Conformational Transition of a Hairpin Structure to G-Quadruplex within the WNT1 Gene Promoter

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · DECEMBER 2014

Impact Factor: 12.11 · DOI: 10.1021/ja5089327 · Source: PubMed

7 AUTHORS, INCLUDING:

Zi-Fu Wang
Academia Sinica
13 PUBLICATIONS **80** CITATIONS
SEE PROFILE

Ting-Yuan Tseng
Academia Sinica
8 PUBLICATIONS **44** CITATIONS
SEE PROFILE

Shang-Te Danny Hsu
Academia Sinica
72 PUBLICATIONS **1,537** CITATIONS
SEE PROFILE

Ta-Chau Chang
Academia Sinica
103 PUBLICATIONS **1,198** CITATIONS
SEE PROFILE

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Conformational Transition of a Hairpin Structure to G-Quadruplex within the *WNT1* Gene Promoter

Margaret Hsin-Jui Kuo,[†,∥] Zi-Fu Wang,[†,∥] Ting-Yuan Tseng,[†] Ming-Hao Li,[†] Shang-Te Danny Hsu,[‡] Jing-Jer Lin,[§] and Ta-Chau Chang*[,†]

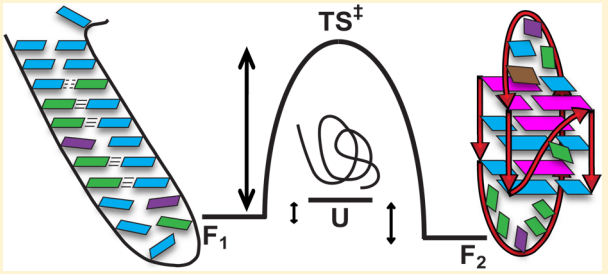
[†]Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 106, Taiwan, R.O.C.

[‡]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan, R.O.C.

[§]Institute of Biochemistry and Molecular Biology, National Taiwan University College of Medicine, Taipei 100, Taiwan, R.O.C.

Supporting Information

ABSTRACT: The role of G-quadruplexes (G4s) in biological systems has been widely studied. It is found that they have an important function in gene transcription and regulation. In this work, we have identified two topologies of hairpin and G4 structures formed by a native G-rich sequence (WT22: 5′-GGGCCACCGGGCAGGGGGCGGG-3′) from the *WNT1* promoter region using nuclear magnetic resonance (NMR) spectroscopy. With the help of site-specific isotope labeling, the topologies of these two structures are unambiguously characterized. Circular dichroism and NMR results are analyzed to determine the kinetics associated with the potassium ion-induced hairpin-to-G4 transition, which is very slow—on the time scale of 4800 s—compared to the previously reported folding kinetics of G4 formation. In addition, the free energies of the unfolding of these two structures are obtained using differential scanning calorimetry. Combining the kinetic and thermodynamic data, we have established the free energy landscape of this two-state folding system. Considering that similar conformational change may exist in other native G-rich sequences, this work highlights an important hairpin to G4 conformational transition which can be used in manipulation of gene regulation or ligand modulation *in vivo*.



INTRODUCTION

Single-stranded guanine (G)-rich DNA and RNA sequences are capable of forming G-quadruplexes (G4s) via Hoogsteen hydrogen bonds.[1−4] Human telomeric G4 structures have been widely studied under physiological conditions both *in vitro*[5−8] and *in vivo*[9,10] because of their importance in protecting the ends of chromosomes and their potential as promising targets for cancer therapy.[4,11] In addition, a large number of putative G4 forming sequences have been found in the human genome through bioinformatic searches.[12,13] Particularly, G4 structures are experimentally identified in the promoter regions of many oncogenes, such as BCL2,[14] VEGF,[15] KRAS,[16] *c*-MYC,[17] *c*-KIT,[18] and WNT1.[19] These G4 forming motifs are found to play important roles in gene transcription and regulation. Because of their functional importance, a large number of G4 stabilizers have been developed to regulate the expression levels of these oncogenes.[11,20,21]

Due to the importance and polymorphic nature of G4 structures, many experiments have been focused on identifying the existence of G4 structures *in vivo*.[9,10] Much effort has been focused on the study of the folding stabilities of G4 forming DNA and RNA sequences.[22−25] However, the folding/unfolding kinetics of these G4 motifs have been less documented,[26−29] which evidently plays an integral part in the regulations of biological activities. Of great interest is that the kinetics and thermodynamics of G4 formation have been found to be cation-dependent.[1,3,23] The problem is further complicated by the fact that some DNA and RNA sequences can coexist in hairpin and G4 structures.[30,31] To our knowledge, no report has been documented for the kinetics and thermodynamics study of conformational transition between a hairpin and a G4 structure for a genomic G-rich sequence.

Recently, a G4-forming sequence, 5′-$G_5CCACCG_3CAG_5CG_5$-3′, located at the −193 to −167 region of *WNT1* promoter has been identified.[19] Importantly, the Wnt1-mediated migration and invasion activities of cancer cells can be inhibited by G4 stabilizers via interaction with G4 structures of *WNT1*. In this work, we illustrate a conformational transition from a hairpin structure to a G4 structure upon addition of K^+ within the G-rich sequence located in the promoter region of the *WNT1* gene. Considering the highly polymorphic nature of the WNT1 G4s due to three runs of five consecutive G-bases, a shorter sequence, 5′-$G_3CCACCG_3CAG_5CG_3$-3′ (WT22, Figure 1A), is used to prevent the formation of intermolecular G4s since they are unlikely to form in living cell. We have used NMR spectroscopy

