9152 *Biochemistry* **2010,** *49,* 9152–9160
DOI: 10.1021/bi100946g

BIOCHEMISTRY
including biophysical chemistry & molecular biology
*Article*

# Thermodynamic Stability and Folding Kinetics of the Major G-Quadruplex and Its Loop Isomers Formed in the Nuclease Hypersensitive Element in the Human c-Myc Promoter: Effect of Loops and Flanking Segments on the Stability of Parallel-Stranded Intramolecular G-Quadruplexes[†]

Emmanuel Hatzakis,[‡] Keika Okamoto,[‡] and Danzhou Yang*[,‡,§]

[‡]*College of Pharmacy, The University of Arizona, 1703 East Mabel Street, Tucson, Arizona 85721, United States, and* [§]*BIO5 Institute, The University of Arizona, Tucson, Arizona 85721, United States, Arizona Cancer Center, 1515 North Campbell Avenue, Tucson, Arizona 85721, United States, and Department of Chemistry, The University of Arizona, Tucson, Arizona 85721, United States*

*Received June 11, 2010; Revised Manuscript Received September 15, 2010*

ABSTRACT: Overexpression of the c-Myc proto-oncogene is associated with a broad spectrum of human cancers. Nuclease hypersensitivity element $III_1$ (NHE $III_1$) of the c-Myc promoter can form transcriptionally active and silenced forms, and the formation of DNA G-quadruplex structures has been shown to be critical for c-Myc transcriptional silencing. The major G-quadruplex formed in c-Myc NHE $III_1$ is a mixture of four loop isomers, which have all been shown to be biologically relevant to c-Myc transcriptional control. In this study, we performed a thorough thermodynamic and kinetic study of the four c-Myc loop isomers in a $K^+$ solution. The four loop isomers all form parallel-stranded G-quadruplexes with short loop lengths. While the parallel-stranded G-quadruplex has been known to favor short loop lengths, our results show that the difference in thermodynamic and kinetic properties of the four loop isomers, and hence between the parallel G-quadruplexes with similar loop lengths, is more significant than previously recognized. At 20 mM $K^+$, the average difference in the $T_m$ values between the most stable loop isomer 14/23 and the least stable loop isomer 11/20 is more than 10 °C. In addition, the capping structures formed by the extended flanking segments are shown to contribute to a stabilization of 2−3 °C in $T_m$ for the c-Myc promoter G-quadruplex. Understanding the intrinsic thermodynamic stability and kinetic properties of the c-Myc G-quadruplex loop isomers can aid in our understanding of their biological roles and drug targeting.

Overexpression of the c-Myc proto-oncogene is associated with a broad spectrum of human cancers, including colon, breast, prostate, cervical, and lung carcinomas, osteosarcomas, lymphomas, and leukemias (*1*−*9*). In addition, elevated levels of c-Myc expression are often associated with poor therapeutic prognosis. c-Myc overexpression can be caused by different mechanisms, including gene amplification (*10*, *11*), translocation (*12*−*14*), and simple upregulation of transcription (*1*, *15*). The transcriptional regulation of c-Myc expression is complex and involves multiple promoters and transcriptional start sites, with P1 and P2 being the predominant promoters (*16*). A highly conserved NHE $III_1$, a 27 bp sequence located at base pairs 115−142 upstream from the P1 promoter, has been shown to be required for 80−95% of c-Myc transcription, regardless of whether the P1 or P2 promoter is used (*17*, *18*). This NHE $III_1$ element can form transcriptionally active and silenced forms in the promoter (*19*); the formation of DNA G-quadruplex structures is critical for c-Myc transcriptional silencing (*20*−*22*), and compounds that stabilize the G-quadruplex can repress c-Myc gene expression (*20*, *23*).

DNA G-quadruplexes make up a family of secondary DNA structures that consist of stacked G-tetrads connected by Hoogsteen hydrogen bonds and stabilized by monovalent cations such as potassium and sodium. Intramolecular G-quadruplexes have been found in a number of G-rich regions with biological significance, such as human telomeres and oncogene promoters (*24*, *25*). The G-rich strand of c-Myc NHE $III_1$ is a 27-nucleotide segment (mycPu27) comprised of five consecutive runs of guanines (Figure 1A). It has been previously shown by mutational analysis in conjunction with a luciferase reporter system that the major G-quadruplex structure formed in MycPu27 in a $K^+$ solution involves four consecutive 3′ runs of guanines (G7−G9, G11−G14, G16−G18, and G20−G23) but not the 5′ run of guanines (G2−G5) (Figure 1A) (*20*, *26*). However, the sequence containing four consecutive 3′ guanine runs [Myc2345 (Figure 1A)] can form a mixture of four loop isomers because only three consecutive guanines of the two four-guanine runs (G11−G14 and G20−G23) are involved in G-tetrad formation (*26*) (Figure 1B). The four loop isomers have different arrangements of one- and two-nucleotide loops and can be separately formed by dual G-to-T substitutions at positions 14 and 23, 11 and 23, 14 and 20, and 11 and 20, respectively. While the 14/23 isomer is the major conformation of Myc2345, all of the four loop isomers have been shown to be biologically relevant to c-Myc transcriptional control (*26*). We have determined the molecular structure of the major loop isomer (1:2:1) of the c-Myc G-quadruplex in a $K^+$ solution using the 14/23 sequence (Figure 1A) (*27*). The four loop isomers have been shown to adopt parallel-stranded folding topology in a $K^+$ solution (*26*, *28*).

[†]This research was supported by National Institutes of Health Grants 1S10 RR16659 and CA122952.

*To whom correspondence should be addressed. Telephone: (520) 626-5969. Fax: (520) 626-6988. E-mail: yang@pharmacy.arizona.edu.

 Published on Web 09/17/2010 

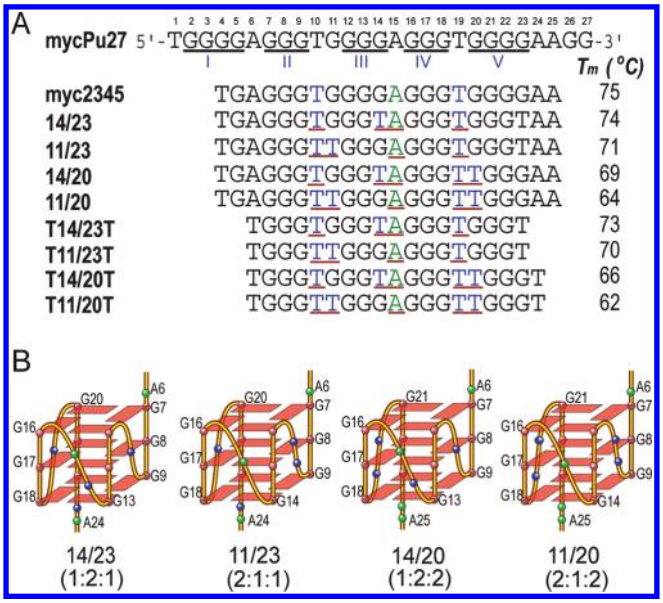


FIGURE 1: (A) Promoter sequence of the NHE $III_1$ element of the c-Myc gene and its modifications. MycPu27 is the wild-type 27-mer G-rich sequence of c-Myc NHE $III_1$. The five runs of guanines are underlined and numbered from the 5′ end. The numbering system is shown above mycPu27. Myc2345 is the 22-nucleotide G-rich sequence that adopts the major c-Myc promoter G-quadruplex. Myc2345 has four loop isomers, which can be isolated by dual G-to-T substitutions at positions 14 and 23, 11 and 23, 14 and 20, and 11 and 20, namely, 14/23, 11/23, 14/20, and 11/20, respectively. 14/23 is the mutant Myc2345 that forms the major loop isomer and was used for NMR structure determination (*26*). The truncated sequences of four loop isomers have only one flanking T at each end. The three loops of each loop isomer are underlined. The melting temperatures ($T_m$) of all the sequences in a 20 mM $K^+$ solution are also shown. (B) Schematic drawing of the folding topologies of the four loop isomers of the major c-Myc G-quadruplex Myc2345 in a $K^+$ solution. The four loop isomers all adopt parallel-stranded folding and have different arrangements of one- and two-nucleotide loops. Base colors: red for guanine, blue for thymine, and green for adenine.

The wild-type sequence, which forms a mixture of all four loop isomers, shows the strongest inhibition of c-Myc transcription, indicating that each loop isomer contributes to this inhibitory effect (*26*). Thus, the biological function of the c-Myc promoter G-quadruplex depends on the equilibrium of the four loop isomers; however, it is not known yet what the relative stability and folding kinetics of each loop isomer are and how each loop isomer contributes to the equilibrium.

