JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

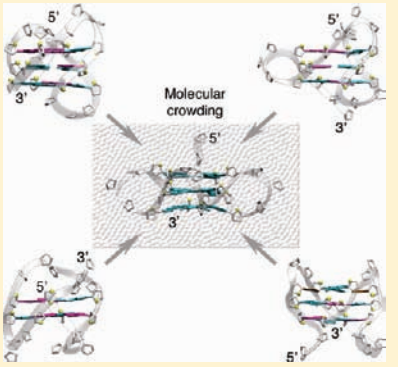
Structure of Human Telomeric DNA in Crowded Solution

Brahim Heddi and Anh Tuân Phan*

School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore

Supporting Information

ABSTRACT: G-quadruplex structures formed by DNA at the human telomeres are attractive anticancer targets. Human telomeric sequences can adopt a diverse range of intramolecular G-quadruplex conformations: a parallel-stranded conformation was observed in the crystalline state, while at least four other forms were seen in K^+ solution, raising the question of which conformation is favored in crowded cellular environment. Here, we report the first NMR structure of a human telomeric G-quadruplex in crowded solution. We show that four different G-quadruplex conformations are converted to a propeller-type parallel-stranded G-quadruplex in K^+-containing crowded solution due to water depletion. This study also reveals the formation of a new higher-order G-quadruplex structure under molecular crowding conditions. Our molecular dynamics simulations of solvent distribution provide insights at molecular level on the formation of parallel-stranded G-quadruplex in environment depleted of water. These results regarding human telomeric DNA can be extended to oncogenic promoters and other genomic G-rich sequences.



INTRODUCTION

Telomeres are the protective ends of linear eukaryotic chromosomes.[1] Human telomeric DNA consist mainly of thousands of TTAGGG repeats terminated with a 100−200-nt 3′-end overhang.[2] In vitro, human telomeric DNA sequences can form four-stranded helical structures called G-quadruplexes,[3−6] built from the stacking of multiple G·G·G·G tetrads.[7] G-quadruplexes at telomeres have been detected in vivo,[8] and their existence in living cells can be regulated by a number of proteins.[8,9] Intramolecular G-quadruplexes formed by human telomeric DNA sequences are promising anticancer targets,[10,11] because formation of such structures by the telomeric 3′-end overhang inhibits the activity of telomerase,[12,13] an enzyme necessary for the proliferation of most human cancer cells.[14]

G-quadruplex structures are highly polymorphic:[4−6,15] different G-quadruplex topologies correspond to very different shapes and dimensions of various structural elements, such as grooves and loops. Therefore, to understand the functions of G-quadruplexes in telomeres and to efficiently target them, it is important to know their detailed structures. Extensive research has been dedicated to the structures formed by sequences containing four human telomeric repeats, the minimum length required for intramolecular G-quadruplex folding.[15] At least five different intramolecular G-quadruplexes have been reported for DNA sequences containing TTAGGG repeats (Figure S1, Supporting Information). The d[AGGG(TTAGGG)$_3$] sequence forms in Na^+ solution a basket-type antiparallel-stranded G-quadruplex,[16] and in K^+-containing crystal a propeller-type parallel-stranded G-quadruplex.[17] In K^+ solution, multiple G-quadruplex conformations have been observed: the d[TAGGG(TTAGGG)$_3$] and d[TAGGG(TTAGGG)$_3$TT] sequences form predominantly (3 + 1)-type G-quadruplexes Form 1[18−20] and Form 2,[21] respectively, which differ from each other only by the order of loop arrangements; the d[GGG(TTAGGG)$_3$T] sequence forms predominantly Form 3, which is another basket-type G-quadruplex involving only two G-tetrads instead of three G-tetrads as in other forms.[22] Human telomeric DNA can be interspersed with some sequence-variant repeats, such as CTAGGG, which have been shown to cause genetic instability.[23] The d[AGGG(CTAGGG)$_3$] sequence, containing CTAGGG repeats, forms in K^+ solution a chair-type antiparallel-stranded G-quadruplex involving two G-tetrads and a G·C·G·C tetrad[24] (Figure S1), which is the sixth intramolecular G-quadruplex conformation determined so far for human telomeric DNA.

In living cells, biological macromolecules function in a crowded intracellular environment.[25−27] Molecular crowding could affect the structure, stability, and activity of biomolecules.[25−27] It is important, yet challenging, to know which G-quadruplex conformation of the human telomeric sequences is favored in a crowded cell-like environment.[28] Several groups used CD spectra to interpret structural transitions of G-quadruplexes under molecular crowding conditions:[29−33] positive peaks at 260 and 295 nm are characteristic of parallel and antiparallel-stranded G-quadruplexes, respectively.[34] Miyoshi et al.[29] reported that molecular crowding simulated by polyethylene glycol (PEG) induced conformational transition of a *Oxytricha* telomeric sequence from an antiparallel to a parallel-stranded G-quadruplex. Similarly, PEG was shown to induce conformational transition in a human telomeric sequence,[30,33]