Received: August 29, 2014

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

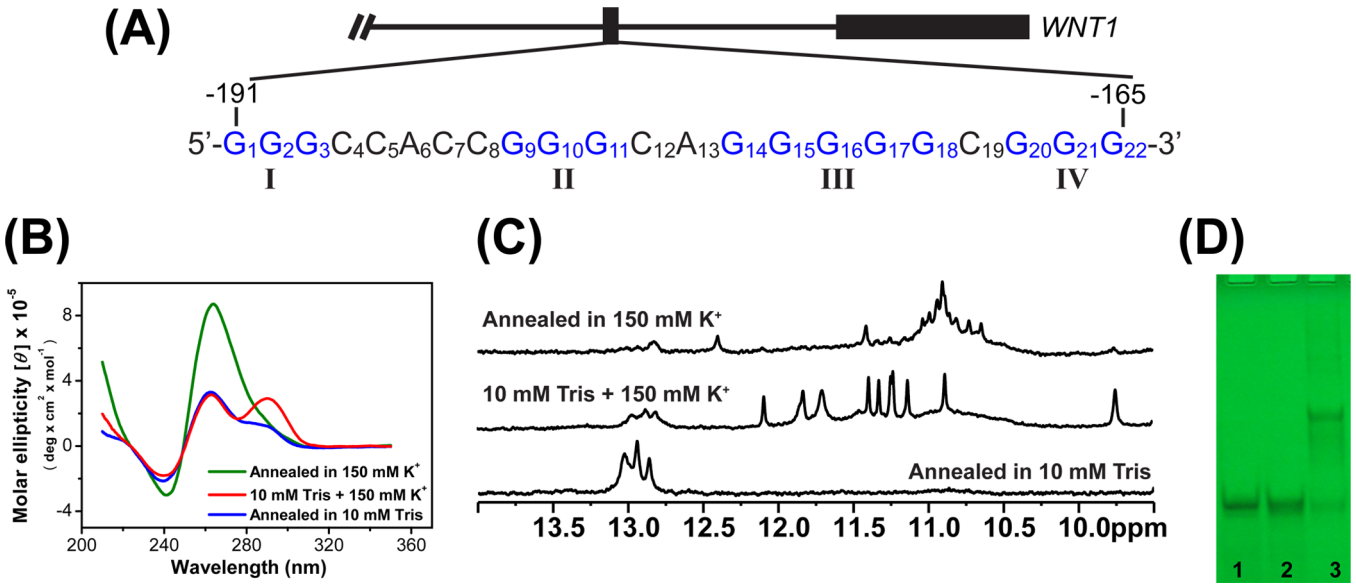


Figure 1. (A) A G4-forming sequence located at Wnt1 gene promoter named WT22. The CD profile (B) and NMR spectrum (C) of WT22 annealed in 10 mM Tris buffer, with the addition of 150 mM K^+ for 90 min at 37 °C, and annealed in 150 mM K^+ at DNA concentration of 4 μM for CD and 100 M for NMR. (D) UV shadowing of gel assay of WT22 sequences in 10 mM Tris buffer (lane 1), with addition of 150 mM K^+ for 90 min at 37 °C (lane 2), and annealed in 150 mM K^+ (lane 3) at DNA concentration of 80 μM.

to determine the Watson−Crick base pairings of the hairpin structure of WT22 in Tris buffer solution as well as the Hoogsteen hydrogen-bonding patterns of the G4 structure induced by the addition of K^+. It is important to obtain the structural information on the individual states involved in the conformational transition to better elucidate the underlying mechanism of such structural change.[32] Kinetic analysis via CD and NMR shows that the folding kinetics of WNT1 G4 is much slower than that of telomeric G4 upon addition of K^+. The WNT1 transition from hairpin to G4 is temperature-sensitive, which is associated with the unfolding of the hairpin structure. Finally, a simple energy diagram of the transition from the hairpin to G4 is proposed.

EXPERIMENTAL SECTION

DNA Preparation. All oligonucleotides were purchased from Bio Basic (Ontario, Canada) without further purification. The oligonucleotides were dissolved in 10 mM Tris-HCl (pH 7.5), followed by heat-denaturation at 95 °C for 10 min and slowly annealed to room temperature at a rate of 1 min/°C. The annealed oligonucleotides were stored at 4 °C overnight before experiments. The DNA concentrations were determined using a UV−vis absorption nanophotometer (Implen, Germany). A stock solution of 3 M KCl (Sigma-Aldrich, USA) was added to the DNA samples to reach a final K^+ concentration of 150 mM. The samples were kept at 37 °C for 90 min before experiments. The site-specific ^{15}N labeled oligonucleotides were synthesized using a solid-phase oligonucleotide synthesizer (Dr. Oligo, USA) as previous described.[32,33] The synthesized oligonucleotides were further purified by HPLC.

Circular Dichroism (CD) Spectroscopy. The CD spectra were recorded using a spectropolarimeter (J-815, Jasco, Japan) with a bandwidth of 2 nm, at a scan speed of 50 nm/min and a step resolution of 0.2 nm to scan over the spectral range of 210−350 nm. The DNA sample concentrations were 4 μM in the specified buffer solution. The time-resolved CD spectra of each sample were recorded at a set temperature controlled by a Peltier thermal coupler chamber (PFD-425S/15, Jasco, Japan). The instrument was purged with nitrogen flow during data collection.

Polyacrylamide Gel Electrophoresis (PAGE). PAGE experiment was carried out using 20% polyacrylamide gel, and electrophoresis of the gel was conducted at 25 mA for 3.5 h at 4 °C. The gel was then photographed under ultraviolet (UV) light at 254 nm using a digital camera.

Nuclear Magnetic Resonance (NMR) Spectroscopy. All NMR experiments were performed on Bruker AVIII 500 MHz and AVIII 800 MHz spectrometers equipped with a prodigy and a cryo-probe, respectively. One-dimensional (1D) imino proton NMR spectra were recorded using a WATERGATE[34] or jump-return pulsed sequence.[35] The ^{15}N-edited 1D heteronuclear multiquantum correlation (HMQC) spectra[36] were used to assign individual imino proton resonances using the site-specifically ^{15}N-enriched (8%) guanine base at one of the 11 G-tetrad-forming guanine residues. The strand concentrations of the NMR samples were typically 0.1−1 mM containing 10% D_2O in Tris-HCl or specific K^+ salt conditions with an internal reference of 0.1 mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid). Resonances of exchange protons were assigned using 2D nuclear Overhause effect spectroscopy (NOESY)[37] (mixing time of 250 ms). Total correlation spectroscopy (TOCSY)[38] (mixing times of 50 and 150 ms) spectra were used to cross-check the assignments of the NOEs of cytosine H5−H6. Through-bond correlations at natural abundance (H1/H8−C5) and heteronuclear multiple-bond correlation (JR-HMBC)[39] were used to assign aromatic proton (H8). For the hydrogen−deuterium exchange (HDX) experiments, the oligonucleotides were lyophilized after being prepared in the desired conditions and then were resuspended in D_2O (99%) immediately before NMR measurements to reach a DNA concentration of 100 μM.

Differential Scanning Calorimetry (DSC). An N-DSC III calorimeter (New Castle, DE, USA) was used to obtain the DSC thermographs. The data acquisition and analysis were carried out using the built-in software (NanoDSC Run version 4.1.12 and NanoAnalyze version 2.3.6). Each calorimetric experiment was performed by scanning sample prepared in desired buffer solution at a concentration of 200 to 250 μM from 10 to 115 °C at a heating rate of 1.0 °C/min. The corresponding buffer−buffer baseline scans were subtracted from the buffer-sample scans before normalization and analysis.

RESULTS

WT22: A G4-Forming Sequence. CD is a commonly used tool to examine the existence of G4 structures. The different spectral features result from the different modes of guanine stacking due to various G4 topologies. For example, two

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

positive CD bands at 295 and 240 nm with a negative band at 265 nm is typically the CD profile of antiparallel G4 structures such as the basket and the chair forms, while a positive band at 265 nm accompanied by a negative band at 240 nm are attributed to a parallel G4 structure such as the propeller form.[40−42] The CD spectrum of WT22 annealed in 10 mM Tris shows a major peak at 260 nm with shoulder at ∼285 nm (Figure 1B). Upon addition of 150 mM K^+ for 90 min, the peak at ∼285 nm shifts to 290 nm, and the intensity increases dramatically. A sample of WT22 annealed in 150 mM K^+ shows a very different CD profile with only a major peak at 264 nm.

NMR spectroscopy was further conducted to unambiguously differentiate the various structures formed by WT22. In general, the imino proton NMR spectrum shows the typical chemical shifts of the Watson−Crick base pairing at around 13 ppm for guanine imino protons (^{1}H),[43] while the chemical shifts of the imino protons of Hoogsteen G-G pairings of a G4 structure are located between 10 and 12 ppm.[44,45] In the absence of K^+, the imino proton NMR spectrum of WT22 in 10 mM Tris buffer shows several well-defined resonances at around 13 ppm (Figure 1C), indicating the presence of hairpin-like structure in Tris buffer. After addition of 150 mM K^+, the imino proton resonances at around 13 ppm diminish, and distinct G4 imino proton resonances in the range of 10−12 ppm appear concomitantly, indicating the formation of intramolecular G4 of WT22. However, the proton NMR spectrum of annealed WT22 in 150 mM K^+ exhibit some weak but distinct resonances on top of a broad, undefined imino proton resonances between 10 and 12 ppm, implying the coexistence of multiple G4 structures of WT22 including possible formation of intermolecular G4.