In this work, we performed a complete thermodynamic and folding kinetic study of the four loop isomers of the major G-quadruplex formed in the c-Myc promoter in a $K^+$ solution. We also examined the effect of the flanking sequences by comparing sequences with and without flanking segments. The different loop isomers may have different protein interactions and could be recognized by different small molecule compounds. Understanding the intrinsic thermodynamic stability and folding kinetics of the four loop isomers may thus aid in our understanding of the biological role of the c-Myc loop isomers and their drug targeting. In addition, it has been found that the intramolecular G-quadruplex structures formed in gene promoters are in general a mixture of multiple loop isomers because of the unequal number of guanines in each G-tract, and that parallel-stranded structural motifs with short loops are prevalent in promoter quadruplexes (*29*). While the parallel-stranded structural motifs have been shown to favor short loop lengths in a $K^+$ solution (*24*, *27*, *30*−*32*), it is unclear how much difference exists between the one- or two-nucleotide loops. Thus, our results also provide information about the effects of the loop sizes and arrangements, as well as of the flanking segments, on the thermodynamic stability and folding kinetics of parallel-stranded G-quadruplexes with short loops in general.

## MATERIALS AND METHODS

*Oligonucleotides*. The DNA oligonucleotides were synthesized using β-cyanoethyl phosphoramidite solid-phase chemistry on an Expedite 8909 Nucleic Acid Synthesis System (Applied Biosystem, Inc., Foster City, CA) in DMT-on mode and were purified using MicroPure II columns from BioSearch Technologies (Novato, CA) according to the provided protocol. Concentrations were determined using UV spectroscopy by recording the absorbance at 260 nm. Oligonucleotide extinction coefficients ($\varepsilon_{260}$) were calculated using the nearest neighbor method (*33*) as follows: 266170 $M^{-1}$ $cm^{-1}$ for Myc2345, 258933 $M^{-1}$ $cm^{-1}$ for 14/23, 11/23, 14/20, and 11/20, 201329 $M^{-1}$ $cm^{-1}$ for T14/23T and T11/23T, and 209732 $M^{-1}$ $cm^{-1}$ for T14/20T and T11/20T.

*Circular Dichroism (CD) Spectroscopy*. CD spectroscopy of the oligonucleotides was performed on a Jasco J-810 spectropolarimeter (Jasco, Easton, MD) equipped with a thermoelectrically controlled cell holder. DNA samples were prepared at 10 μM in 10 mM $Li_3PO_4$ (pH 7.0) with various concentrations of KCl in a total volume of 200 μL. The DNA samples were annealed by being heated at 95 °C for 10 min and cooled to room temperature and then transferred into a quartz cell with an optical path length of 1 mm for CD measurements. A blank sample containing only buffer was used for the baseline correction.

CD spectroscopy measurements were the averages of three scans collected between 350 and 200 nm at 25 °C. The scanning speed of the CD instrument was 100 nm/min, and the response time was 1 s.

For CD melting or annealing experiments, CD measurements were performed at 265 nm, with a heating or cooling rate of 0.5−2 °C/min. A heating or cooling rate lower than 0.5 °C/min gave notable instrument-related noises.

*NMR Spectroscopy*. NMR experiments were performed on a Bruker DRX-600 spectrometer as described previously (*30*). Samples in water were prepared in a 10%/90% $D_2O/H_2O$ solution. The final NMR samples contained 0.1−2.5 mM DNA in 25 mM potassium phosphate buffer (pH 7.0) and 70 mM KCl. Before the experiments were conducted, samples were annealed by being heated to 95 °C for 10 min and then cooled to room temperature. The one-dimensional (1D) $^1H$ NMR spectra were recorded at 25 °C with a Watergate pulse sequence to suppress the water signal. The spectral width was 25 ppm, and 256 scans were collected. In each scan, 16K data points were used.

*Data Analysis*. (*i*) *Thermodynamic Analysis*. The melting temperatures ($T_m$) were calculated from van't Hoff analysis of the CD melting profiles and from the maximum in the first derivative of the melting curves using SigmaPlot 8.0. We also used SigmaPlot 8.0 to determine the $T_{1/2}$ values ($T_{1/2}$ is the temperature for which the normalized absorbance is 0.5). These two methodologies gave similar results for $T_m$ and $T_{1/2}$, which is expected for intramolecular quadruplexes. The folded fraction was determined from the difference between the absorbance of the upper and lower asymptotes as previously described (*34*). To test the reproducibility of the technique, we repeated all the experiments at least twice; the $T_m$ values of the repeats differed by $< 1$ °C.

In the case of a two-state folding−unfolding process with no significant intermediate formed, the process can be symbolized by

$$\mathrm{UF} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{F} \quad K_{eq} = k_1/k_{-1} \tag{1}$$

where F and UF refer to the folded and the unfolded fractions of the oligonucleotides, respectively, $K_{eq}$ is the equilibrium constant, and $k_1$ and $k_{-1}$ are the folding and unfolding rate constants, respectively. The equilibrium constant ($K_{eq}$ = [F]/[UF]) can be calculated from the equation

$$K_{eq} = a/(1-a) \tag{2}$$

where $a$ is the folded fraction obtained from the normalized experimental absorbance in melting and annealing curves. For complete folding, $a = 1$, and for complete unfolding, $a = 0$. To calculate the equilibrium constants precisely, the analysis must be restricted to the temperature range for which $0.03 < a < 0.97$ (*34*). From the equations

$$\Delta G = \Delta H - T\Delta S \tag{3}$$

$$\Delta G = -RT \ln K_{eq} \tag{4}$$

where $T$ is the temperature in kelvin and $R$ is the gas constant (8.314 J $K^{-1}$ $mol^{-1}$), the following equation can be obtained:

$$\ln K_{eq} = -\Delta H/RT + \Delta S/R \tag{5}$$

On the basis of this equation, in a two-state model, $\Delta H$ and $\Delta S$ values can be determined by plotting the line of $\ln K_{eq}(T)$ ($y$ axis) versus $1/T$ ($x$ axis), where $-\Delta H/R$ is the slope and $\Delta S/R$ is the $y$ intercept. $\Delta G$ values can be calculated from eq 3 (assuming that $\Delta C_p = 0$). $T_m$ values can also be determined from the same equation. At the melting temperature $T_m$, $\Delta G = 0$, and therefore, $T_m = \Delta H/\Delta S$.

(*ii*) *Number of Bound Potassium Cations*. As previously described (*35*), assuming a simple model of the transition between the folded and the unfolded state:

$$\text{folded} \Leftrightarrow \text{unfolded} + \Delta n$$

where $\Delta n$ is the difference between the number of monovalent ions bound in the folded and unfolded states and thus is the number of cations specifically bound to the folded state. The $\Delta n$ value can be calculated from the rate of change in $\Delta G_{25}$ as a function of log[$K^+$] based on the following equation:

$$\Delta n = -\mathrm{d}(\ln K_{eq})/\mathrm{d}(\ln[K^+]) = \Delta\Delta G/(2.3RT\Delta\log[K^+]) \tag{6}$$

where $\Delta\Delta G_{25}/\Delta\log[K^+]$ is the slope of the line that relates $\Delta G_{25}$ and log[$K^+$].

(*iii*) *Kinetic Analysis*. $k_1$ and $k_{-1}$ rate constants for folding and unfolding processes, respectively, can be derived using the hysteresis between the CD melting and annealing curves, as previously described (*34*, *36*, *37*).

The rate equation associated with a two-state model (eq 1) is

$$\mathrm{d}[F]/\mathrm{d}t = k_1[UF] - k_{-1}[F] \tag{7}$$

The use of folded fraction $a$ in eq 7 gives

$$\mathrm{d}(a_a)/\mathrm{d}t = k_1(1-a_a) - k_{-1}a_a \tag{8}$$

$$\mathrm{d}(a_m)/\mathrm{d}t = k_1(1-a_m) - k_{-1}a_m \tag{9}$$

where $a_m$ and $a_a$ are the fractions of the folded DNA at any temperatures in the melting and annealing curves, respectively. Additionally, $\mathrm{d}(a_a)/\mathrm{d}t$ and $\mathrm{d}(a_m)/\mathrm{d}t$ can be determined from the equation

$$\mathrm{d}a/\mathrm{d}t = \mathrm{d}a/\mathrm{d}T\ \mathrm{d}T/\mathrm{d}t \tag{10}$$

where $\mathrm{d}T/\mathrm{d}t$ is the temperature gradient used in melting and annealing experiments and the term $\mathrm{d}a/\mathrm{d}T$ can be determined from the experimental melting and annealing curves. By solving eqs 8 and 9, we can calculate $k_1$ and $k_{-1}$ values over a wide range of temperatures. Because $k_1 = k_{-1}$ at the melting temperature, $T_m$ values can be calculated for oligonucleotides with very slow kinetics, for which hysteresis occurs even at a very slow rate of heating and cooling. Activation energies can be determined using the Arrhenius equation

$$k = Ae^{-E_a/RT} \tag{11}$$

where $E_a$ is the activation energy and $A$ is the pre-exponential factor (or frequency factor). Equation 11 can be transformed to

$$\ln k = -E_a/RT + \ln A \tag{12}$$

In the Arrhenius relationship ($\ln k$ vs $1/T$), the slope of the line is $-E_a/R$ and the $y$ intercept is $\ln A$.