Received: January 26, 2011
Published: May 07, 2011

dx.doi.org/10.1021/ja200786q | *J. Am. Chem. Soc.* 2011, 133, 9824–9833

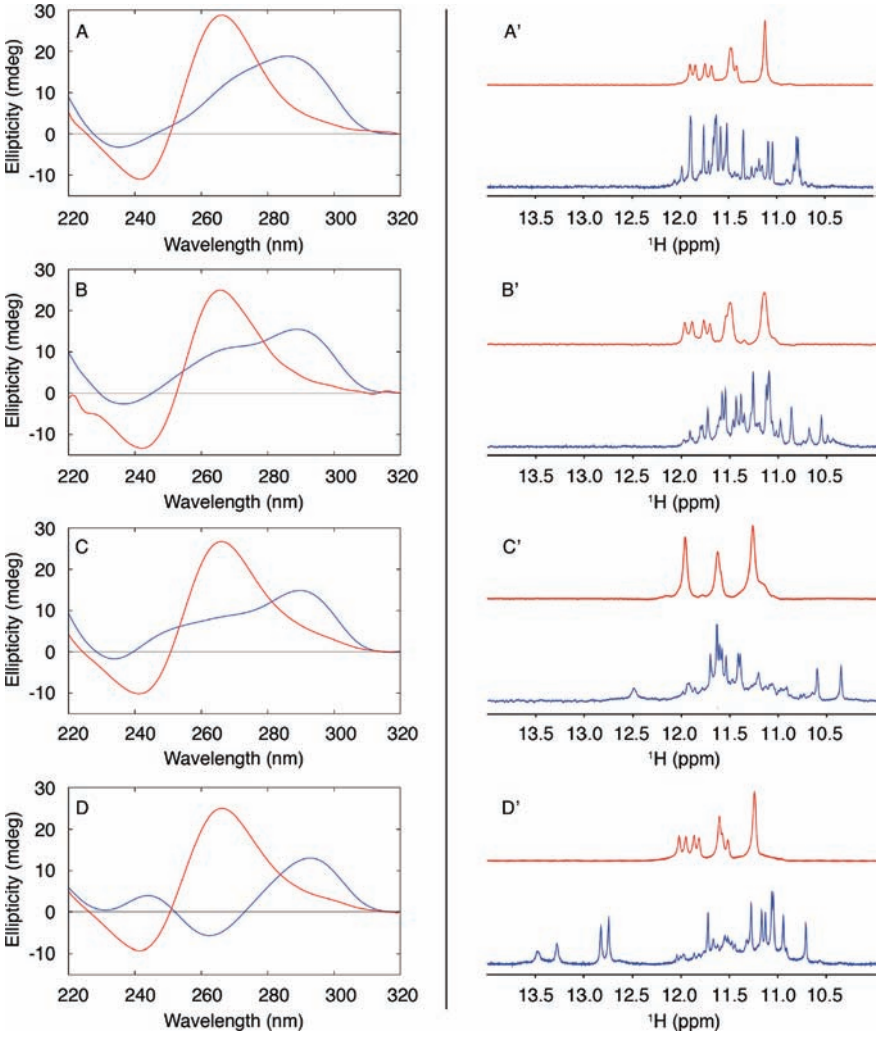


Figure 1. Detection of conformational transition in K^+-containing crowded solution. CD and NMR imino proton spectra of four different human telomeric DNA sequences in dilute (colored in blue) and crowded (colored in red) solutions. (A and A′) Sequence *Htelo1* d[$TAG_3(TTAG_3)_3$]; (B and B′) sequence *Htelo2* d[$TAG_3(TTAG_3)_3TT$]; (C and C′) sequence *Htelo3* d[$G_3(TTAG_3)_3T$]; (D and D′) sequence *Htelo4* d[$AG_3(CTAG_3)_3$]. The crowding condition was created by adding 40% (v/v) PEG 200.

and it was suggested that 40% (w/v) PEG 200 induced conformational switch of this sequence from an antiparallel to a parallel-stranded G-quadruplex.[33] The role of hydration on G-quadruplexes under molecular crowding conditions was discussed in the works of Miyoshi et al.[31] using various carbohydrate cosolutes and Vorlikova et al.[32] using ethanol cosolute. During the writing of this Article, a paper by Miller et al.[35] appeared online showing that 50% (v/v) acetonitrile could induce conformational transition in a human telomeric sequence, supporting the role of hydration in determining the G-quadruplex conformation. However, the authors concluded that the G-quadruplex conformation formed in this condition[35] was not identical to the parallel form observed in the crystalline state.[17] So far, high-resolution structure of human telomeric sequences in crowded solution has not been determined, and the conformational transition induced by molecular crowding condition is not fully understood.

In this work, we used CD and NMR spectroscopy to characterize the structures of human telomeric sequences in K^+-containing crowded solution. We determined the first NMR structure of a human telomeric G-quadruplex in K^+-containing crowded solution. We showed that four different G-quadruplex topologies found in K^+ solution were converted to parallel-stranded G-quadruplexes under molecular crowding conditions, simulated by the addition of various cosolutes, due to water depletion. This conformational transition was reversible, and the kinetics of the conversion could be measured in real-time. Molecular dynamics simulations of solvent distribution advanced our understanding on how the parallel-stranded G-quadruplex is favored in condition where the water content is reduced. We also found evidence for the formation of a new higher-order G-quadruplex structure in crowded solution.

RESULTS AND DISCUSSION

Molecular Crowding, Simulated by Different Cosolutes, Induces Conformational Transition of Human Telomeric Sequences from Four Different G-Quadruplexes to a Parallel-Stranded G-Quadruplex. In dilute solution (water solution containing 90 mM potassium salt), CD and NMR imino proton spectra (blue curves, Figure 1) of the four human telomeric sequences d[$TAGGG(TTAGGG)_3$], d[$TAGGG(TTAGGG)_3TT$],