To verify the formation of intermolecular G4s of WT22 induced by annealing, PAGE assay was conducted. Figure 1D shows the UV shadowing of WT22 (lane 1−3) in 10 mM Tris (lane 1), after addition of 150 mM K^+ (at 37 °C for 90 min; lane 2), and annealed in 150 mM K^+ (lane 3). A major single band is clearly visible for WT22 in 10 mM Tris (lane 1) and after addition of 150 mM K^+ at 37 °C for 90 min (lane 2), while WT22 that was annealed in the presence of 150 mM K^+ reveals a major formation of dimer structure with minor formation of intramolecular G4 (lane 3). Therefore, the experiments hereafter were all carried out using the WT22 sequence in the presence of 150 mM K^+ at 37 °C for 90 min without annealing. We hypothesize that the spectral change of WT22 is due to the structural conversion from a hairpin-like structure in Tris buffer to an intramolecular G4 structure upon addition of K^+. We therefore set out to determine the Watson−Crick pairings of the hairpin-like structure formed by WT22 in Tris buffer solution.

Determination of Hairpin Structure of WT22. The imino proton resonances at 13 ppm that correspond to Watson−Crick base pairings are observed for the WT22 sequence, which most likely originate from the base pairing between the segment of $C_4C_5A_6C_7C_8$ with the consecutive five guanines, G_{14} to G_{18}, with an A (adenine)-G mismatch at the middle. Here, a number of mutants were designed to further verify the G-C pairings within the sequence (Table 1). To eliminate the G-C pairings for Watson−Crick signals, pairs of guanines ($G_{14}G_{15}$ or $G_{17}G_{18}$) were replaced by a pair of thymines; alternatively, pairs of cytosines (C_4C_5 or C_7C_8) were replaced by a pair of adenines. These mutants are named M1, M2, M3, and M4, respectively. We further designed three additional sequences that contain additional G-C pairings by replacing A_{13} with C_{13} for M5, A_6 with C_6 for M6, or both A_6 and A_{13} with C_6 and C_{13} for M7. Figure S1A shows the respective NMR spectra of these mutants in 10 mM Tris buffer. Mutants M1, M2, M3, and M4 did not exhibit the Watson−Crick signals at around 13 ppm in 10 mM Tris buffer. On the other hand, mutants M5, M6, and M7 exhibited strong Watson−Crick signals at around 13 ppm, which are similar to that of WT22. Figure S1B shows the corresponding NMR spectra of these mutants in 10 mM Tris buffer after addition of 150 mM K^+ at 37 °C for 90 min. Because of the additional G-C pairing in mutants M5, M6, and M7, they did not form G4 even at 150 mM K^+, while mutants M1 to M4 exhibit quite different NMR spectra. It is well established that a change of a single base could result in different G4 structures due to the highly polymorphic nature of G-rich sequences as reflected in the very different NMR spectra.[46,47] In particular, the imino proton NMR spectrum of M1 G4 (in 150 mM K^+) most closely resembles that of WT22 G4, implying that G_{14} and G_{15} might not be involved in the formation of G-quartet.

Table 1. Sequences of Wild-Type *WNT1* Gene Promoter and Its Mutants Studied in This Work[a]

abbreviation	sequence
WNT1	5′-GGGGGCCACCGGGCAGGGGGCGGGGG
WT22	5′-GGGCCACCGGGCAGGGGGCGGG
M1	5′-GGGCCACCGGGCATTGGGCGGG
M2	5′-GGGCCACCGGGCAGGGTTCGGG
M3	5′-GGGAAACCGGGCAGGGGGCGGG
M4	5′-GGGCCAAAGGGCAGGGGGCGGG
M5	5′-GGGCCACCGGGCCGGGGGCGGG
M6	5′-GGGCCCCCGGGCAGGGGGCGGG
M7	5′-GGGCCCCCGGGCCGGGGGCGGG
M8	5′-GGGCCACCGGGCAGTGGGCGGG
M9	5′-GGGCCACCGGGTAGGGGGCGGG

[a]Underline indicates G-rich tracts.

To further obtain structural insights, we used the sequence-specific assignments of the imino proton resonances to identify the G-bases that contributed to the Watson−Crick signals at around 13 ppm and to pinpoint the exact Watson−Crick base pairings for the hairpin-like structure of WT22 in Tris buffer. Four WT22 samples were synthesized, each of which was site-specifically labeled with ^{15}N-enriched guanine. Figure 2A shows ^{15}N-edited imino proton spectra of WT22 for site-specific assignments of imino proton resonances of these four guanines. The ^{15}N-edited imino proton spectra of the guanines of G_{14}, G_{15}, G_{17}, and G_{18} indeed contribute to the signal at around 13 ppm. In addition, 2D-NOESY spectra show four correlating peaks (Figure 2B), confirming our hypothesis that the Watson−Crick signal at 13 ppm is due to these four G-C pairings in the sequence. This finding is consistent with the mutation analysis as described above and further supports our hypothesis that the G-C pairings in the sequence resulted in the formation of a hairpin structure. Figure 2C shows the hairpin structure of WT22 in Tris buffer.

Topology of WT22 G4. With the residue-specific imino proton NMR assignments, we can then determine the topology of WT22 G4. Upon addition of K^+ ion (Figure 3A), ^{15}N-edited imino proton spectra show the presence of the imino proton signals of G_{17} and G_{18} as well as other guanine imino protons in the range of 10−12 ppm. The absence of G_{14} and G_{15} implies that G_{14} and G_{15} in the consecutive five guanines of G_{14}-G_{18} are not involved in the formation of the G-quartets. This finding is

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

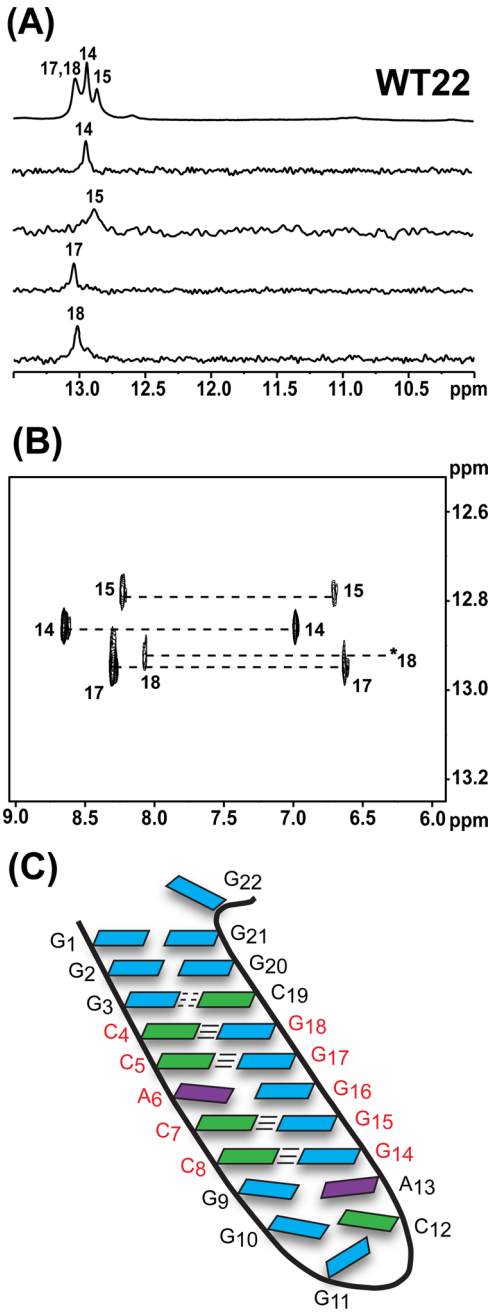


Figure 2. (A) Site-specific assignments of imino proton resonances of WT22 hairpin structure. The imino proton NMR spectrum of WT22 in the absence of K^+ ion is on the top, and the residue-specific assignments are labeled on top of individual peaks. The 1D ^{15}N-edited HMQC spectra of 8% ^{15}N enriched oligonucleotide samples are shown below with assignments to the labeling sites. All spectra were recorded at 25 °C in 10 mM Tris-HCl buffer. (B) NOESY spectrum of exchangeable proton (mixing time, 250 ms) at 25 °C was showing the cross-peaks that identified the arrangements of the G-C base pairs. (C) The hairpin structure of WT22.