## RESULTS

*Four Loop Isomers of the Major c-Myc G-Quadruplex Myc2345 Formed Well-Defined Parallel-Stranded Conformations under Physiologically Relevant Conditions, As Shown by CD and NMR*. For the two four-guanine runs (G11−G14 and G20−G23) of the c-Myc promoter NHE $III_1$ element (Figure 1A), only three consecutive guanines are involved in formation of the G-tetrad; therefore, the sequence containing the four consecutive 3′ guanine runs (Myc2345) (Figure 1B) can form a mixture of four loop isomers. Strategic dual G-to-T substitutions at positions 14 and 23, 11 and 23, 14 and 20, and 11 and 20 lead to DNA sequences that can form only one of the four possible loop isomers, i.e., the 1:2:1, 2:1:1, 1:2:2, and 2:1:2 loop isomers, respectively (Figure 1B). The 1:2:1 loop isomer has been shown to be the major loop isomer of the c-Myc promoter G-quadruplex, with parallel-stranded folding (*26*). We have determined the NMR structure of this major loop isomer in a physiologically relevant $K^+$ solution using the 14/23 sequence, a 22-mer c-Myc promoter sequence with dual G-to-T substitutions at positions 14 and 23 (Figure 1A) (*27*). The 14/23 sequence exhibits a NMR spectrum with a sharp line width and good spectral resolution in 100 mM $K^+$ at pH 7 (Figure 2A, top). Twelve sharp imino peaks in the region of 10.5−12.5 ppm, which is a characteristic region for tetrad guanines in a G-quadruplex, indicate the formation of a single stable intramolecular G-quadruplex structure with three G-tetrads (Figure 1B, left) (*27*).

To systematically examine the stability of the four c-Myc G-quadruplex loop isomers, we modified the Myc2345 sequence with different dual G-to-T substitutions at positions 11 and 23, 14 and 20, and 11 and 20, namely, 11/23, 14/20, and 11/20 for the 2:1:1, 1:2:2, and 2:1:2 loop isomers, respectively (Figure 1A). The same 5′ flanking segment is used for all four loop isomers for a fair comparison. We applied NMR spectroscopy for each sequence to monitor the formation of the G-quadruplex. The imino proton regions of the 1D $^1$H NMR spectra of the four loop isomer sequences in a 100 mM $K^+$ solution at pH 7 are shown in

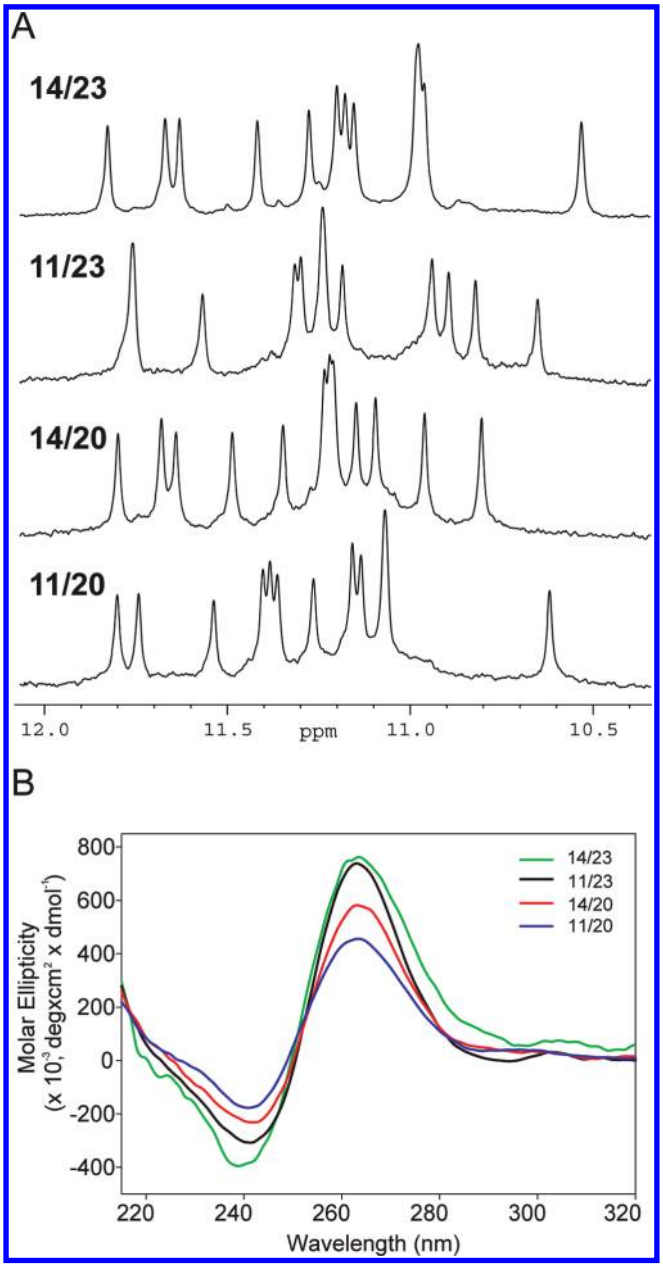


FIGURE 2: (A) Imino regions of 1D $^1$H NMR spectra of the 14/23, 11/23, 14/20, and 11/20 sequences that form four loop isomers of the major c-Myc G-quadruplex. All four modified 22-nucleotide Myc2345 sequences show high-quality $^1$H NMR spectra, indicating the formation of a single and stable G-quadruplex structure. Experimental conditions: 25 °C, 30 mM potassium phosphate, 70 mM KCl, pH 7.0, and 0.2 mM DNA. (B) CD spectra of 14/23, 11/23, 14/20, and 11/20 in the presence of 10 mM $Li_3PO_4$ (pH 7) and 2 mM $K^+$ at 25 °C. Each sample contained 10 $\mu$M DNA. All four loop isomers are shown to adopt a parallel-stranded topology with all *anti* glycosidic sugar conformations for tetrad guanines, as indicated by a positive peak around 265 nm and a negative peak around 240 nm.

Figure 2A. All of the four modified c-Myc promoter sequences give rise to excellent quality NMR spectra with sharp peaks. As anticipated, 12 well-resolved and sharp imino proton peaks are observed for each loop isomer sequence, indicating the formation of a single stable intramolecular G-quadruplex with three G-tetrads (Figure 1B).

The CD spectra of the four loop isomer sequences in a $K^+$ solution, which showed a positive peak at 265 nm and a negative peak at 240 nm (Figure 2B), confirmed that they all adopt a parallel-stranded topology as previously suggested (*26*). CD spectroscopy has become a useful method for inferring the topology of a DNA G-quadruplex (*31*, *38*−*41*). The stacking tetrad guanines with the same *anti* glycosidic bond conformation, as seen in G-quadruplexes with a parallel-stranded topology, give rise to a CD profile with a positive peak around 265 nm and a negative peak around 240 nm, whereas the stacking tetrad guanines with alternating *syn* and *anti* glycosidic conformations, as seen in G-quadruplexes with antiparallel-stranded topologies, give rise to a positive CD peak at 295 nm.

The CD profiles of the four c-Myc G-quadruplex loop isomers were not affected by the concentrations of $K^+$ in the range of 2−100 mM. At very low $K^+$ concentrations, such as 0.5 mM, a decrease in molecular ellipticity was observed, likely caused by partial unfolding of G-quadruplexes because of the low melting temperatures.

