimum length of 21 nucleotides is required to form a stable intramolecular G-quadruplex unit.[3]

The AFM image of DNA 2, which has the human telomere TTAGGG-repeats sequence at both ends, shows a dumbbell-shaped structure at the base of a globular end protrusion (Figure 2a and Figure S2 in the Supporting Information). The higher-resolution AFM image clearly shows a dumbbell-shaped structure with clear bright blobs at the ends (Figure 2c). G-quadruplex structures can be recognized at the ends of both DNA 1 and DNA 2 as a protrusion dot with a height of 2–3 nm (averaged over 28 images; Figure 2b,c), which is consistent with the reported values for G-quadruplex determined by AFM and close to the diameter of G-quartets (2.4 nm) determined by X-ray crystallography.[8] In independent AFM experiments, higher-order G-quadruplex structures were observed at almost 100% frequency for DNA 1 and DNA 2. Apparently, the superhelix structures are formed by human single-stranded telomeric-overhang DNA.

Next, FRET and CD experiments were performed to further confirm the higher-order G-quadruplex structure. Formation of a higher-order G-quadruplex was accompanied by dimer formation by two G-quadruplex units. A fluorescence color change induced by FRET between the donor (6-FAM) and the acceptor (6-TAMRA) in 46 nt human telomere DNA 3 suggests that a dimeric structure is formed by two G-quadruplex units (Figure S3 in the Supporting Information). The formation of a higher-order G-quadruplex structure was further demonstrated by comparing the CD spectra of single-stranded DNA having no telomeric sequence and telomere-

sequence-containing DNA (DNA 4–8, Figure S4 in the Supporting Information).[9]

In previous studies, a mixed-type G-quadruplex structure has been demonstrated to form by a 22 nt human telomere sequence.[3] This structure should easily fold into a higher-order packing structure by consecutive formation of G-quadruplex units, because the 5′ and 3′ ends of the mixed G-quadruplex point in opposite directions, allowing the structure to fold end-to-end in the elongated telomeric DNA strand (Figure 1, Figure 2 f).[3] End-to-end connection is simply accomplished by a TTA linker to connect each 22 nt G-quadruplex structure (Figure 2 e,f).[3] In further support of the dimeric G-quadruplex, a study by molecular dynamics simulation showed a stable dimeric G-quadruplex structure for human telomeric DNA in which the interface between quadruplex units is stabilized by specific stacking interactions of loop nucleotides,[10] in contrast to a "string of beads" model.[11] A crystal structure study also showed that a dimeric structure with two G-quadruplex units formed by human telomeric DNA is stabilized by an acridine drug, BRACO-19, that is sandwiched between the interface of two G-quadruplex units.[12]

One important function of telomeres is to protect chromosome ends from being recognized as double-strand breaks.[1] Indeed, perturbation of telomeres can efficiently induce a DNA damage response characteristic of DNA double-strand breaks.[13] It has been suggested that the function of a telomere might depend on its state: whether it is a "capping" or an "uncapping" structure.[1] The consecutive formation of G-quadruplex units in human single-stranded telomeric-overhang DNA, proposed above on the basis of structural studies, may provide a protective capping structure for the telomere end. To test this hypothesis, we designed a series of DNA substrates (DNA 9–11) and tested their capability to activate DNA damage signals in an in vitro system (Figure 3). This nuclear extract system, which has been employed in a previous study,[14] is capable of activating various DNA damage signals, including phosphorylations of p53 (Ser-15), in response to exogenous linear DNA. DNA 9, a 24 nt duplex DNA substrate with both ends exposed, was highly efficient in activating phosphorylation of p53 (Ser-15; lanes 2 and 3). DNA 10, structurally identical to DNA 9 except that the duplex ends contain two single G-quadruplex units, was less efficient (lanes 4 and 5). DNA 11, a duplex DNA substrate with both ends protected by four G-quadruplex units, exhibited no phosphorylation of p53 (Ser-15; lanes 6 and 7). DNA 9 activated DNA damage signals, thus suggesting that exposure of the duplex ends is recognized by a sensor mechanism, presumably a nuclear protein or protein complex, and that this mechanism initiates a signaling cascade that includes a key regulator such as DNA-dependent protein kinase (DNA-PK).[15] In contrast, DNA 11 did not activate DNA damage signals. A simple explanation for the protective action of DNA 11 invokes a steric interference model in which the higher-order G-quadruplex at the ends of DNA 11 may block the access of key regulatory molecules such as DNA-PK and ataxia telangiectasia mutated (ATM) to the ends of the duplex.[14,15] A control experiment suggested that the G block in the telomeric repeat sequence is important for the inhibitory activity of the 3′ overhang. DNA 12, which was structurally similar to DNA 10 except that the G-quadruplex unit at the duplex ends was replaced with a mutant having no telomeric repeat sequence, stimulated phosphorylation of p53 (Ser-15; Figure S5 in the Supporting Information).