consistent with the mutation analyses and enables us to propose an intramolecular G4 topology of WT22 upon addition of K^+. Based on the NOEs of imino−imino and imino−H8 protons (Figure 3B,C), we discover that WT22 G4 contains two G-quartets. The hydrogen-bonding directionality of G-quartets are clockwise ($G_2 \rightarrow G_{17} \rightarrow G_{21} \rightarrow G_{10}$ and $G_3 \rightarrow G_{18} \rightarrow G_{22} \rightarrow G_9$). However, there was a lack of NOE connectivity for the third G-quartet of G_1, G_{16}, G_{11}, G_{20} in imino−H8 NOESY spectrum, which warrant further clarification. We therefore recorded imino proton spectra at different temperatures (Figure 3D) and found that four of the imino peaks, G_1, G_2, G_{11} and G_{20}, are convoluted at room temperature and can be resolved upon increasing temperatures. Hence, it is likely that the upper layer is distorted by a possible G_{15}-C_{12} pairing on the second loop (Figure 3E).

To examine whether this G-C pairing on the second loop is responsible for the upper layer distortion, two mutants were designed (M8 and M9) to eliminate the G-C pairing on the loop. Figure 4A shows the NMR spectra of these mutants together with that of the wild type. M8 with a single base G_{15} replaced by T shows 12 distinct imino proton resonances without Watson−Crick signals at around 13 ppm, while M9 with a single base C_{12} replaced by T retains some of the Watson−Crick signals at around 13 ppm. The Watson−Crick signals of M9, which are similar to that of WT22, should correspond to the hairpin structure. Of importance is that the 12 distinct imino proton peaks of M8 indicate a single form of intramolecular G4. These peaks are quite similar to that of the wild type.

Furthermore, NMR HDX results indicate that both WT22 and M8 shared the same central G-tetrad topology that resulted in the same imino proton spectra under D_2O solution (Figure 4B). We therefore resolved the G4 topology of M8, in order to infer the topology of the wild type without the complication of the coexistence of the hairpin structure. The ^{15}N-edited imino proton spectrum of M8 (Figure S2A) along with the NOE connectivities of H1−H1, H1−H8 (Figure S2B,C), and H8/H6−H1′ (Figure S3A), as well as the scalar coupling connectivities observed in JR-HMBC (Figure S3B) and TOCSY (Figure S3C) helped to establish the topology of M8 unambiguously. In addition, the glycosidic conformations are confirmed by the H1′−H8 NOE patterns (Figure S3D) where the top G-quartet is syn-syn-anti-syn, while the other two are anti-anti-syn-anti. The syn guanine bases are G_1, G_9, G_{10}, G_{16}, and G_{20} as evidenced by the strong H1′−H8 NOEs, compared to the H5−H6 NOE of cytosine (Figure S3C).

We therefore conclude that the G4 topology of M8 involves three G-quartets, the top layer consists of $G_1 \rightarrow G_{11} \rightarrow G_{20} \rightarrow G_{16}$ with clockwise hydrogen-bonding directionality, while the middle and bottom layers consist of $G_2 \rightarrow G_{17} \rightarrow G_{21} \rightarrow G_{10}$ and $G_3 \rightarrow G_{18} \rightarrow G_{22} \rightarrow G_9$, respectively, with anticlockwise hydrogen-bonding directionality (Figure S3E). Moreover, the asymmetric hydrogen-bonding directionalities between the top and central layers are in consistent with H1−H1 NOEs of G_{11}-G_{21}, G_{16}-G_2, G_{20}-G_{17}, and G_1-G_{10}. As such, the three connecting loops adopt different conformations that correspond to a lateral loop, a diagonal loop, followed by a double-chain reversal loop. Together, we establish the topology of M8 to be the (3+1) G4 with three different loops (Figure S3F).

Comparing the NMR results of WT22 and M8 G4s further confirms that the G4 topology of WT22 wild type is similar to that of M8, except for the top G-quartet. This difference is due to distortion of the top quartet by the G-C pairing on the diagonal loop of WT22. An important conclusion that can be drawn thus far is that the native genomic DNA sequence within the promoter region of the *WNT1* gene can simultaneously adopts two distinct higher order structures, namely a hairpin structure and a G4 structure.

Kinetics of Structural Conversion of WT22 DNA. We then studied the folding kinetics of WT22 G4 upon addition of