*CD Melting Studies of the Four c-Myc Loop Isomers with Flanking Segments Show a Clear Difference in Their Stability, with 14/23 Being the Most Stable*. To examine the stability of the four c-Myc G-quadruplex loop isomers, CD melting experiments were performed at 265 nm, which was the maximum in CD profiles for all four parallel-stranded loop isomers. Lithium phosphate at pH 7 was used as the buffer for all DNA samples in the CD melting study, because the lithium cation has been shown not to stabilize the G-quadruplex structures (*42*) and thus does not interfere with accurate measurements of the stabilizing effects of the potassium cation. No aggregation was observed for any DNA samples. The melting temperatures ($T_m$) of a G-quadruplex, at which half of the DNA molecules are unfolded, were determined using the CD melting curve. Additionally, thermodynamic parameters, such as $\Delta H$, $\Delta S$, and $\Delta G$, were calculated from melting profiles using van't Hoff analysis. At concentrations of $<0.5$ mM, the c-Myc G-quadruplex loop isomers were too unstable to have their melting temperatures accurately measured. At $K^+$ concentrations of $>50$ mM, the $T_m$ values of some loop isomers, such as 14/23, are too high to be measured by CD melting ($T_m > 90$ °C). Therefore, $K^+$ concentrations between 0.5 and 50 mM were used for the CD melting studies and thermodynamic calculations.

No hysteresis was observed in CD melting experiments for any of the four loop isomers at $K^+$ concentrations of $\geq 5$ mM with a heating and cooling rate of 2 °C/min, indicating that the folding and unfolding processes are at a thermodynamic equilibrium, and hence, the $T_m$ values and thermodynamic parameters can be accurately determined. At $K^+$ concentrations of $<5$ mM, a heating and cooling rate of 0.5 °C/min was needed to prevent any hysteresis. For the 11/20 sequence, even at the rate of 0.5 °C/min, which was the low limit for the CD instrument, some hysteresis was observed (Figure S1 of the Supporting Information), indicating that the temperature change is still too fast for the 11/20 G-quadruplex loop isomer to reach the thermodynamic equilibrium.

As shown in Table 1, the melting temperatures of the four c-Myc G-quadruplex loop isomers were all independent of concentration, confirming the intramolecular (monomeric) nature of the molecular systems. CD melting experiments performed at various concentrations of $K^+$ ion showed a strong dependence of the $T_m$ values on $K^+$ concentration (Table 2). As expected, the increase in the $K^+$ concentration markedly increases the $T_m$ values of the four c-Myc loop isomers. The CD melting curves of 14/23 at various $K^+$ concentrations are shown as an example in Figure 3A.

A clear difference in stability was observed among the four c-Myc loop isomers. CD melting profiles of the four loop isomers with flanking segments in 20 mM $K^+$ are shown in Figure 3B.

The 14/23 loop isomer (1:2:1) is the most stable, followed by 11/23 (2:1:1), 14/20 (1:2:2), and 11/20 (2:1:2), as indicated by the $T_m$ values (Figure 3B and Table 2). The 14/23 and 11/23 loop isomers, which have one two-nucleotide loop and two one-nucleotide loops with a total loop length of four, have a higher melting temperature than the 14/20 and 11/20 loop isomers, which have two two-nucleotide loops and one one-nucleotide loop with a total loop length of five. The difference between the $T_m$ values of the most stable loop isomer, 14/23, and the least stable loop isomer, 11/20, is greater than 10 °C. In addition, within the two loop isomers with the same total loop length, 14/23 (T:TA:T loop isomer) is clearly more stable than 11/23 (TT:A:T) and 14/20 (T:TA:TT) is clearly more stable than 11/20 (TT:A:TT) (Figure 1A), indicating that in addition to the loop length, the loop sequence also contributes to the stability of a G-quadruplex.

Thermodynamic parameters for the quadruplex−single strand transition for the four c-Myc promoter loop isomers with flanking segments were derived from van't Hoff analysis of the melting curves at various $K^+$ concentrations (Table 2 and Table S1 of the Supporting Information). For a two-state model in which there is no significant population of intermediates, the plot

Table 1: $T_m$ (degrees Celsius) Values for 14/23, 11/23, 14/20, and 11/20 Using Different Concentrations of DNA[a]

| [DNA] ($\mu$M) | 14/23 | 11/23 | 14/20 | 11/20 |
|---|---|---|---|---|
| 5 | 59.0 | 54.1 | 51.5 | 48.0 |
| 10 | 58.8 | 53.3 | 51.0 | 47.5 |
| 20 | 59.0 | 54.1 | 51.3 | 48.0 |
| 50 | 58.0 | 53.5 | 50.8 | 47.2 |

[a]All experiments were performed in the presence of 10 mM $Li_3PO_4$ (pH 7.0) containing 2 mM $K^+$, at a heating rate of 2 °C/min.

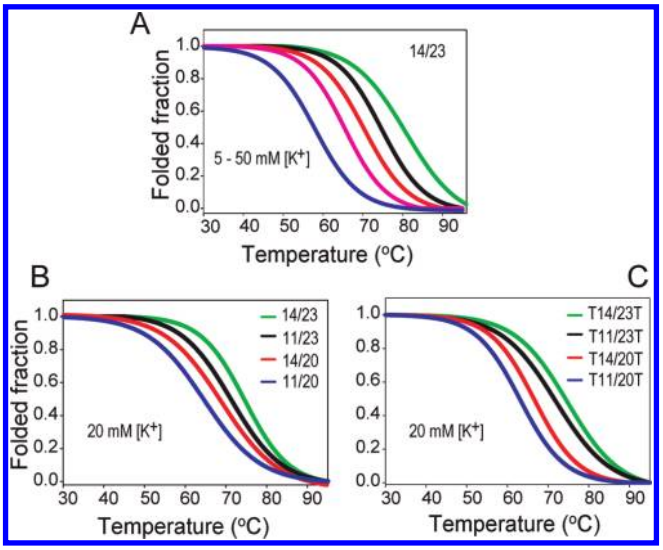


FIGURE 3: (A) CD melting curves for 14/23 in the presence of 10 mM $Li_3PO_4$ (pH 7) and different concentrations of $K^+$ (2−50 mM, from left to right). 14/23 shows $K^+$ concentration-dependent melting temperatures. The $T_m$ of 14/23 at 50 mM $K^+$ is ~80 °C: 2 (blue), 5 (pink), 10 (red), 20 (black), and 50 mM $K^+$ (green). (B and C). CD melting curves of 11/20 (blue), 14/20 (red), 11/23 (black), and 14/23 (green) (B) and of the truncated sequences T11/20T (blue), T14/20T (red), T11/23T (black), and T14/23T (green) (C) of four loop isomers in the presence of 10 mM $Li_3PO_4$ (pH 7) and 20 mM $K^+$. The 14/23 loop isomer is shown to be the most stable, followed by 11/23, 14/20, and 11/20. The heating rate for CD melting is 2 °C/min.

Table 2: Thermodynamic Data for the c-Myc G-Quadruplex Loop Isomers Formed by the Extended and Truncated DNA Sequences, Determined in the Presence of 10 mM $Li_3PO_4$ (pH 7) Containing Different Concentrations of $K^+$[a]