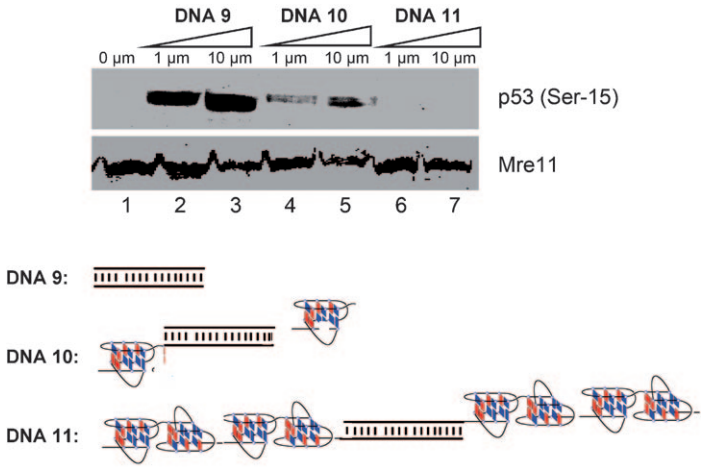

Figure 3. Human higher-order DNA G-quadruplex inhibits end-mediated DNA damage signals. DNA 9–11 (see structures in the top panel) were incubated with nuclear extracts, and p53 (Ser-15) phosphorylation was measured by immunoblotting with phospho-p53 (Ser-15) antibody (top). Reaction mixtures were analyzed, after incubation, by western blotting using phospho-p53 (Ser-15) antibody. For loading controls, reactions were also immunoblotted with Mre11-specific antibody (bottom). DNA concentrations: 0, 1, 10 μM.

Our results are consistent with previous studies suggesting that telomere overhang loss as a molecular signal triggers senescence.[1] Overhang loss leads to eventual telomere uncapping, that is, disruption of the proper structure of the protective cap at the end of the telomere.[1] It is reasonable to conclude that when the structures are perturbed, the uncapped duplex ends can be efficiently recognized by key DNA damage regulators such as DNA-PK and ATM, leading to activation of various DNA damage signals. Without dedicated protection, linear chromosome ends are identified by the cell's DNA damage surveillance machinery as DNA breaks, which can lead to many types of genome aberrations.[16]

The protective effect was further confirmed by hydrolysis assay using T4 polymerase.[17] This enzyme with 3′ to 5′ exonuclease activity hydrolyzes both single- and double-stranded DNA. The assay was performed on DNA 1 and a control, DNA 13, having no telomeric sequence. The hydrolysis was arrested in DNA 1 as compared with control DNA 13 (Figure S6 in the Supporting Information), indicating that DNA 1 was protected from digestion and thus suggesting that higher-order G-quadruplex formation induces a strong exonuclease resistance. These results indicate that consecutive G-quadruplex units in human single-stranded telomeric-overhang DNA may constitute a protective structure against such hydrolysis by, for example, exonuclease.

The results herein reveal the consecutive formation of G-quadruplex in human single-stranded telomeric-overhang DNA. The higher-order DNA G-quadruplex structure protects DNA double-strand ends from being recognized as double-strand breaks and directs against nuclease hydrolysis,

suggesting that this superhelix structure might be important for effectively packing telomeric DNA into a protective capping state. It is noteworthy that different structural forms may be important for participation in different biological functions of telomeres. The higher-order structure and T-loop may provide conformational flexibility for chromosome ends in response to environmental conditions such as protein binding. These results may possibly solve the puzzles pointed out above. Several telomeric DNA-binding proteins have been found to associate with telomere G-quadruplex. Pot1 is known to disrupt telomeric G-quadruplex, allowing telomerase extension.[18] Conversely, Rap1 is believed to promote G-quadruplex formation.[19] By coordination with such proteins, the forming and opening of higher-order DNA G-quadruplex might provide a reasonable mechanism for protecting and regulating chromosome ends. The proposed structure might also have clinical implications. Structure-based design of telomerase inhibitors requires molecular understanding of G-quadruplex topologies, and this structure is valuable in providing such information.

Received: July 14, 2009
Published online: September 15, 2009

Keywords: DNA · DNA structures · G-quadruplexes · telomeres

[1] a) E. H. Blackburn, *Cell* **2001**, *106*, 661 – 673; b) T. R. Cech, *Cell* **2004**, *116*, 273 – 279; c) M. A. Blasco, *Nat. Rev. Genet.* **2005**, *6*, 611 – 622; d) J. Karlseder, R. E. Verdun, *Nature* **2007**, *447*, 924 – 931; e) T. de Lange, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 323 – 329; f) W. E. Wright, V. M. Tesmer, K. E. Huffman, S. D. Levene, J. W. Shay, *Genes Dev.* **1997**, *11*, 2801 – 2809; g) N. Maizels, *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055 – 1059.