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

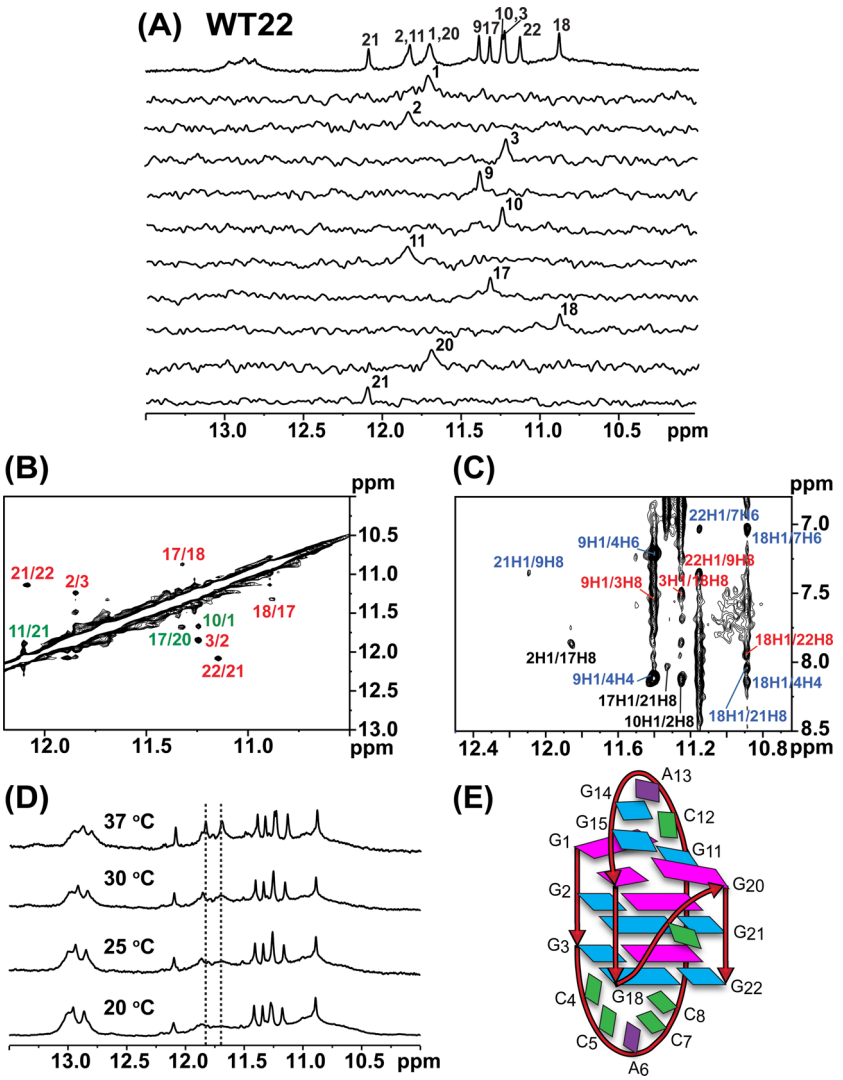


Figure 3. (A) Site-specific assignments of imino proton resonances of WT22 G4 structure. The imino proton NMR spectra of WT22 G4 in the presence of K^+ ion on the top, and the residue-specific assignments are labeled on top of individual peaks. The 1D ^{15}N-edited HMQC spectra of 8% ^{15}N-enriched oligonucleotide samples are shown below with the assignments and site-specific labels that corresponds to labeling sites. (B) The 2D NMR spectra of WT22 G4. The NOESY spectrum of exchangeable (250 ms mixing time) was showing the H1−H1 correlations. Green and red labels correspond to inter- and intrastrand correlations, respectively. (C) The NOESY spectrum (250 ms mixing time) was showing the H1−H8 correlations. Black, red, and blue labels correspond to the middle, the bottom, and interquartet correlations, respectively. (D) The imino proton NMR spectra of WT22 in 150 mM K^+ solution at different temperatures. (E) Topology of the WT22 G4 structure with the distorted top layer.

K^+. CD spectroscopy was used to determine the folding rate (Figure 5A). The increase of the 290 nm CD band of WT22 upon the addition of 150 mM K^+ suggests the formation of a G4 structure. The change of CD signal at 290 nm is fitted to a single exponential function with a time constant of 70 ± 3 min (Figure 5B). Such folding rate is significantly slower than that of a variety of human telomeric sequences as well as those of the other G4 forming gene promoter sequences. For example, the G4 folding kinetics of human telomeric sequences upon addition of 150 mM K^+ is on the time scale of milliseconds to seconds.[26,27]

Given the slow folding kinetics, imino proton NMR spectra of WT22 were recorded in real time to monitor the loss of the hairpin imino proton resonances, and the increase of the G4 imino proton resonances, upon addition of K^+ (Figure 5C). Although the decay rate of the hairpin signals is slightly faster than the arising rate of the G4 signals, the time constants of the two processes are very close (ca. 80 min; Figure 5D), which is consistent with the CD-derived reaction time constant. The results suggest that the unfolding rate of the hairpin structure upon addition of K^+ could be the rate-determining step of the formation of WT22 G4.

To further illustrate that the unfolding of the hairpin structure is the rate-determining step for the formation of WT22 G4, we measured the time of G4 formation in mutants without the Watson−Crick base pairings (Table S1). Mutants M1 to M4 lack the G4 imino proton NMR signals at around 13 ppm in Tris buffer (Figure S1A). These mutants also exhibit much faster folding rates (<2 min, Table S1). In contrast, M5 and M6, which harbors an additional G-C pairing, have a significantly longer reaction time constant. Moreover, M7 contains additional G-C pairings, and its G4 formation kinetic is markedly slower, with a reaction time constant of greater than 700 min (nearly 10-fold longer than that of wild type). These results further confirm that the Watson−Crick base-pairing from the secondary hairpin structure is the reason for

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

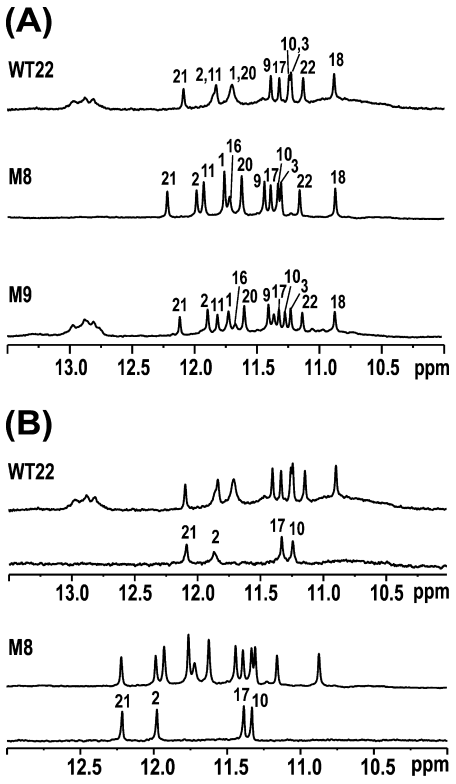


Figure 4. (A) Imino proton NMR spectra of WT22, M8, and M9 in the presence of K^+ ion are shown on the top, middle, and bottom, respectively. (B) Imino proton spectra of WT22 and M8 G4 after 4 and 12 h, respectively, in D_2O. All spectra were recorded at 37 °C in 10 mM Tris plus 150 mM K^+ solution.

the slow formation of WT22 G4 upon addition of 150 mM K^+, and the rate-determining step is the unfolding of this hairpin structure.

To determine the energy barrier of the structural conversion of WT22 from hairpin to G4, we conducted the folding kinetics analysis of WT22 at seven temperatures, ranging between 25 to 65 °C (Table S2). The folding kinetics of WT22 G4 follows a typical Arrhenius behavior with an activation energy $E_a \approx 25.6$ kcal/mol (Figure 5E), according to the definition of the Arrhenius plot, where $\ln(k) = \ln(A) - E_a/RT$, of which R is the gas constant (1.987 cal/mol·K) and T is the sample temperature. This is different from the anomalous curvatures in Eyring plots observed for the folding of DNA and RNA G4s,[26,27] where folding rates are slower at higher temperature. Thus, the existing activation energy of WT22 serves as a proof for the presence of a secondary hairpin structure in Tris, and the unfolding of hairpin structure is prerequisite for the formation of G4 structure upon addition of 150 mM K^+.

Thermal Stability and Proposed Simple Energy Diagram. Here, we have verified that the WT22 sequence derived from the promoter sequence of the *WNT1* gene can form a hairpin structure in Tris buffer (a folded state, F1; Scheme 1). This hairpin structure contains four G-C pairings (black solid lines) of C_4 to G_{18}, C_5 to G_{17}, C_7 to G_{15}, and C_8 to G_{14}. It seems that G_3 may also pair with C_{19} (black dashed lines, Figure 2C); however, 2D NOSEY shows four correlating peaks (Figure 2B). According to the mutational experiments, this hairpin structure ceases to exist in Tris buffer when two bases of the sequence of G_{14} and G_{15} were changed (Figure S1A). Upon addition of 150 mM K^+, the hairpin structure of WT22 starts to unfold to an unfolded state (U) with an activation energy barrier of ~25.6 kcal/mol and takes ~70−80 min to convert to the G4 structure (F2).