| [$K^+$] (mM) | DNA | $T_m$ (°C) | $CT_m$[b] (°C) | $\Delta H$ (kJ/mol) | $\Delta S$ (kJ mol$^{-1}$ K$^{-1}$) | $\Delta G_{37}$ (kJ/mol) | DNA | $T_m$ (°C) | $CT_m$[b] (°C) | $\Delta H$ (kJ/mol) | $\Delta S$ (kJ mol$^{-1}$ K$^{-1}$) | $\Delta G_{37}$ (kJ/mol) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.5[c] | 14/23 | 43 | 42 | −129 | −0.412 | −1.86 | | | | | | |
| 2[c] | 14/23 | 57 | 56 | −173 | −0.526 | −9.92 | T14/23T | 57 | 57 | −167 | −0.508 | −9.96 |
| 5 | 14/23 | 64 | 63 | −181 | −0.540 | −14.03 | T14/23T | 65 | 65 | −165 | −0.491 | −13.40 |
| 10 | 14/23 | 69 | 69 | −185 | −0.542 | −17.57 | T14/23T | 69 | 69 | −162 | −0.475 | −14.92 |
| 20 | 14/23 | 75 | 74 | −212 | −0.612 | −22.59 | T14/23T | 73 | 72 | −184 | −0.535 | −18.32 |
| 50 | 14/23 | 80 | 79 | −192 | −0.548 | −22.93 | | | | | | |
| 0.5[c] | 11/23 | 39 | 37 | −120 | −0.390 | 0.05 | | | | | | |
| 2[c] | 11/23 | 52 | 52 | −154 | −0.476 | −6.90 | T11/23T | 53 | 53 | −156 | −0.481 | −7.18 |
| 5 | 11/23 | 60 | 60 | −146 | −0.439 | −10.03 | T11/23T | 61 | 60 | −166 | −0.501 | −11.60 |
| 10 | 11/23 | 64 | 63 | −173 | −0.516 | −13.65 | T11/23T | 65 | 64 | −159 | −0.473 | −12.37 |
| 20 | 11/23 | 72 | 71 | −189 | −0.550 | −18.50 | T11/23T | 70 | 69 | −189 | −0.553 | −17.45 |
| 50 | 11/23 | 75 | 75 | −158 | −0.454 | −17.51 | | | | | | |
| 0.5[c] | 14/20 | 34 | 34 | −129 | −0.421 | 1.41 | | | | | | |
| 2[c] | 14/20 | 50 | 49 | −154 | −0.479 | −5.70 | T14/20T | 48 | 48 | −165 | −0.515 | −5.60 |
| 5 | 14/20 | 58 | 57 | −159 | −0.483 | −9.70 | T14/20T | 57 | 56 | −172 | −0.523 | −9.95 |
| 10 | 14/20 | 62 | 61 | −159 | −0.480 | −11.10 | T14/20T | 61 | 61 | −180 | −0.540 | −12.78 |
| 20 | 14/20 | 69 | 69 | −178 | −0.521 | −16.49 | T14/20T | 66 | 65 | −187 | −0.553 | −15.45 |
| 50 | 14/20 | 72 | 71 | −158 | −0.462 | −15.39 | | | | | | |
| 0.5[c] | 11/20 | 28 | 26 | −108 | −0.363 | 4.03 | | | | | | |
| 2[c] | 11/20 | 47 | 46 | −115 | −0.362 | −3.33 | T11/20T | 44 | 44 | −150 | −0.476 | −3.00 |
| 5 | 11/20 | 53 | 52 | −118 | −0.364 | −5.49 | T11/20T | 53 | 53 | −164 | −0.505 | −7.76 |
| 10 | 11/20 | 57 | 56 | −126 | −0.386 | −7.16 | T11/20T | 58 | 58 | −165 | −0.501 | −10.15 |
| 20 | 11/20 | 65 | 64 | −148 | −0.442 | −11.61 | T11/20T | 62 | 62 | −173 | −0.518 | −12.81 |
| 50 | 11/20 | 67 | 66 | −138 | −0.408 | −11.96 | | | | | | |

[a]Samples were heated and cooled at a rate of 2 °C/min. The samples were prepared at a DNA concentration of 10 $\mu$M. All experiments were performed at least twice, and $T_m$ values are accurate to within 1 °C. $\Delta H$, $\Delta S$, and $\Delta G$ values are given with an experimental error of ±5% determined on the basis of at least two independent experiments. Missing values at low concentrations of $K^+$ correspond to quadruplexes for which the $T_m$ values were too low to measure or for which too much hysteresis was observed between the melting and annealing profiles even by using the lowest rate of cooling and heating. [b]Melting temperatures ($CT_m$) calculated from the equation $T_m = \Delta H/\Delta S$. [c]Samples were heated and cooled at a rate of 0.5 °C/min.

of ln $K$ versus $1/T$ provides a linear relationship with a slope of $-\Delta H/R$ and a $y$ intercept of $\Delta S/R$. Note that thermodynamic calculations at $K^+$ concentrations of $<2$ mM were less accurate, because of the presence of hysteresis between the melting and annealing curves as well as the fact that the quadruplexes were less stable under these conditions. Consistent with the $T_m$ data, the thermodynamic data indicated that 14/23 is the most stable loop isomer whereas 11/20 is the least stable. The $\Delta\Delta G_{25}$ between the two loop isomers is approximately $-13$ kJ/mol at 20 mM $K^+$. Again, the sequences with a total loop length of four, i.e., 14/23 and 11/23, were more stable than the sequences with a total loop length of five5, i.e., 14/20 and 11/20, whereas 14/23 appeared to be more stable than 11/23 and 14/20 more stable than 11/20. In addition, the $\Delta H$ values of all four c-Myc loop isomers clearly decreased as the concentration of $K^+$ increased from 0.5 to 20 mM, indicating the markedly enhanced stability of the c-Myc G-quadruplexes when the $K^+$ concentration increased to 20 mM. This is consistent with the presence of specific cation binding sites within the G-quadruplex as previously reported (*35*). However, at 50 mM $K^+$, although the $T_m$ values still increased from those at 20 mM $K^+$, the decrease of the $\Delta G$ values became much less obvious (14/23 and 11/20) or even absent (11/23 and 14/20) (Table 2). This phenomenon was also manifested by the shapes of the CD melting curves, which were less sharp at 50 mM $K^+$ (Figure 3A), suggesting the presence of minor conformations (see Discussion).

*CD Melting Studies of the Four Truncated c-Myc Loop Isomers Indicated That the Contribution of the Longer Flanking Segments Was ~2−3 °C in $T_m$.* The flanking segments have been found to be important for the formation of a well-defined c-Myc promoter G-quadruplex in NMR spectroscopic studies (*27*). As shown in our previous NMR structure study, specific capping structures are formed by the flanking segments at both ends of the 14/23 loop isomer (*27*), and the stable formation of capping structures has been suggested to contribute to the stability of this parallel-stranded major c-Myc G-quadruplex. To examine how the c-Myc flanking segments and the capping structures contribute to the stability of the c-Myc G-quadruplex loop isomers, we conducted CD melting studies on truncated c-Myc promoter sequences. While the same sample preparation and annealing processes were used for all DNA samples, completely truncated sequences without any flanking segments (i.e., start and end with a guanine in a G-run) often gave unrepeatable results in CD melting studies, implying the strong likelihood of the formation of higher-order structures. The formation of higher-order structures in completely truncated sequences was indicated by $^1$H NMR spectra, which showed a broader envelope of imino resonances in a much narrower range of 10.5−11 ppm (Figure S2 of the Supporting Information). Aggregation was also observed as a gel-like solution after multiple melting−annealing processes in completely truncated DNA sequences. Therefore, we added a flanking thymine at the 5′ and 3′ ends of the terminal G-runs for the four c-Myc loop isomers, namely, T14/23T, T11/23T, T14/20T, and T11/20T (Figure 1A). CD melting studies with the four loop isomers with terminal thymines appeared to be repeatable, and no aggregation was observed for multiple melting−annealing experiments with those DNA samples. Figure 3C shows CD melting curves of the four truncated sequences at 20 mM $K^+$. The results showed that the truncated sequences for the four loop isomers have a lower $T_m$ than the sequences with the longer flanking segments (Figure 1A), with the differences in $T_m$ being 2, 2, 3, and 3 °C for 14/23, 11/23, 14/20, and 11/20, respectively, at 20 mM $K^+$. Thus, it appeared that the longer flanking segments contributed to a stabilization of 2−3 °C in $T_m$ for the major c-Myc promoter G-quadruplexes. The thermodynamic parameters of the four truncated sequences at various $K^+$ concentrations are summarized in Table 2. The relative stabilities for the four loop isomers are the same as those observed in sequences with longer flanking segments as described above. In particular, T14/23T is the most stable, followed by T11/23T, T14/20T, and T11/20T, as measured by both the $T_m$ and $\Delta G_{37}$ values.

The number of monovalent potassium cations ($\Delta n$) specifically bound to a G-quadruplex could be calculated by the slopes of the plots of $\Delta G_{25}$ versus log[$K^+$] (see eq 6) (*35*, *43*). $\Delta G$ values at 25 °C were used to ensure the complete folding of all four c-Myc loop isomers at the $K^+$ concentrations used for the calculation. $\Delta n$ values were calculated to be 2.63, 2.24, 1.90, and 2.00 for 14/23, 11/23, 14/20, and 11/20, respectively, and 2.00, 1.91, 1.90, and 1.85 for the truncated sequences T14/23T, T11/23T, T14/20T, and T11/20T, respectively. A $\Delta n$ value of 2 appears likely for the four truncated loop isomers, as two potassium cations are located between the three stacked G-tetrads of the c-Myc quadruplexes that contain only short loops. The higher $\Delta n$ values of the sequences with longer flanking sequences are likely caused by the binding of the potassium cation within the capping structure formed by the flanking segments (*27*).