[2] a) S. Neidle, G. Parkinson, *Nat. Rev. Drug Discovery* **2002**, *1*, 383 – 393; b) L. H. Hurley, *Nat. Rev. Cancer* **2002**, *2*, 188 – 200; c) S. Neidle, *Curr. Opin. Struct. Biol.* **2009**, *19*, 239 – 250; d) S. Balasubramanian, S. Neidle, *Curr. Opin. Chem. Biol.* **2009**, *13*, 345 – 353.

[3] a) Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263 – 282; b) G. N. Parkinson, M. P. Lee, S. Neidle, *Nature* **2002**, *417*, 876 – 880; c) Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg. Med. Chem.* **2006**, *14*, 5584 – 5591; d) A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *FEBS J.* **2007**, *274*, 3545 – 3556; e) A. Ambrus, D. Chen, J. X. Dai, T. Bialis, R. A. Jones, D. Z. Yang, *Nucleic Acids Res.* **2006**, *34*, 2723 – 2735; f) K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, *128*, 9963 – 9970; g) Y. Xu, H. Sato, Y. Sannohe, K. I. Shinohara, H. Sugiyama, *J. Am. Chem. Soc.* **2008**, *130*, 16470 – 16471.

[4] J. D. Griffith, L. Comeau, S. Rosenfield, R. M. Stansel, A. Bianchi, H. Moss, T. de Lange, *Cell* **1999**, *97*, 503 – 514.

[5] R. M. Stansel, T. de Lange, J. D. Griffith, *EMBO J.* **2001**, *20*, 5532 – 5540.

[6] S. A. Stewart, I. Ben-Porath, V. J. Carey, B. F. O'Connor, W. C. Hahn, R. A. Weinberg, *Nat. Genet.* **2003**, *33*, 492 – 496.

[7] B. Drake, C. Prater, A. Weisenhorn, S. Gould, T. Albrecht, C. Quate, D. Cannell, H. Hansma, P. Hansma, *Science* **1989**, *243*, 1586 – 1589.

[8] a) T. Marsh, J. Vesenka, E. Henderson, *Nucleic Acids Res.* **1995**, *23*, 696 – 700; b) C. H. Kang, X. Zhang, R. Rathff, R. Moyzis, A. Rich, *Nature* **1992**, *356*, 126 – 131.

[9] a) S. Kumari, A. Bugaut, J. L. Huppert, S. Balasubramanian, *Nat. Chem. Biol.* **2007**, *3*, 218 – 221; b) I. N. Rujan, J. C. Meleney, P. H. Bolton, *Nucleic Acids Res.* **2005**, *33*, 2022 – 2031; c) J. L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* **2005**, *33*, e138; d) J. Qi, R. H. Shafer, *Biochemistry* **2007**, *46*, 7599 – 7606.

[10] L. Petraccone, J. O. Trent, J. B. Chaires, *J. Am. Chem. Soc.* **2008**, *130*, 16530 – 16532.

[11] H. Q. Yu, D. Miyoshi, N. Sugimoto, *J. Am. Chem. Soc.* **2006**, *128*, 15461 – 15468.

[12] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Am. Chem. Soc.* **2008**, *130*, 6722 – 6724.

[13] a) D. Churikov, C. Wei, C. M. Price, *Mol. Cell. Biol.* **2006**, *26*, 6971 – 6982; b) G. B. Celli, T. de Lange, *Nat. Cell Biol.* **2005**, *7*, 712 – 718.

[14] Y. C. Tsai, H. Qi, L. F. Liu, *J. Biol. Chem.* **2007**, *282*, 18786 – 18792.

[15] S. M. Bailey, M. A. Brenneman, J. Halbrook, J. A. Nickoloff, R. L. Ullrich, E. H. Goodwin, *DNA Repair* **2004**, *3*, 225 – 233.

[16] A. Lechel, A. Satyanarayana, Z. Ju, R. R. Plentz, S. Schaetzlein, C. Rudolph, L. Wilkens, S. U. Wiemann, G. Saretzki, N. P. Malek, M. P. Manns, J. Buer, K. L. Rudolph, *EMBO Rep.* **2005**, *6*, 275 – 281.

[17] Y. Yao, Q. Wang, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2007**, *35*, e68.

[18] A. J. Zaug, E. R. Podell, T. R. Cech, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10864 – 10869.

[19] R. Giraldo, M. Suzuki, L. Chapman, D. Rhodes, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7658 – 7662.