We further used DSC[22,23] to determine the melting temperature (T_m) and the free energy at 37 °C (ΔG_{310}) associated with thermal unfolding of the WT22 hairpin structure and WT22 G4 (Figure S4A). The T_m of WT22 hairpin structure in Tris (black line) is 67.2 ± 0.3 °C, with an enthalpy ΔH = 43.2 ± 0.2 kcal/mol and an associated entropy ΔS = 126.9 ± 0.5 cal/mol·K, and thus the free energy of unfolding, ΔG = 3.8 ± 0.1 kcal/mol at T = 310 K (Table S3). Upon addition of 150 mM K^+ to WT22, the DSC thermograph exhibits a biphasic transition (Figure S4A). It indicates the formation of higher order or intermolecular structures during the heating process, which is similar to the annealing process during sample preparation. Such a behavior prohibits the determination of the exact free energy of WT22 G4 unfolding in the presence of 150 mM K^+. Nonetheless, our NMR results show that M8 adopts a similar intramolecular G4 topology as WT22 wild type. Moreover, the DSC show similar T_m for the intramolecular G4 of WT22 (~77 °C), and M8 (~78 °C). We therefore used the DSC thermograph of M8 (red line) to estimate the free energy of M8 G4 unfolding, which should be close to that of WT22 wild type. The T_m of M8 is 78.1 ± 0.9 °C, with enthalpy ΔH = 46.8 ± 4.2 kcal/mol, which gives the associated entropy ΔS = 133.4 ± 11.7 cal/mol·K, and the free energy of unfolding of M8 G4 ΔG = 5.5 ± 0.6 kcal/mol. Furthermore, we measured DSC thermograph of mutant, M1, where it also exhibits similar T_m as well as ΔG_{310} (Figure S4B, Table S3). In addition, the biphasic transition of WT22 in 150 mM K^+ is deconvoluted to estimate the free energy of intramolecular G4 unfolding, which is 5.3 ± 0.2 kcal/mol. Thus, we believe that the intramolecular G4 of WT22 should also have similar free energy of unfolding as M1 and M8 in the range of 5.3−5.6 kcal/mol. Therefore, a simple energy diagram is proposed, as shown in Scheme 1. The G4 structure formed by WT22 is thermodynamically more stable than its hairpin structure assuming both structures unfold to a similar random coil.

In addition, we have examined the ionic effect by the addition of Na^+ in Tris buffer. The DSC results show that the addition of 150 mM Na^+ increases the melting temperature of WT22 to 78 °C due to electrostatic interaction (Figure S4C). We obtain an enthalpy ΔH = 40.3 ± 4.2 kcal/mol and its associated entropy ΔS = 114.8 ± 1.2 cal/mol·K, and thus the free energy of unfolding ΔG_{310} = 4.7 ± 0.5 kcal/mol at T = 310 K in the presence of 150 mM Na^+ (Table S3). Of interest is that imino proton NMR spectra show no spectral change from hairpin to G4 of WT22 at 37 °C (Figure S4D), indicating that the addition of 150 mM Na^+ stabilizes the hairpin structure of WT22 and does not induce structural conversion. In contrast, NMR spectra of WT22 in 150 mM Na^+ solution show the structural conversion from the hairpin structure to the intramolecular G4 structure after addition of 150 mM K^+ (Figure S4D).

DISCUSSION

Much effort has been devoted to tackle the Wnt1 signal transduction pathway, since the Wnt1 pathway is recognized to play a major role in cancer progression.[48,49] In the previous study,[19] we found that the WNT1 G-rich sequence at the upstream region of *WNT1* promoter is capable of forming G4 structure. This G4 structure can be stabilized by the carbazole

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

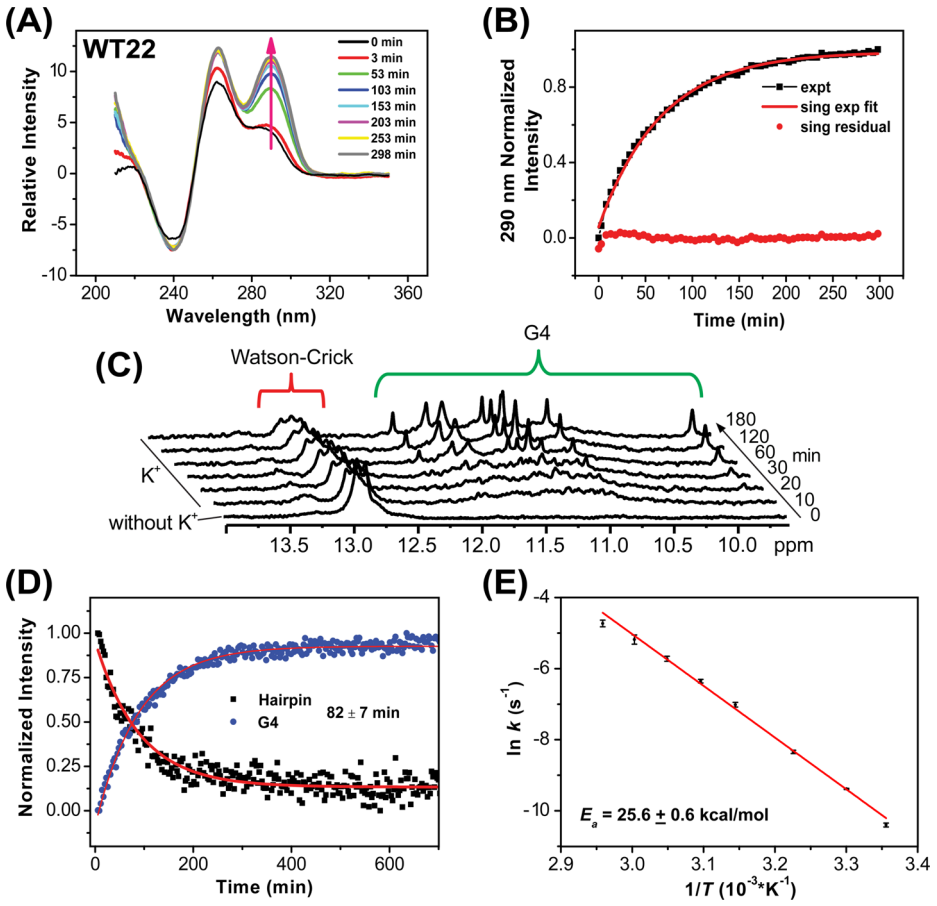
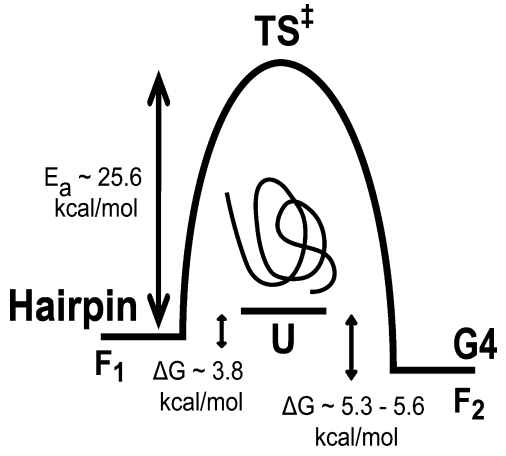


Figure 5. (A) CD spectra of WT22 after adding 150 mM K^+ revealing the formation of G4 by the increasing intensity of 295 nm peak. (B) The arising time for folding of the G4-forming WT22 obtained as 70 ± 3 min. (C) Time-resolved imino proton spectra of WT22 recorded at 0, 10, 20, 30, 60, 120, and 180 min after the addition of 150 mM K^+. (D) The decay of NMR hairpin signal and arising of G4 signal were fitted with 1st exponential to obtain a rate of 82 ± 7 min. (E) Arrhenius plot of the G4 formation rate against temperature with a slope of 25.6 ± 0.6 kcal/mol as the energy barrier for the unfolding of hairpin structure.