*Kinetic Calculations Indicated That the Four Loop Isomers Differ in Their Folding and Unfolding Kinetics.* No hysteresis was observed in CD melting experiments for all of the four loop c-Myc isomers with flanking segments at $K^+$ concentrations of $\geq 5$ mM with a heating and cooling rate of 2 °C/min. A clear hysteresis was observed for the four loop isomers in 2 mM $K^+$ at a heating and annealing rate of 2 °C/min. Figure 4 shows the melting and annealing curves for 14/23 at heating and annealing rates of 0.5 and 2 °C/min and for 14/20 at heating and annealing rates of 2 and 4 °C/min, respectively. The melting and annealing curves for 11/23 and 11/20 at heating and annealing rates of 2 and 4 °C/min are shown in Figure S3 of the Supporting Information. A clear difference in the levels of hysteresis of the four c-Myc loop isomers was observed. 14/23 and 11/23 showed the least hysteresis, 14/20 a larger hysteresis, and 11/20 a markedly larger hysteresis (Figure 4 and Figure S3 of the Supporting Information). At 2 mM $K^+$, the $\Delta T_m$ values between the melting and annealing curves are approximately 3.5, 4, 4.7, and 6 °C for 14/23, 11/23, 14/20, and 11/20, respectively. Note that the $T_m$ from the melting curve is always higher than that from the annealing curve. The presence of hysteresis between melting and annealing curves of a G-quadruplex indicates that the molecules are not at thermodynamic equilibrium because of the slow folding and/or unfolding processes.

The hysteresis data can be used to calculate the kinetic parameters, the rate constants of folding ($k_1$) and unfolding ($k_{-1}$), for unimolecular folding, as described previously (*34*, *37*). The kinetic parameters of the four c-Myc loop isomers with flanking segments were obtained from the melting and annealing curves of the four loop isomers in 2 mM $K^+$ at a rate of 2 °C/min. Arrhenius plots for folding and unfolding rates obtained from these data for the four loop isomers are shown in Figure 5. On the basis of the slopes of the plots, the activation energies of folding and unfolding processes can be determined. The calculated melting temperature ($CT_m$) for each loop isomer can also be determined using the Arrhenius plots, i.e., the temperature at which $k_1$ equals $k_{-1}$. The kinetic parameters derived form the

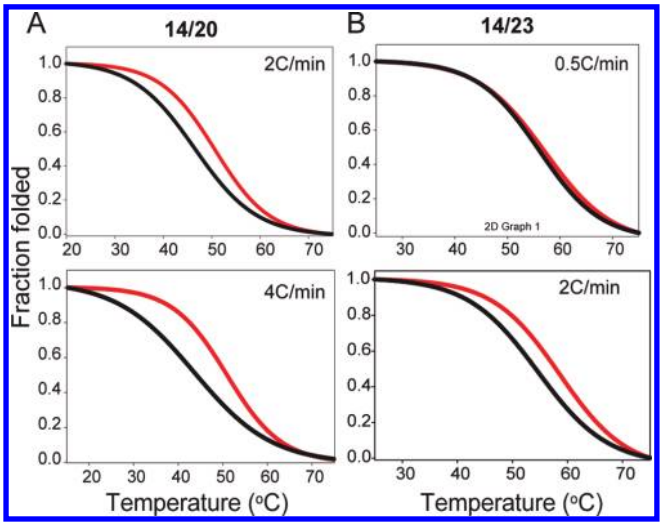


FIGURE 4: CD melting (red) and annealing (black) curves of 14/20 and 14/23 loop isomers with different temperature gradients. Hysteresis was observed between the melting and annealing curves for (A) 14/20 with temperature gradients of 2 °C/min (top) and 4 °C/min (bottom) and (B) 14/23 with temperature gradients of 2 °C/min (top) and 0.5 °C/min (bottom). No hysteresis was observed in 14/23 with a temperature gradient of 0.5 °C/min, while the greatest hysteresis was seen in 14/20 with a temperature gradient of 4 °C/min. The presence of hysteresis between melting and annealing profiles indicates the molecule is not at thermodynamic equilibrium during melting and annealing processes. Conditions: 10 mM $Li_3PO_4$, pH 7, and 2 mM $K^+$.

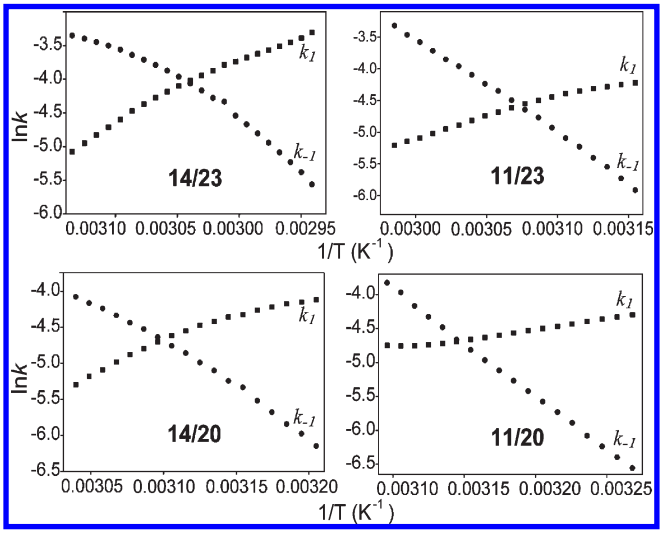


FIGURE 5: Arrhenius plots for 14/23, 11/23, 14/20, and 11/20 showing the temperature dependence of the kinetic parameters of the folding and unfolding processes, based on the analysis of the hysteresis between the melting and annealing profiles. The $k_1$ and $k_{-1}$ values were obtained from the annealing and melting profiles. Using the slopes of these curves, the activation energies of folding and unfolding reactions can be determined. The theoretical melting temperatures can be determined at the intersection of $k_1$ and $k_{-1}$ (curves), $k_1$ (■), and $k_{-1}$ (●). All experiments were performed in the presence of 10 mM $Li_3PO_4$ (pH 7.0) containing 2 mM $K^+$, at a cooling and heating rate of 2 °C/min.

Arrhenius plots are listed in Table 3. The association reaction, or the folding process, appears to be dominant at lower temperatures, as previously reported (*37*). The association reaction shows an apparent negative activation energy, indicating the reaction is faster at lower temperatures, as noted by others who suggested a nucleation-zipper reaction mechanism (*34*, *37*). Among the four c-Myc loop isomers, 14/23 had the most negative activation energy for the folding process, indicating that the 14/23 loop isomer has the fastest folding. 11/20 had the least negative activation energy for the folding process, indicating that the folding process of 11/20 is the slowest.

## DISCUSSION

The major G-quadruplex formed in the c-Myc promoter, the Myc2345 G-quadruplex, has four loop isomers, namely, 1:2:1, 2:1:1, 1:2:2, and 2:1:2, which can be isolated by 14/23, 11/23, 14/20, and 11/20 sequences, respectively (Figure 1A). The four c-Myc loop isomers have been previously shown to form parallel-stranded G-quadruplexes (*26*). While the 14/23 loop isomer is the major conformation of the c-Myc promoter G-quadruplex, all four loop isomers have been shown to be biologically relevant to c-Myc transcriptional control (*26*). The contribution of each loop isomer to c-Myc transcriptional control is likely to be related to the stability and dynamic equilibrium of the four loop isomers. In this study, we performed a thorough thermodynamic and kinetic study of the four loop isomers of the major Myc2345 G-quadruplex formed in the c-Myc promoter in a $K^+$ solution (*44*).

The four loop isomers of the major c-Myc promoter G-quadruplex all have short loop lengths, i.e., one nucleotide or two nucleotides. Specifically, the loop length arrangements of the four loop isomers, 14/23, 11/23, 14/20, and 11/20, are 1:2:1, 2:1:1, 1:2:2, and 2:1:2, respectively (Figure 1). Among the four loop isomers, 14/23 and 11/23 have a total loop length of four and 14/20 and 11/20 have a total loop length of five. The NMR and CD data showed that each of the four loop isomers forms a stable parallel-stranded G-quadruplex under physiological $K^+$ conditions (Figure 2 and Table 2). $T_m$ and $\Delta G$ values calculated from melting curves and van't Hoff analysis revealed that the thermodynamic stability of c-Myc G-quadruplexes depends strongly on loop lengths, while the loop sequences and the flanking segments also have a considerable effect. The 14/23 loop isomer is the most stable, followed by 11/23, 14/20, and 11/20 (Figure 3 and Table 2). Loop isomers with only one two-nucleotide loop such as 14/23 and 11/23 form more stable G-quadruplexes (higher $T_m$ and more negative $\Delta G$), whereas loop isomers with two two-nucleotide loops such as 14/20 and 11/20 form less stable G-quadruplexes. This is consistent with the previous reports (*31*, *32*) that the stability of a parallel-stranded G-quadruplex is correlated with its total loop length. It is generally considered that the parallel-stranded G-quadruplexes in a $K^+$ solution favor short loop lengths, i.e., one nucleotide or two nucleotides (*24*, *27*, *30*−*32*). Interestingly, our results show that the difference in thermodynamic and kinetic properties of the four loop isomers, which are the parallel G-quadruplexes with similar short loop lengths, is more significant than previously recognized. At 20 mM $K^+$, the average difference in the $T_m$ values between the most stable loop isomer, 14/23, and the least stable loop isomer, 11/20, is greater than 10 °C, while the $\Delta\Delta G_{37}$ between the two loop isomers is approximately −11 kJ/mol. It is thus suggested that a notable difference exists between one- and two-nucleotide loops, which may explain the differential contribution of the four loop isomers to the c-Myc transcriptional control. It is noteworthy that the $T_m$ of wild-type Myc2345, which is a mixture of four loop isomers, is only slightly higher than the $T_m$ of 14/23, the most stable loop isomer (Figure 1A).

Within the two loop isomers with the same total loop length, 14/23 (T:TA:T) is clearly more stable than 11/23 (TT:A:T) and 14/20 (T:TA:TT) is clearly more stable than 11/20 (TT:A:TT).

Table 3: Kinetic Parameters for the Folding and Unfolding Processes for 14/23, 11/23, 14/20, and 11/20, Determined from the Analysis of the Hysteresis between the Melting and Annealing Profiles[a]

| DNA | $k_1$ | | $k_{-1}$ | | $CT_m$[b] (°C) | $T_m$ (°C) (melting/annealing) |
|---|---|---|---|---|---|---|
| | $E_a$ (kJ/mol) | ln $A$ ($s^{-1}$) | $E_a$ (kJ/mol) | ln $A$ ($s^{-1}$) | | |
| 14/23 | −48 | −29 | 118 | 35 | 56 | 59.0/55.5 |
| 11/23 | −34 | −27 | 120 | 41 | 52 | 53.0/49.0 |
| 14/20 | −28 | −23 | 119 | 42 | 49 | 51.0/46.3 |
| 11/20 | −21 | −12 | 112 | 35 | 45 | 47.0/41.0 |

[a]All experiments were performed in the presence of 10 mM $Li_3PO_4$ (pH 7.0) containing 2 mM $K^+$ at a cooling and heating rate of 2 °C/min. [b]Calculated melting temperatures ($CT_m$) are the temperatures at which $k_1$ equals $k_{-1}$.

It has been previously shown that the loop position does not affect the parallel-stranded G-quadruplex stability (*31*). Thus, it is suggested that the stability of the four parallel-stranded c-Myc G-quadruplex loop isomers is also loop sequence-dependent, which has been suggested previously (*45*). While the G-to-T substitutions were used in this study, it would also be interesting to study the effect of other substitutions, such as G-to-A substitutions.

For the four loop isomers with extended flanking segments, the $T_m$ values increased while $\Delta G_{37}$ became more negative as the $K^+$ concentration increased to 20 mM (Table 2), indicating a clearly enhanced stability. The observation that the $\Delta G_{37}$ values are dependent on the cation concentration has been reported previously (*31*, *35*) and is consistent with the presence of specific ion binding sites in the G-quadruplex. However, when the $K^+$ concentration further increased to 50 mM, although the $T_m$ values still increased, the $\Delta G_{37}$ values did not change much as compared to those at 20 mM $K^+$. This phenomenon was also shown by the shape of the CD melting curves, which was less sharp at 50 mM $K^+$ (Figure 3A). This is likely to be caused by the copresence of minor conformations, such as different loop conformations of the flanking segments, which can be stabilized by a higher $K^+$ concentration. Generally, the coexistence of two conformations gives biphasic melting curves, but if the stability and melting temperatures of the two conformations are close enough, a superimposed melting profile with less sharpness is observed, which gives $\Delta H$, and therefore $\Delta G$, values that are lower than expected (*44*). Moreover, as discussed previously (*34*, *44*), if two G-quadruplexes are completely folded at 37 °C, then the one with the sharper melting profile may have more negative thermodynamic values at 37 °C even if its melting temperature is lower, as the shape of the melting curves affects the values of $\Delta H$ and therefore $\Delta G$.

We have also examined the effect of the flanking sequences by comparing sequences with extended flanking segments and truncated sequences with only a single thymine at both ends. The difference in $T_m$ values between the truncated and full flanking sequences is ~about 2−3 °C. Completely truncated sequences without any flanking residues (i.e., start and end with a guanine in a G-run) were prone to the formation of higher-order structures and thus gave unrepeatable results in CD melting studies. Thus, it needs to be noted that the true contribution of the capping structures may be greater than what was observed here, as a partial capping effect may still be imparted by the single flanking T of the truncated sequences. Using the plots of $\Delta G_{25}$ versus log[$K^+$], the difference in potassium ions bound to folded and unfolded conformations ($\Delta n$) can be calculated as previously described (*35*). $\Delta n$ values for the four truncated sequences are all ~2, consistent with the three-tetrad c-Myc G-quadruplex structures that contain two coordinated potassium cations. Higher $\Delta n$ values were shown for the 14/23 and 11/23 sequences with longer flanking segments, but not for the 14/20 and 11/20 sequences (Figure 1A), suggesting an additional potassium ion may be partially retained by the capping conformation formed with the 3′-flanking TAA of 14/23 and 11/23, which has been well-defined in the NMR structure of 14/23 (*27*). However, the $\Delta n$ values need to be used with caution, as factors such as the presence of hysteresis or the coexistence of multiple conformations may affect the accurate determination of $\Delta n$ values.

Kinetic analysis showed that the four c-Myc loop isomers had clear differences in their kinetic parameters for folding and unfolding processes. The folding process appears to be dominant under the physiologically relevant conditions and temperature. Sequences with one two-nucleotide loop (14/23 and 11/23) appear to have a markedly smaller amount of hysteresis than sequences with two two-nucleotide loops (14/20 and 11/20) (Figure 4 and Table 3). While 11/20 is the slowest in folding, it is also the slowest in the unfolding process. Between the loop isomers with the same total loop length, 14/23 appears to fold faster than 11/23 based on a more negative activation energy of 14/23 and 14/20 appears to fold faster than 11/20. This result suggests that, for the parallel-stranded loop isomers with two one-nucleotide loops and the same total loop length, the folding kinetics is faster for the loop isomer with both one-nucleotide loops at the ends versus one in the middle and one at the end. Thus, the folding of a parallel-stranded G-quadruplex may start at the two ends, which is favored by the parallel motifs with one-nucleotide loops at both ends.

In conclusion, we investigated the thermodynamic and kinetic properties of the G-quadruplex loop isomers that coexist in the nuclease hypersensitive element of the c-Myc promoter. This study may aid in our understanding of the dynamic equilibrium between the four loop isomers and the effect of ligand interactions, such as with a small molecule compound or a protein, on the stability and folding kinetics. Our results showed that the differences between G-quadruplexes with similar loop lengths were more significant than previously realized and also suggested that the loop constitution affected the stability of a G-quadruplex. Results also showed that the flanking segments had a clear effect on the stabilization of a G-quadruplex structure.

## ACKNOWLEDGMENT

We thank Dr. Megan Carver for proofreading the paper.

## SUPPORTING INFORMATION AVAILABLE

$\Delta G_{25}$ values for the c-Myc G-quadruplex loop isomers, hysteresis between CD melting and annealing curves of 11/20

with a temperature gradient of 0.5 °C/min, imino regions of 1D $^1$H NMR spectra of the completely truncated sequences for the four loop isomers, and hysteresis between CD melting and annealing curves of 11/20 and 11/23 loop isomers with temperature gradients of 2 and 4 °C/min. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Spencer, C. A., and Groudine, M. (1991) Control of C-Myc Regulation in Normal and Neoplastic-Cells. *Adv. Cancer Res. 56*, 1–48.
2. Magrath, I. (1990) The pathogenesis of Burkitt's lymphoma. *Adv. Cancer Res. 55*, 133–270.
3. Kinzler, K. W., and Vogelstein, B. (1996) Lessons from hereditary colorectal cancer. *Cell 87*, 159–170.
4. Berns, E., Klijn, J. G. M., Vanputten, W. L. J., Vanstaveren, I. L., Portengen, H., and Foekens, J. A. (1992) C-Myc Amplification Is a Better Prognostic Factor Than Her2/Neu Amplification in Primary Breast-Cancer. *Cancer Res. 52*, 1107–1113.
5. Pertschuk, L. P., Feldman, J. G., Kim, D. S., Nayeri, K., Eisenberg, K. B., Carter, A. C., Thelmo, W. T., Rhong, Z. T., Benn, P., and Grossman, A. (1993) Steroid-Hormone Receptor Immunohistochemistry and Amplification of C-Myc Protooncogene: Relationship to Disease-Free Survival in Breast-Cancer. *Cancer 71*, 162–171.
6. Nupponen, N. N., Kakkola, L., Koivisto, P., and Visakorpi, T. (1998) Genetic alterations in hormone-refractory recurrent prostate carcinomas. *Am. J. Pathol. 153*, 141–148.
7. Dang, C. V. (1999) c-Myc Target Genes Involved in Cell Growth, Apoptosis, and Metabolism. *Mol. Cell. Biol. 19*, 1–11.
8. Nesbit, C. E., Tersak, J. M., and Prochownik, E. V. (1999) MYC oncogenes and human neoplastic disease. *Oncogene 18*, 3004–3016.
9. Schlagbauer-Wadl, H., Griffioen, M., van Elsas, A., Schrier, P. I., Pustelnik, T., Eichler, H. G., Wolff, K., Pehamberger, H., and Jansen, B. (1999) Influence of increased c-Myc expression on the growth characteristics of human melanoma. *J. Invest. Dermatol. 112*, 332–336.
10. Freier, K., Joos, S., Flechtenmacher, C., Devens, F., Benner, A., Bosch, F. X., Lichter, P., and Hofele, C. (2003) Tissue Microarray Analysis Reveals Site-specific Prevalence of Oncogene Amplifications in Head and Neck Squamous Cell Carcinoma. *Cancer Res. 63*, 1179–1182.
11. Harris, C. P., Lu, X. Y., Narayan, G., Singh, B., Murty, V., and Rao, P. H. (2003) Comprehensive molecular cytogenetic characterization of cervical cancer cell lines. *Genes, Chromosomes Cancer 36*, 233–241.
12. Dallafavera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. C., and Croce, C. M. (1982) Human C-Myc Onc Gene Is Located on the Region of Chromosome-8 That Is Translocated in Burkitt-Lymphoma Cells. *Proc. Natl. Acad. Sci. U.S.A. 79*, 7824–7827.
13. Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. (1982) Translocation of the C-Myc Gene into the Immunoglobulin Heavy-Chain Locus in Human Burkitt-Lymphoma and Murine Plasmacytoma Cells. *Proc. Natl. Acad. Sci. U.S.A. 79*, 7837–7841.
14. Nowell, P., Finan, J., Favera, R. D., Gallo, R. C., Arrushdi, A., Romanczuk, H., Selden, J. R., Emanuel, B. S., Rovera, G., and Croce, C. M. (1983) Association of Amplified Oncogene C-Myc with an Abnormally Banded Chromosome-8 in a Human-Leukemia Cell-Line. *Nature 306*, 494–497.
15. Marcu, K. B., Bossone, S. A., and Patel, A. J. (1992) Myc Function and Regulation. *Annu. Rev. Biochem. 61*, 809–860.
16. Marcu, K. B., Patel, A. J., and Yang, Y. (1997) Differential regulation of the c-MYC P1 and P2 promoters in the absence of functional tumor suppressors: Implications for mechanisms of deregulated MYC transcription. *C-Myc in B-Cell Neoplasia,* 47–56.
17. Berberich, S. J., and Postel, E. H. (1995) Puf/Nm23-H2/Ndpk-B Transactivates a Human C-Myc Promoter-Cat Gene Via a Functional Nuclease Hypersensitive Element. *Oncogene 10*, 2343–2347.
18. Simonsson, T., Pecinka, P., and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res. 26*, 1167–1172.
19. Collins, I., Weber, A., and Levens, D. (2001) Transcriptional Consequences of Topoisomerase Inhibition. *Mol. Cell. Biol. 21*, 8437–8451.
20. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A. 99*, 11593–11598.
21. Hurley, L. H., Von Hoff, D. D., Siddiqui-Jain, A., and Yang, D. (2006) Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Semin. Oncol. 33*, 498–512.
22. Yang, D. Z., and Hurley, L. H. (2006) Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides, Nucleotides Nucleic Acids 25*, 951–968.
23. Ou, T. M., Lu, Y. J., Zhang, C., Huang, Z. S., Wang, X. D., Tan, J. H., Chen, Y., Ma, D. L., Wong, K. Y., Tang, J. C. O., Chan, A. S. C., and Gu, L. Q. (2007) Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. *J. Med. Chem. 50*, 1465–1474.
24. Yang, D., and Okamoto, K. (2010) Structural insights into G-quadruplexes: Towards new anticancer drugs. *Future Med. Chem. 2*, 619–646.
25. Punchihewa, C., and Yang, D. (2009) Therapeutic Targets and Drugs II: G-Quadruplex and G-Quadruplex Inhibitors. In Cancer Drug Discovery and Development: Telomeres and Telomerase in Cancer (Hiyama, K., Ed.) pp 251−280, Humana Press, Totowa, NJ.
26. Seenisamy, J., Rezler, E. M., Powell, T. J., Tye, D., Gokhale, V., Joshi, C. S., Siddiqui-Jain, A., and Hurley, L. H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc. 126*, 8702–8709.
27. Ambrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. Z. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry 44*, 2048–2058.
28. Phan, A. T., Modi, Y. S., and Patel, D. J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc. 126*, 8710–8716.
29. Qin, Y., and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie 90*, 1149–1171.
30. Dai, J. X., Dexheimer, T. S., Chen, D., Carver, M., Ambrus, A., Jones, R. A., and Yang, D. Z. (2006) An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J. Am. Chem. Soc. 128*, 1096–1098.
31. Rachwal, P. A., Findlow, I. S., Werner, J. M., Brown, T., and Fox, K. R. (2007) Intramolecular DNA quadruplexes with different arrangements of short and long loops. *Nucleic Acids Res. 35*, 4214–4222.
32. Bugaut, A., and Balasubramanian, S. (2008) A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. *Biochemistry 47*, 689–697.
33. Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970) Oligonucleotide Interactions. 3. Circular Dichroism Studies of Conformation of Deoxyoligonucleotides. *Biopolymers 9*, 1059–1077.
34. Mergny, J. L., and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides 13*, 515–537.
35. Jing, N. J., Rando, R. F., Pommier, Y., and Hogan, M. E. (1997) Ion selective folding of loop domains in a potent anti-HIV oligonucleotide. *Biochemistry 36*, 12498–12505.
36. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T., and Helene, C. (1992) Kinetics and Thermodynamics of Triple-Helix Formation: Effects of Ionic-Strength and Mismatches. *Biochemistry 31*, 9269–9278.
37. Brown, N. M., Rachwal, P. A., Brown, T., and Fox, K. R. (2005) Exceptionally slow kinetics of the intramolecular quadruplex formed by the Oxytricha telomeric repeat. *Org. Biomol. Chem. 3*, 4153–4157.
38. Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res. 31*, 2097–2107.
39. Xu, Y., Sato, H., Sannohe, Y., Shinohara, K., and Sugiyama, H. (2008) Stable Lariat Formation Based on a G-Quadruplex Scaffold. *J. Am. Chem. Soc. 130*, 16470–16471.
40. Guo, K., Gokhale, V., Hurley, L. H., and Sun, D. (2008) Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res. 36*, 4598–4608.
41. Skolakova, P., Bednarova, K., Vorlickova, M., and Sagi, J. (2010) Quadruplexes of human telomere $dG_3(TTAG_3)_3$ sequences containing guanine abasic sites. *Biochem. Biophys. Res. Commun. 399*, 203–208.
42. De Cian, A., Guittat, L., Kaiser, M., Sacca, B., Amrane, S., Bourdoncle, A., Alberti, P., Teulade-Fichou, M. P., Lacroix, L., and Mergny, J. L. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods 42*, 183–195.
43. Cantor, C. R., and Schimmel, P. R. (1980) Biophysical Chemistry, W. H. Freedman and Co., New York.
44. Rachwal, P. A., and Fox, K. R. (2007) Quadruplex melting. *Methods 43*, 291–301.
45. Rachwal, P. A., Brown, T., and Fox, K. R. (2007) Sequence effects of single base loops in intramolecular quadruplex DNA. *FEBS Lett. 581*, 1657–1660.