dx.doi.org/10.1021/ja200786q | *J. Am. Chem. Soc.* 2011, 133, 9824–9833

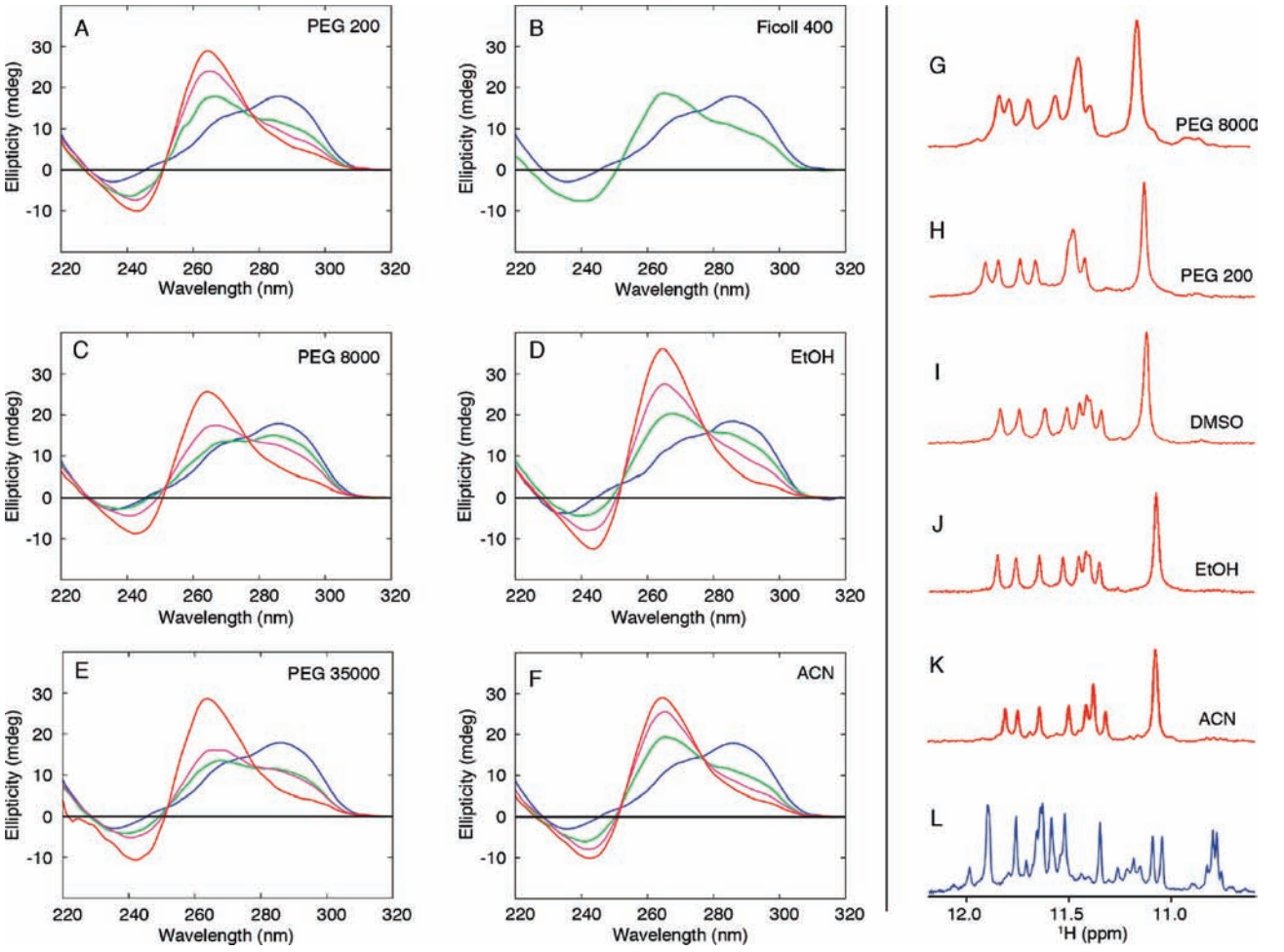


Figure 2. Induction of conformational transition by different cosolutes. CD and NMR imino proton spectra of *Htelo1* in crowded solution induced by (A and H) PEG 200, (B) Ficoll 400, (C and G) PEG 8000, (D and J) ethanol, (E) PEG 35000, (F and K) acetonitrile, and (I) DMSO. CD spectra of *Htelo1* were recorded under 20% (v/v) (green), 30% (v/v) (magenta), and 40% (v/v) (red) of PEG 200, ethanol, acetonitrile, and DMSO; and under 200 g/L (green), 300 g/L (magenta), and 400 g/L (red) of PEG 8000, PEG 35000, and Ficoll 400. CD reference spectrum of *Htelo1* in dilute solution (blue) is shown in each case for comparison. NMR spectra of *Htelo1* were recorded in crowded solution induced by (G) 400 g/L of PEG 8000 and (H−K) 40% (v/v) of other cosolutes. NMR reference spectrum of *Htelo1* in dilute solution is shown in (L).

Table 1. Human Telomeric Sequences Used in This Study and Their Structures under Molecular Crowding Conditions, Simulated by Various Concentrations of PEG 200

name	sequence 5′ −3′	Structure		
		40% PEG 200	50% PEG 200	60% PEG 200
Htelo3	$G_3(TTAG_3)_3T$	High-order	High-order	High-order
Htelo3b	$AG_3(TTAG_3)_3T$	Monomer	High-order	High-order
Htelo1	$TAG_3(TTAG_3)_3$	Monomer	High-order	High-order
Htelo4	$AG_3(CTAG_3)_3$	Monomer	High-order	High-order
Htelo2	$TAG_3(TTAG_3)_3TT$	Monomer	Monomer	High-order

d[$GGG(TTAGGG)_3T$], and d[$AGGG(CTAGGG)_3$] (designated *Htelo1*, *Htelo2*, *Htelo3*, and *Htelo4*, respectively, Table 1) were distinct and corresponded to four different G-quadruplex conformations reported previously.[19,21,22,24] In crowded solution, simulated by the addition of 40% (v/v) PEG 200, the CD and NMR profiles of these four sequences became similar (red curves, Figure 1): CD spectra gave a positive peak at 260 nm and a negative peak at 240 nm, characteristic of a parallel-stranded G-quadruplex, while NMR imino proton spectra were less disperse with peaks regrouped in three groups. These observations suggested that molecular crowding induced conformational transition from all four different G-quadruplex conformations to a parallel-stranded G-quadruplex (see structure and discussion below).

Other CD and NMR data (Figure 2 and Figures S2−S4) showed that similar conformational transitions from these four different G-quadruplexes to a parallel-stranded G-quadruplex could occur under molecular crowding conditions, generated by the addition of different cosolutes, including polyethylene glycol (PEG) of various molecular weights (200, 8000, and 35 000 Da), a polysaccharide Ficoll 400, ethanol (EtOH), acetonitrile (ACN), and dimethyl sulfoxide (DMSO).

For PEG 200, EtOH, and ACN, the amount of cosolutes required for full conversion of *Htelo1* to a parallel-stranded G-quadruplex was about 40% of the volume (Figure S5). The effect of different crowding agents was observed to be additive, as the combination of 20% (v/v) PEG and 20% (v/v) ACN (Figure S6) or 20% (v/v) PEG and 200 g/L Ficoll (Figure S7) almost

dx.doi.org/10.1021/ja200786q | *J. Am. Chem. Soc.* 2011, 133, 9824–9833

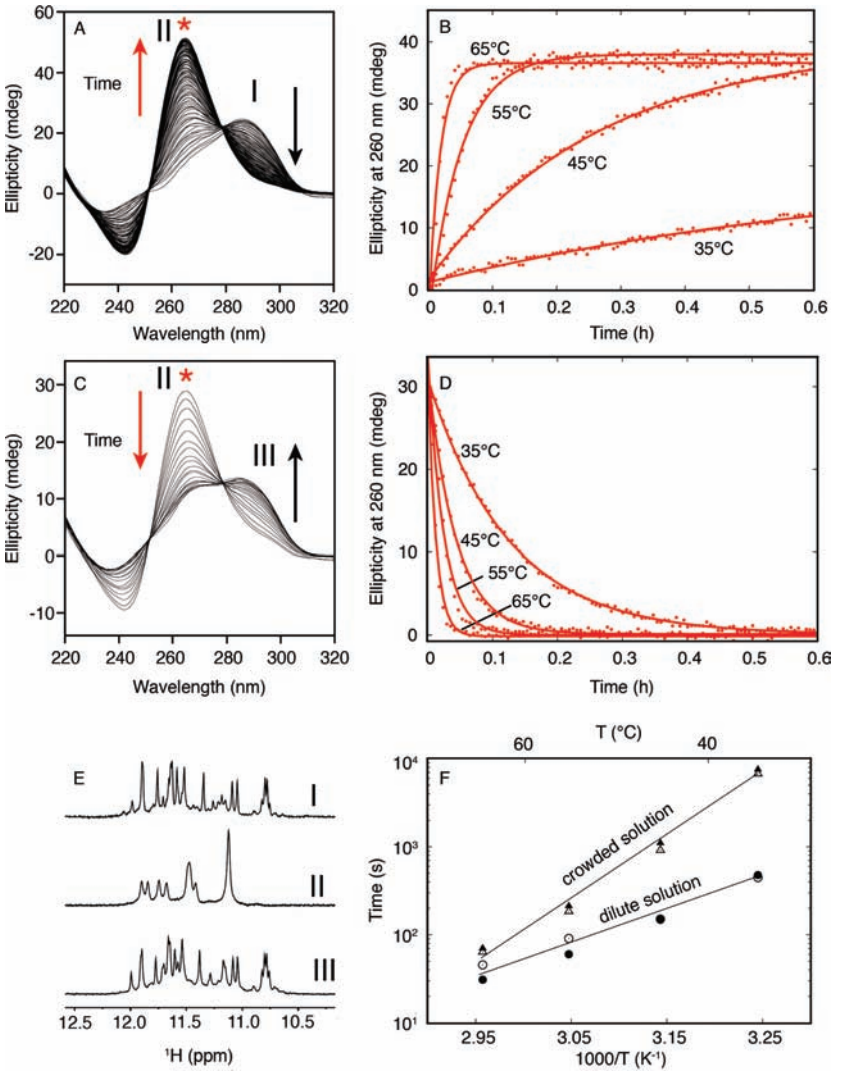


Figure 3. Kinetics of conformational transition. (A, B) Conformational transition of *Htelo1* from a (3 + 1) G-quadruplex in dilute solution to a parallel-stranded G-quadruplex in crowded solution monitored by CD spectra: (A) CD spectra of *Htelo1* recorded at different times following addition of 40% (v/v) PEG 200 at 45 °C and (B) the 260-nm CD peak intensity as a function of time at different temperatures. (C, D) Conformational transition of *Htelo1* from a parallel-stranded G-quadruplex in crowded solution to a (3 + 1) G-quadruplex in dilute solution monitored by CD spectra: (C) CD spectra of *Htelo1* recorded following 100-fold dilution of a concentrated PEG-containing sample and (D) the 260-nm peak intensity as a function of time at different temperatures. Spectra were recorded every 13 s. (E) NMR imino proton spectra of *Htelo1* before adding PEG 200 (top; state I), after equilibrium with 40% (v/v) PEG 200 (middle; state II), and after dialysis against 20 mM KCl using membrane of 1 kDa cutoff (bottom; state III). (F) The time constants for conformational transition between the two G-quadruplex forms in crowded (▲ and △) and dilute solution (● and ○) at different temperatures.

fully converted *Htelo1* to a parallel-stranded G-quadruplex, while individually the same amount of each agent was not sufficient for a full conformational conversion. The observations of similar effects induced by different types and sizes of cosolutes indicated the predominant role of water depletion rather than specific cosolute−DNA interactions on this conformational transition.

Kinetics of G-Quadruplex Conformational Transitions. The conformational transition of *Htelo1* from a (3 + 1) G-quadruplex[19] (Figure S1) to a parallel-stranded G-quadruplex (see structure below) under molecular crowding was followed by monitoring the CD spectra as a function of time (Figure 3A). Upon addition of 40% (v/v) PEG 200, the CD signal at 290−295 nm (a marker of the (3 + 1) form) decreased, while that at 260 nm (a marker of the parallel-stranded form) increased (Figure 3A). This conformational transition was investigated at different temperatures, ranging from 35 to 65 °C (Figure 3B and Figure S8). The identification of each G-quadruplex conformation was verified by NMR (Figure 3E): the spectrum at the state I, recorded at the beginning of the experiment, corresponded to the (3 + 1) G-quadruplex;[19] while the spectrum at the state II, recorded at the end of the experiment, belonged to the parallel-stranded G-quadruplex (see below).

The reverse conformational transition of *Htelo1*, from the parallel-stranded G-quadruplex to the (3 + 1) G-quadruplex, was achieved by a quick sample dilution, which resulted in a negligible PEG concentration. This conformational transition was followed with a decrease in the 260 nm CD signal and an increase in the 290−295 nm CD signal (Figure 3C and Figure S8). The transition kinetics was determined at different temperatures, ranging from 35 to 65 °C (Figure 3D and Figure S8). Formation of the (3 + 1) G-quadruplex (the state III) at the end of the reverse conformational transition was also verified by NMR (Figure 3E).

dx.doi.org/10.1021/ja200786q | *J. Am. Chem. Soc.* 2011, 133, 9824–9833

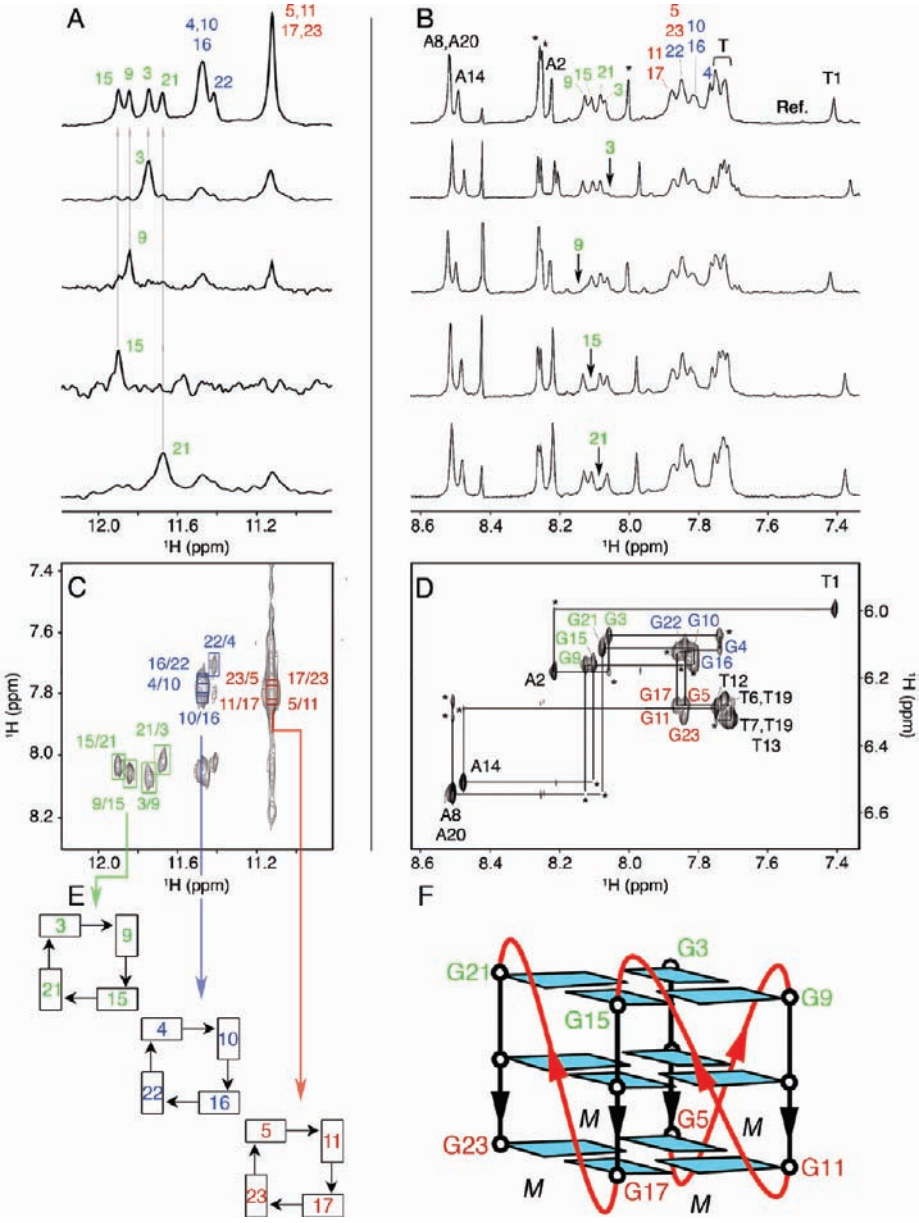


Figure 4. NMR spectral assignments and topology determination in crowded solution. (A) Examples of guanine imino proton assignments of *Htelo1* in ^{15}N filtered spectra of samples, ~2% ^{15}N labeled at the indicated positions with the reference spectrum shown at the top. Residual signals from nonlabeled guanines with ^{15}N natural abundance could be pronounced due to peak superposition. (B) Examples of guanine H8 proton assignments by site-specific ^{2}H labeling at the indicated positions with the reference spectrum shown at the top. (C−F) Determination of G-quadruplex folding topology. (C) NOESY spectrum (mixing time, 200 ms), showing imino-H8 connectivities. Cross-peaks that identify G-tetrad alignments are framed and labeled with the residue number of imino proton in the first position and H8 proton in the second position. (D) H8/6-H1′ proton region of NOESY sectrum (mixing time, 300 ms), showing the assignments and H8/6-H1′ NOE sequential connectivities. (E) Guanine imino-H8 NOE connectivity patterns around tetrads as indicated with arrows, observed for G3·G9·G15·G21, G4·G10·G16·G22, and G5·G11·G17·G23 tetrads. (F) Schematic structure of *Htelo1* in K^+-containing crowded solution. Anti guanines are colored cyan; TTA linkers are colored red.

The time constants for conformational transition between the two G-quadruplex forms in crowded and dilute solution are plotted in Figure 3F. In crowded solution, the time constants measured for the formation of the parallel-stranded form (▲, Figure 3F) were found to coincide with those measured for the disappearance of the (3 + 1) form (△, Figure 3F). In dilute solution, the time constants measured for the disappearance of the parallel-stranded form (●, Figure 3F) were found to coincide with those measured for the formation of the (3 + 1) form (○, Figure 3F). At 35 °C, the formation of the parallel form under molecular crowding is in the order of an hour, while its unfolding in dilute solution is in the order of a minute. The corresponding enthalpy activation of the transitions was 136 and 79 kJ/mol, respectively.

NMR Structure Determination of the Parallel-Stranded G-Quadruplex Form. We decided to determine the solution structure of the G-quadruplex formed by the human telomeric sequence *Htelo1* under molecular crowding condition, simulated by 40% (v/v) PEG 200. Guanine imino and H8 protons were unambiguously assigned using site-specific low-enrichment ^{15}N labeling[36] and site-specific ^{2}H labeling[37] (Figure 4A, B and

dx.doi.org/10.1021/ja200786q |*J. Am. Chem. Soc.* 2011, 133, 9824–9833

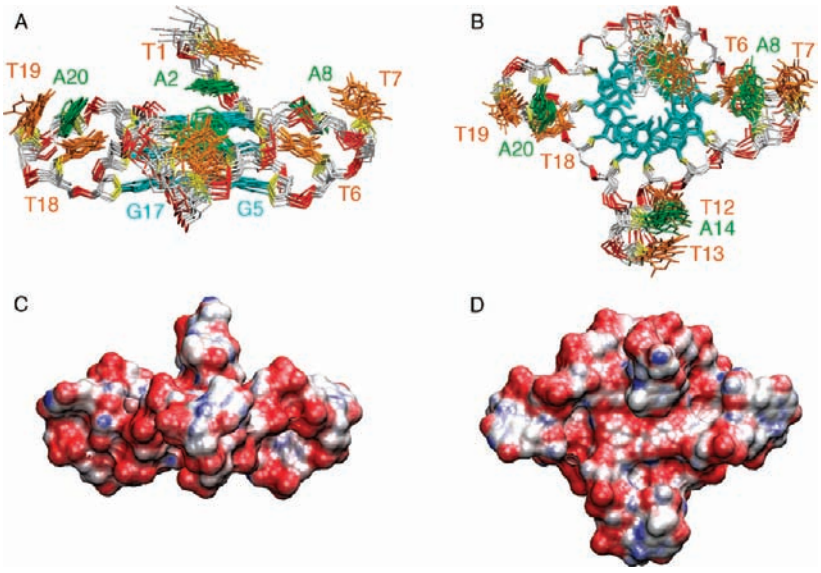


Figure 5. NMR structure of the human telomeric G-quadruplex in crowded solution. (A) Front and (B) top views of 10 superimposed refined structures of *Htelo1* in K^+-containing crowded solution. Anti guanines are colored cyan; adenines green; thymines orange; backbone gray; sugar ring oxygens yellow; phosphorus atoms red. (C) Front and (D) top surface views of a representative refined structure. Positive and negative electrostatic potentials are shown in blue and red, respectively.

Table 2. NMR Restraints and Structure Statistics

(A) NMR Restraints		
distances restraints	^2H_2O	H_2O
intraresidue distance restraints	141	0
sequential (i,$i+1$) distance restraints	100	4
long range (i,$>i+1$) distance restraints	51	17
other restraints		
hydrogen-bond restraints	24	
dihedral restraints	35	
repulsive restraints	15	
(B) Structure Statistics		
NOE violations		
number (>0.2 Å)	2	
maximum violation (Å)	0.26 ± 0.02	
rmsd of violations (Å)	0.02 ± 0.00	
deviations from the ideal covalent geometry		
bond lengths (Å)	0.00 ± 0.00	
bond angles (deg)	0.76 ± 0.01	
impropers (deg)	0.42 ± 0.01	
pairwise all heavy atom rmsd values (Å)		
G-tetrad core	0.66 ± 0.17	
all residues	1.15 ± 0.23	

Figures S9−S11; Table S1). Thymine H6 protons were assigned following site-specific T-to-U substitutions[38] (Figure S12). The H8/H6−H1′ NOE sequential connectivity could be traced from T1 through G23 (Figure 4D). The intensity of intraresidue H8−H1′ NOE cross-peaks (Figure 4D) indicated anti glycosidic conformation for all guanines.

The imino-H8 connectivity patterns (Figure 4C,E) pointed to the formation of an intramolecular propeller-type parallel-stranded G-quadruplex structure (Figure 4F): the core consists of three G-tetrads, G3·G9·G15·G21, G4·G10·G16·G22, and G5·G11·G17·G23; four grooves are of medium size; three TTA loops are double-chain-reversal. The central placing of the G4·G10·G16·G22 tetrad was consistent with the observation of the imino protons of G4, G10, G16, and G22 being well protected from the exchange with solvent (Figure S13). We distinguished three groups of resonances each belonging to a G-tetrad (Figure 4). The peaks of the 5′-end G-tetrad layer (colored in green) were resolved, while those of the middle G-tetrad (colored in blue) and 3′-end G-tetrad (colored in red) were partially degenerated, consistent with higher symmetry of the 3′-end of the structure. The observation of sharp peaks for aromatic protons of adenines (Figure 4B) indicated higher flexibility of loop residues as compared to the core of the G-quadruplex.

The structure of the propeller-type G-quadruplex in K^+-containing crowded solution (Figure 5) was calculated on the basis of NMR data (Table 2); with a number of NOE restraints unambiguously obtained using site-specific ^2H labeled samples (see examples in Figures S10 and S11). The calculated structure is a propeller-type parallel-stranded G-quadruplex with two 5′-end bases being stacked on the top of the G-tetrad core (Figure 5). Continuous stacking of T1, A2, and G3 was supported by numerous NOEs between protons of these bases (Figure S14). The three TTA double-chain-reversal loops adopt similar structures (pairwise rmsd, 1.2 Å). Bases in each loop stack partially; the adenine is sandwiched between the two thymines (Figure S15).

A propeller-type parallel-stranded G-quadruplex was previously determined by X-ray crystallography[17] for the human telomeric d[AGGG(TTAGGG)$_3$] sequence (PDB ID: 1KF1). Figure 6 shows the superposition of this X-ray structure (colored in red) and our NMR structure (colored in blue). The G-tetrad cores of the two structures are similar (rmsd for heavy atoms, 1.4 Å), while the loops are somewhat different (rmsd for heavy atoms, 5.6 Å). The loops in the X-ray structure appear more flattened as compared to our NMR structure (Figure 6). In the

dx.doi.org/10.1021/ja200786q | *J. Am. Chem. Soc.* 2011, 133, 9824–9833

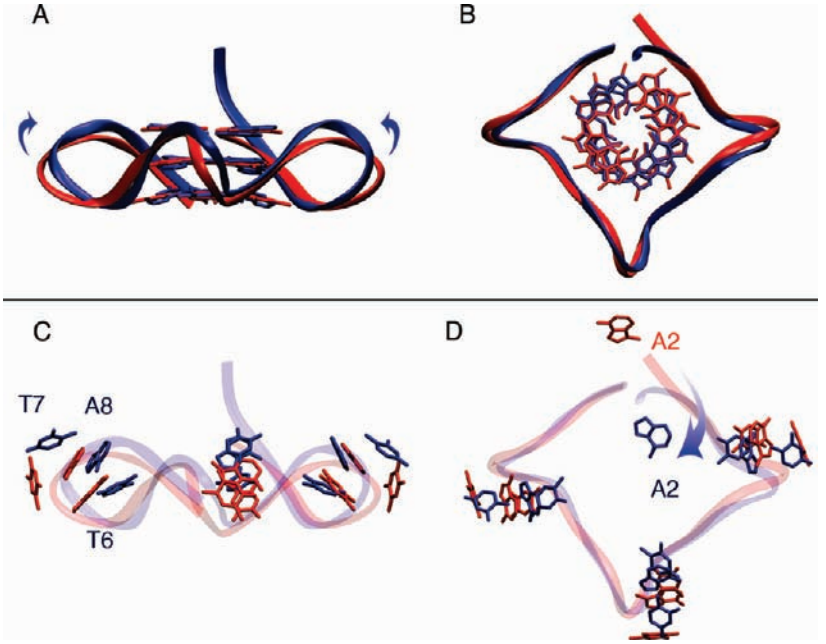


Figure 6. Comparison between human telomeric G-quadruplexes determined in crowded solution and crystalline state. Superposition of the NMR structure from this work (blue) and the X-ray crystal structure (PDB ID: 1KF1)[17] (red) based on the best fit of the G-tetrad core. (A and C) Front and (B and D) top views of the two structures. Bases of the G-tetrad core are shown in (A) and (B); bases of the loops are shown in (C) and (D). Residues in the X-ray crystal structure were renumbered according to the NMR structure. Arrows indicate some differences between the two structures.

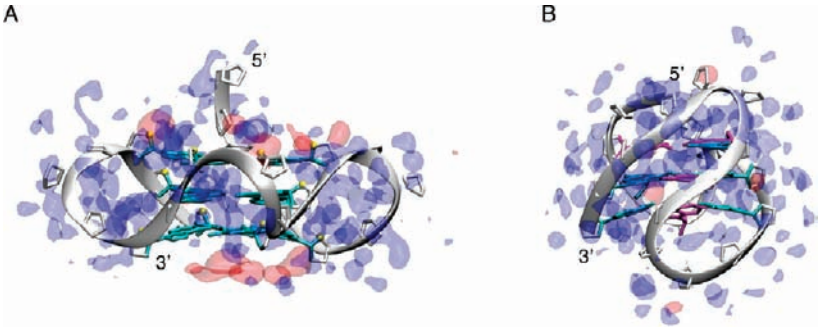


Figure 7. Computed solvent distribution at the surface of different G-quadruplexes. Solvent representations around (A) the parallel-stranded G-quadruplex (this work) and (B) the (3 + 1) G-quadruplex,[39] obtained from molecular dynamics simulations in a water-ethanol mixture. A pixel (unit volume) of 0.5 Å × 0.5 Å × 0.5 Å was colored in blue or red if it was occupied by water or ethanol molecules, respectively, for more than 55% of simulation time.

loops, the adenine base is positioned between the two thymine bases for both structures, but their arrangements and stacking patterns are different (Figure 6). At the 5′-end, the A2 base stacks on the top of the G-tetrad core in the NMR structure, but flips out in the X-ray structure (Figure 6). Some differences between the two structures could be partially explained in view of the stacking of the second T bases in the loops and the A2 base at the 5′-end with symmetry-related neighboring molecules in the crystal lattice (Figure S16).

Computational Study of Solvent Distribution. To better understand the drastic conformational change induced by molecular crowding, we performed molecular dynamics (MD) simulations on the (3 + 1) G-quadruplex (PDB ID: 2JSM)[39] and the parallel-stranded G-quadruplex structure (this work), both in a K^+-containing mixture of water and ∼40% (v/v) EtOH. Ethanol was chosen as the cosolute in this computational study, because the available model for EtOH molecules has already been shown to reproduce the physical properties of a water−ethanol mixture.[40] The simulations started with the DNA structures and K^+ counterions immersed in a box of water, surrounded by EtOH molecules forming 40% of the total volume. During the MD simulations, water and EtOH molecules rearranged by diffusion and reached an equilibrium distribution after ∼1.5 ns (Figure S17). At equilibrium, the distribution of EtOH and water molecules in the inner shell of the two G-quadruplexes was dramatically different for the two structures (Figure 7). Although no specific interaction was found between EtOH molecules and DNA, we observed an enrichment of EtOH molecules (or depletion of water) within 4 Å from the surface of the parallel-stranded G-quadruplex (mostly located at the two ends of the G-tetrad core) and, in the reverse, a depletion of EtOH molecules (or enrichment of water) within 4 Å from the

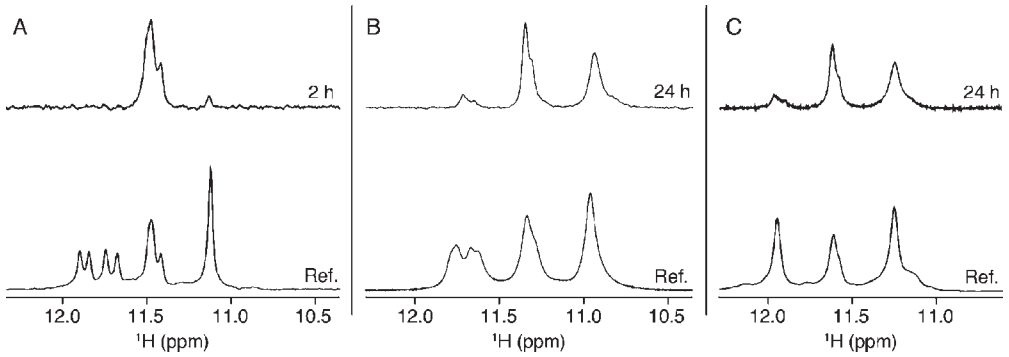


Figure 8. Solvent exchange ($H_2O \rightarrow {}^2H_2O$) experiments. Imino proton spectra of (A) *Htelo1* under 40% (v/v) PEG 200 before and at 2 h after dissolving the sample in ^2H_2O; (B) *Htelo1* under 50% (v/v) PEG 200 before and at 24 h after dissolving the sample in ^2H_2O; and (C) *Htelo3* under 40% (v/v) PEG 200 before and at 24 h after dissolving the sample in ^2H_2O.

surface of the (3 + 1) G-quadruplex. The observed difference in the solvent distribution for the two forms is consistent with the difference in their hydration ability inferred from their structures. In the parallel-stranded conformation, the hydrophobic aromatic rings of guanine bases at the ends of the G-tetrad core are more exposed to the solvent, while in the (3 + 1) conformation, these ends are partially covered by hydrophilic backbones. Crowded solution, in which water content is reduced being replaced by cosolutes, would provide more favorable condition for the formation of the parallel-stranded form over the (3 + 1) form.

We estimated the solvent content in the crystal, used by Parkinson et al.[17] for the X-ray structure determination of the parallel-stranded human telomeric G-quadruplex (PDB ID: 1KF1), to be ~55% of the total volume. In that case, together with PEG, DNA molecules themselves were crowding agents that depleted water from the surface of the neighboring counterparts. This self-induced crowding and the resulting water depletion found in the crystal should be the reason for the observation of the parallel-stranded G-quadruplex in the crystalline state, while other forms are observed in dilute solution.

Higher-Order and Higher-Symmetry Structures. Imino proton spectra of the human telomeric sequence *Htelo3* in K^+-containing crowded solution were degenerated and showed only three peaks (Figure 1C′), indicating the formation of a highly symmetric structure. Full conversion of *Htelo3* to this structure was detected at 40% (v/v) PEG 200 (Figure S18). Interestingly, a similar spectrum was observed for *Htelo1* in the presence of 50% (v/v) PEG 200 (Figure S19). Solvent exchange ($H_2O \rightarrow {}^2H_2O$) experiment (Figure 8) showed that these structures (i.e., *Htelo3* and *Htelo1* at 40% and 50% (v/v) of PEG 200, respectively) shared the same pattern of imino proton protection, in contrast to the structure of *Htelo1* in 40% (v/v) PEG 200. The former (Figure 8B,C) have two groups of imino protons protected from the exchange with solvent, while the latter (Figure 8A) only has one, which consists of imino protons from the central G-tetrad layer. This finding suggested the formation of a higher-order structure for *Htelo3* at 40% (v/v) PEG 200 and *Htelo1* at 50% (v/v) PEG 200. Such a structure could be formed by the stacking of two parallel-stranded G-quadruplex blocks.[41] Our data (e.g., Figures S18−S20) indicated that different telomeric sequences require different concentrations of PEG 200 for the formation of this higher-order structure (Table 1). It seems that sequences with non-G residues at the ends require higher concentrations of crowding agents to form this higher-order structure (Figures S18−S20; Table 1). The effect of non-G terminal residues to hinder the formation of higher-order structures seems to be larger for the 5′-end than the 3′-end (Figure S21, Table 1), coinciding with the higher stacking propensity at the 5′-end of parallel-stranded G-quadruplexes.[41]

CONCLUSION

We have solved the first NMR structure of a human telomeric G-quadruplex in crowded solution. Different G-quadruplex conformations can be converted to a propeller-type parallel-stranded G-quadruplex in K^+-containing crowded solution due to water depletion and these conformational transitions are reversible. Our molecular dynamics simulations provided insights at molecular level on the formation of parallel-stranded form in crowded solution. Finally, we showed the formation of higher-order structure in highly water-depleted solution at high PEG concentration.

METHODS

DNA Sample Preparation. Unlabeled, site-specific 2% ^{15}N labeled, and site-specific ^2H labeled DNA oligonucleotides were chemically synthesized and purified as described previously.[22,24] Samples were dialyzed successively against KCl solution and against water. Unless otherwise stated, the samples contained 70 mM KCl and 20 mM potassium phosphate buffer (pH 7). Unless otherwise stated, samples under molecular crowding conditions were equilibrated at a given temperature before taking measurements.

Crowding Agents. PEG 200, PEG 8000, PEG 35000, Ficoll 400, and DMSO were purchased from Sigma-Aldrich. Ethanol and acetonitrile were purchased from Merck and Fisher Scientific, respectively. Deuterated compounds, acetonitrile-d_3, ethanol-d_6, and DMSO-d_6, were purchased from Cambridge Isotope Laboratories.

CD Measurements. CD spectra were recorded on a Jasco-815 CD spectropolarimeter at 20 °C, unless otherwise stated. Scans from 220 to 320 nm were performed with 200 or 500 nm/min scanning speed, 1 nm pitch, and 1 nm bandwidth. For each spectrum, an average of three scans was taken, spectral contribution from the buffer was subtracted, and the data were zero-corrected at 320 nm. DNA concentration was 5−10 μM. For the measurements of conformational transition kinetics, the time constants were determined by fitting the time-dependent functions of the 260 and 290 nm peak intensity with an exponential, $y = A + B \times \exp(-\tau/t)$, where t represents the time variable, τ is the time constant, and A and B are constants.

NMR Spectroscopy. Experiments were performed on 600 and 700 MHz Bruker spectrometers equipped with a cryoprobe at 37 °C, unless

otherwise stated. DNA concentration was typically 0.1−1.0 mM. Most NMR experiments were performed on samples containing 5 mM KCl and 5 mM potassium phosphate (pH 7); NMR spectra of these samples were observed to be very similar to the spectra of samples in ∼100 mM K^+. The NMR signal of PEG was eliminated using a Jump-return-type[38,42] pulse sequence. For experiments performed in H_2O, the carrier frequency was set to 4.0 ppm, midway between the PEG signal (3.7 ppm) and the water signal (4.7 ppm). For experiments performed in ^2H_2O, the carrier frequency was placed at the PEG signal.

NMR-Restrained Structure Calculation. Interproton distances in the d[TAGGG(TTAGGG)$_3$] G-quadruplex were deduced from NOESY experiments performed in H_2O (mixing time, 200 ms) and ^2H_2O (mixing times, 100, 200, and 300 ms). Distance-geometry simulated annealing and distance-restrained molecular dynamics calculations were successively performed leading to 10 best structures in vacuo, using the XPLOR-NIH program[43] as previously described.[22,24] The final structures were subjected to molecular dynamics refinement in explicit solvent using the AMBER 10.0 program.[44] The system was neutralized by 22 K^+ cations (two of them were in the center of the G-tetrad core) and solvated with 8074 water molecules (TIP3P)[48] in a truncated octahedral box. An attempt to perform molecular dynamics refinement in a box containing 5551 water molecules and 1930 ethanol molecules, resulted in a similar structure as compared to the one refined in the box of 8074 water molecules (rmsd for all atoms, 1.29 Å). Hydrogen-bond restraints, interproton distance restraints, dihedral restraints, planarity restraints, and repulsive restraints were imposed during the structure calculations by the XPLOR-NIH program. During the molecular dynamics refinement in explicit solvent, planarity restraints and repulsive restraints were removed. Structures were displayed using the PyMOL[45] and VMD programs.[46]

Computation of Solvent Content. Molecular dynamics simulations were performed on the parallel-stranded G-quadruplex (this work) and the (3 + 1) G-quadruplex[39] of *Htelo1* for 8 ns at 300 K, using the AMBER 10.0 program[44] with the Parmbsc0 force field[47] and an all-atom flexible ethanol model.[40] The detailed protocol is described in the Supporting Information. Briefly, the system was neutralized by 22 K^+ cations (two of them were in the center of the G-tetrad core) and solvated with 5551 water molecules (TIP3P)[48] and 1930 EtOH molecules[40] in a truncated octahedral box. During the entire simulation, harmonic positional restraints were applied on DNA.

Data Deposition. The coordinates of the 10 best structures of the human telomeric sequence d[TAGGG(TTAGGG)$_3$] in K^+-containing crowded solution have been deposited in the Protein Data Bank (accession code 2LD8).

■ ASSOCIATED CONTENT

Supporting Information. Supplementary methods, tables, and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
phantuan@ntu.edu.sg

■ ACKNOWLEDGMENT

This research was supported by Singapore Biomedical Research Council grant 07/1/22/19/542 to A.T.P.

■ REFERENCES

(1) de Lange, T. *Science* **2009**, *326*, 948–952.
(2) Makarov, V. L.; Hirose, Y.; Langmore, J. P. *Cell* **1997**, *88*, 657–666.
(3) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164–168.
(4) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668–698.
(5) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429–7455.
(6) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239–250.
(7) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 2013–2018.
(8) Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 847–854.
(9) Maizels, N. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055–1059.
(10) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. *J. Med. Chem.* **1997**, *40*, 2113–2116.
(11) Mergny, J. L.; Helene, C. *Nat. Med.* **1998**, *4*, 1366–1367.
(12) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718–720.
(13) Oganesian, L.; Moon, I. K.; Bryan, T. M.; Jarstfer, M. B. *EMBO J.* **2006**, *25*, 1148–1159.
(14) 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.; Shay, J. W. *Science* **1994**, *266*, 2011–2015.
(15) Phan, A. T. *FEBS J.* **2010**, *277*, 1107–1117.
(16) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263–282.
(17) Parkinson, G. N.; Lee, M. P.; Neidle, S. *Nature* **2002**, *417*, 876–880.
(18) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723–2735.
(19) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970.
(20) Xu, Y.; Noguchi, Y.; Sugiyama, H. *Bioorg. Med. Chem.* **2006**, *14*, 5584–5591.
(21) Phan, A. T.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2006**, *34*, 5715–5719.
(22) Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 4301–4309.
(23) Mendez-Bermudez, A.; Hills, M.; Pickett, H. A.; Phan, A. T.; Mergny, J. L.; Riou, J. F.; Royle, N. J. *Nucleic Acids Res.* **2009**, *37*, 6225–6238.
(24) Lim, K. W.; Alberti, P.; Guedin, A.; Lacroix, L.; Riou, J. F.; Royle, N. J.; Mergny, J. L.; Phan, A. T. *Nucleic Acids Res.* **2009**, *37*, 6239–6248.
(25) Zimmerman, S. B.; Trach, S. O. *J. Mol. Biol.* **1991**, *222*, 599–620.
(26) Ellis, R. J.; Minton, A. P. *Nature* **2003**, *425*, 27–28.
(27) Zhou, H. X.; Rivas, G.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375–397.
(28) Miyoshi, D.; Sugimoto, N. *Biochimie* **2008**, *90*, 1040–1051.
(29) Miyoshi, D.; Nakao, A.; Sugimoto, N. *Biochemistry* **2002**, *41*, 15017–15024.
(30) Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, 4649–4659.
(31) Miyoshi, D.; Karimata, H.; Sugimoto, N. *J. Am. Chem. Soc.* **2006**, *128*, 7957–7963.
(32) Vorlickova, M.; Bednarova, K.; Kejnovska, I.; Kypr, J. *Biopolymers* **2007**, *86*, 1–10.
(33) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185–11191.
(34) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061–4067.
(35) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105–17107.
(36) Phan, A. T.; Patel, D. J. *J. Am. Chem. Soc.* **2002**, *124*, 1160–1161.
(37) Huang, X.; Yu, P.; LeProust, E.; Gao, X. *Nucleic Acids Res.* **1997**, *25*, 4758–4763.
(38) Phan, A. T.; Gueron, M.; Leroy, J. L. *Methods Enzymol.* **2001**, *338*, 341–371.

(39) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2007**, *35*, 6517–6525.
(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(41) Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570–2578.
(42) Plateau, P.; Gueron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310–7311.
(43) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65–73.
(44) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(45) DeLano, W. L. *The PyMOL User's Manual*; DeLano Scientific: Palo Alto, CA, 2002.
(46) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(47) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.
(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.