Scheme 1. Simple Energy Diagram of WT22 Hairpin Conformational Transition to G4



derivatives, BMVC and BMVC4, which can repress the expression of *WNT1* and downstream genes such as *β*-catenin, MMP7, and survivin.[19] Here, we have shown that a native WT22 G-rich sequence within the promoter region of *WNT1* gene could form two distinct and different structures: a hairpin and a G4 upon addition of K^+. It is known that a DNA sequence that contains tracks of four or more consecutive guanines can form multiple G4 structures including dimer, tetramer, and even G-wire, which makes the determination of intramolecular topology of G4s more complicated. Indeed, we have identified multiple conformations of the 26-mer WNT1 sequence, which contains three tracks of five consecutive guanines, in the gel electrophoresis assays (Figure S5). Even the shorter sequence of WT22 with one track of five consecutive guanines can form intermolecular G4 after annealed in 150 mM K^+ (Figure 1D). In comparison, a single band is clearly observed for M8, which contains four tracks of three consecutive guanines, even after annealed in 150 mM K^+ (Figure S5). In this work, solution state NMR spectra show that the addition of 150 mM K^+ without annealing can avoid formation of intermolecular conformations in WT22 (Figure 1D). This finding is important to the study of G4 in living cell since high-temperature annealing cannot take place under physiological conditions. In addition, the formation of intermolecular WT22 G4s is unlikely in living cell. Thus, the WT22 wild type is selected for the study of its topology with careful design of sample preparation and experimental conditions to avoid the formation of intermolecular G4s.

The absence of G_{16} from the isotope labeling NMR spectra suggests that one end of G-quartet of WT22 is distorted by the C_{12}-G_{15} pairing in the diagonal loop (Figure 3E). Lim and Phan[50] reported that a G4 structure containing a hairpin stem on the loop could alter the junction between quadruplex and

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

hairpin. Therefore, it is likely that the C_{12}-G_{15} pairing in the diagonal loop of WT22 could distort the Hoogsteen hydrogen-bonding of the top G-quartet. This is supported by the evidence that three G-quartets of M8 G4 characterized by 12 distinct imino proton peaks, where the G_{15} in WT22 is replaced by T in M8. Since M8 G4 lacks the C_{12}-G_{15} paring, it forms a (3+1) G4 structure with three well-defined G-quartets. In this work, similar imino proton signals between WT22 and M8 together with the identical four guanines in the middle G-quartet between WT22 and M8 G4s detected in the HDX confirm that WT22 and M8 have the same type of G4 topology. The similarity between WT22 and M8 also supports that the G4 topology of a native sequence can be mimicked by the study of mutant with careful design.

We have identified the topology of WT22 G4, which adopt (3+1) topology with a lateral loop of CCACC, a diagonal loop of CAGG, followed by a single C base double-chain reversal loop to connect the four G-tracks. To our knowledge, such G4 topology has not been documented for a native G-rich sequence. Plavec and co-workers demonstrated a G4 structure of $d[G_3ATG_3ACAACG_4ACG_3]$ with 2:5:3 base loops including a double-chain reversal loop of AT, a diagonal loop of ACAAC, followed by a lateral loop of GAC.[51] The differences in the order of loop orientation between WT22 and reported structure of $d[G_3ATG_3ACAACG_4ACG_3]$ are very similar to that of (3+1) hybrid I and hybrid II G4s of human telomeric sequences.[6,7]

Hardin *et al.*[30] reported a cation-dependent equilibrium between an intramolecular hairpin structure and a tetrameric G4 structure for a G-rich DNA sequence (5′-CGCGGGCGC-3′). The presence of K^+ strongly favors the G4 formation. Recently, Balasubramanian and co-workers reported a designed RNA sequence that exists in equilibrium between a hairpin structure and a G4 structure.[31] The hairpin and G4 populations can be regulated by Mg^{2+} and K^+ as well as a small molecule that favors the hairpin structure. In this work, we present a transition from a hairpin to a G4 for a native G-rich sequence of WT22. Moreover, we characterize the kinetics of K^+-induced hairpin-to-G4 transition by NMR. The transition for the G4 formation is slow (~4800 s), and the rate-determining step for the G4 formation of WT22 is the slow unfolding of the hairpin structure. In comparison, the mutants of WT22 without formation of hairpin structure show fast G4 formation (<120 s). Using stopped-flow method, Gary and Chaires[26] observed a single folding time of 20−60 ms for K^+-induced G4 folding of three oligonucleotides including Tel22 ($AG_3[TTAG_3]_3$) and proposed a two-step folding model. Zhang and Balasubramanian[27] also found fast folding time within tens of millisecond to second for the formation of DNA and RNA G4s. They further proposed a multi-step folding model via hairpin and triplex intermediates. It is believed that a single-stranded telomeric sequence can rapidly form G4 in the presence of K^+. Recently, we used time-resolved NMR experiments of HDX and hybridization to measure the unfolding of Tel23 ($TAG_3[TTAG_3]_3$) G4 and found that the unfolding time is ~900 s for the G4 in 150 mM Na^+ solution and ~8000 s for the G4 in 150 mM K^+ solution.[32] Gary *et al.*[29] observed three relaxation times, 0.1, 3700, and 750 s, and proposed a new model, a four-step pathway $U \leftrightarrow I_1 \leftrightarrow I_2 \leftrightarrow I_3 \leftrightarrow F$, where U and F represent unfolded and folded states while I_1, I_2, and I_3 are intermediates, for describing the kinetics of folding and unfolding pathways of Tel22 G4. Together with their previous two-step folding model,[29] they assigned fast relaxation times of <5 ms for $U \rightarrow I_1$ (hairpin) and <1 s for $I_1 \rightarrow I_2$ (antiparallel G4), as well as slow relaxation times of 3700 s for $I_2 \rightarrow I_3$ (triplex intermediate) and 750 s for $I_3 \rightarrow F$ (hybrid G4) for the coexistence of two G4 topologies. Since the coexistence of two G4 structures of Tel22 has not been well characterized, this model deserves further study. Nevertheless, their results indicated that the process of unfolding G4 takes time, even for partially unfolding to a triplex intermediate. Thus, the unfolding of hairpin structure plays a critical role in the folding kinetics of structural transition, and the unfolding time can be set as an upper limit time window for such structural transition.

We further examined the difference in ionic effects between Na^+ and K^+ to WT22, where K^+ induces structure change from hairpin to G4, while Na^+ stabilizes the hairpin structure. According to DSC, the free energy (ΔG_{310}) of the hairpin structure is 3.8 kcal/mol in Tris buffer with a melting temperature (T_m) of 67 °C; the addition of 150 mM Na^+ increased ΔG_{310} to 4.7 kcal/mol and T_m to 78 °C; the corresponding ΔG_{310} and T_m values are 5.5 kcal/mol and 78 °C, respectively, for G4 structure in 150 mM K^+ solution. In the case of the M1 G4, the T_m values are 59 and 79 °C when incubated in 150 mM Na^+ solution and in 150 mM K^+ solution, respectively (data not shown), implying that the WT22 hairpin is more stable than the WT22 G4 in 150 mM Na^+ solution. This may explain why addition of 150 mM Na^+ only stabilizes the WT22 hairpin without converting to WT22 G4.

Recently, Hurley and co-workers demonstrated that the addition of small molecules can modulate the equilibrium of a promoter sequence of BCL2 between forming an i-motif and a hairpin structure.[52] The existence of an i-motif and hairpin structures in the proximal region of BCL2 promoter gave insight into the possible relationship between dynamics of structural conversion and regulation of gene expression.[52,53] Therefore, it will be very interesting to apply our two-state folding system to verify the kinetics and thermodynamics of transition between i-motif and hairpin, especially to further understand the effects of the ligand and respective protein (*hnRNP LL*) on the folding kinetics.

CONCLUSION

In summary, we illustrate conformational change of a native G-rich sequence from a hairpin structure to a G4 structure upon addition of K^+. The folding kinetics of WT22 G4 is very different from that of telomeric G4 upon addition of K^+ because of the hairpin structure formed in Tris buffer. In addition, a thermal energy diagram with the assistance of mutated sequences is established to depict the thermodynamics of this conformational transition. According to the study of RNA hairpin to G4 interconversion,[31] it is suggested that the conformational transition between two distinct structures within a native G-rich sequence may involve a novel mechanism in gene regulation or transcription. Moreover, hairpin structure in promoter region has been shown to be a target for some transcriptional factor and enzyme.[54,55] Furthermore, our results show that the energy barrier for the unfolding of hairpin structure is responsible for the exceedingly slow folding kinetics of WT22 G4 in comparison with other G4 structures that have been investigated so far.[26,27,29] Thus, this large kinetic barrier of spontaneous conformational transition could play a critical role in biological function and opens up additional channel in regulating gene expression by manipulating the hairpin to G4 transition.

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

■ ASSOCIATED CONTENT

Supporting Information

Additional experimental data and analyses (Figures S1−S5 and Tables S1−S3). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

tcchang@po.iams.sinica.edu.tw

Author Contributions

[‖]M.H.-J.K. and Z.-F.W. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Academia Sinica (AS-102-TP-A07) and the National Science Council of the Republic of China (Grant NSC-101-2113-M001-022). We thank Dr. Shing-Jong Huang (Instrumentation Center, National Taiwan University) for assistance in obtaining the Bruker AVIII 500 & 800 MHz FT-NMR data and Dr. Shu-Chuan (Chris) Jao (Biophysical Core Facility, Academia Sinica) for assistance on the DSC experiments.

■ REFERENCES

(1) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703.
(2) Mergny, J. L.; Helene, C. *Nat. Med.* **1998**, *4*, 1366.
(3) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668.
(4) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345.
(5) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263.
(6) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963.
(7) Phan, A. T.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2006**, *34*, 5715.
(8) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723.
(9) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. *Nat. Chem.* **2013**, *5*, 182.
(10) Tseng, T. Y.; Wang, Z. F.; Chien, C. H.; Chang, T. C. . *Nucleic Acids Res.* **2013**, *41*, 10605.
(11) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discov.* **2011**, *10*, 261.
(12) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901.
(13) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908.
(14) Dexheimer, T. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2006**, *128*, 5404.
(15) Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H. *Nucleic Acids Res.* **2005**, *33*, 6070.
(16) Cogoi, S.; Xodo, L. E. *Nucleic Acids Res.* **2006**, *34*, 2536.
(17) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593.
(18) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2005**, *127*, 10584.
(19) Wang, J. M.; Huang, F. C.; Kuo, M. H.; Wang, Z. F.; Tseng, T. Y.; Chang, L. C.; Yen, S. J.; Chang, T. C.; Lin, J. J. *J. Biol. Chem.* **2014**, *289*, 14612.
(20) Grand, C. L.; Han, H.; Munoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Mol. Cancer Ther.* **2002**, *1*, 565.
(21) McLuckie, K. I.; Waller, Z. A.; Sanders, D. A.; Alves, D.; Rodriguez, R.; Dash, J.; McKenzie, G. J.; Venkitaraman, A. R.; Balasubramanian, S. *J. Am. Chem. Soc.* **2011**, *133*, 2658.
(22) Olsen, C. M.; Gmeiner, W. H.; Marky, L. A. *J. Phys. Chem. B* **2006**, *110*, 6962.
(23) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482.
(24) Bugaut, A.; Balasubramanian, S. *Biochemistry* **2008**, *47*, 689.
(25) Zhang, A. Y.; Bugaut, A.; Balasubramanian, S. *Biochemistry* **2011**, *50*, 7251.
(26) Gray, R. D.; Chaires, J. B. *Nucleic Acids Res.* **2008**, *36*, 4191.
(27) Zhang, A. Y. Q.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19297.
(28) Li, M. H.; Wang, Z. F.; Kuo, M. H.; Hsu, S. T.; Chang, T. C. *J. Phys. Chem. B* **2014**, *118*, 931.
(29) Gray, R. D.; Trent, J. O.; Chaires, J. B. *J. Mol. Biol.* **2014**, *426*, 1629.
(30) Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. *Biochemistry* **1992**, *31*, 833.
(31) Bugaut, A.; Murat, P.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19953.
(32) Wang, Z. F.; Li, M. H.; Hsu, S. T.; Chang, T. C. *Nucleic Acids Res.* **2014**, *42*, 4723.
(33) Phan, A. T.; Patel, D. J. *J. Am. Chem. Soc.* **2002**, *124*, 1160.
(34) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661.
(35) Plateau, P.; Gueron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310.
(36) Schanda, P.; Kupce, E.; Brutscher, B. *J. Biomol. NMR* **2005**, *33*, 199.
(37) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.
(38) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521.
(39) Phan, A. T. *J. Biomol. NMR* **2000**, *16*, 175.
(40) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061.
(41) Gray, D. M.; Wen, J. D.; Gray, C. W.; Repges, R.; Repges, C.; Raabe, G.; Fleischhauer, J. *Chirality* **2008**, *20*, 431.
(42) Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M. *Nucleic Acids Res.* **2009**, *37*, 1713.
(43) Patel, D. J.; Tonelli, A. E. *Biopolymers* **1974**, *13*, 1943.
(44) Feigon, J.; Koshlap, K. M.; Smith, F. W. *Methods Enzymol.* **1995**, *261*, 225.
(45) Adrian, M.; Heddi, B.; Phan, A. T. *Methods* **2012**, *57*, 11.
(46) Dai, J.; Carver, M.; Yang, D. *Biochimie* **2008**, *90*, 1172.
(47) Phan, A. T. *FEBS J.* **2010**, *277*, 1107.
(48) Clevers, H.; Nusse, R. *Cell* **2012**, *149*, 1192.
(49) Polakis, P. *EMBO J.* **2012**, *31*, 3375.
(50) Lim, K. W.; Phan, A. T. *Angew. Chem., Int. Ed.* **2013**, *52*, 8566.
(51) Marusic, M.; Sket, P.; Bauer, L.; Viglasky, V.; Plavec, J. *Nucleic Acids Res.* **2012**, *40*, 6946.
(52) Kendrick, S.; Kang, H.-J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S. M.; Hurley, L. H. . *J. Am. Chem. Soc.* **2014**, *136*, 4161.
(53) Kang, H.-J.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. *J. Am. Chem. Soc.* **2014**, *136*, 4172.
(54) Zazopoulos, E.; Lalli, E.; Stocco, D. M.; Sassone-Corsi, P. *Nature* **1997**, *390*, 311.
(55) Gleghorn, M. L.; Davydova, E. K.; Rothman-Denes, L. B.; Murakami, K. S. *Mol. Cell* **2008**, *32*, 707.

DOI: 10.1021/ja5089327
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX