

Effect of Interior Loop Length on the Thermal Stability and pK_a of i-Motif DNA

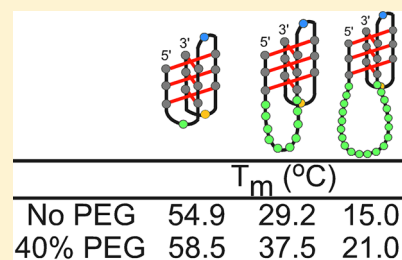
Samantha M. Reilly,[†] Rhianna K. Morgan,[‡] Tracy A. Brooks,[‡] and Randy M. Wadkins^{*,†}

[†]Department of Chemistry and Biochemistry, University of Mississippi, University, Mississippi 38677, United States

[‡]Department of BioMolecular Sciences, University of Mississippi, University, Mississippi 38677, United States

Supporting Information

ABSTRACT: The four-stranded i-motif (iM) conformation of cytosine-rich DNA is important in a wide variety of biochemical systems ranging from its use in nanomaterials to a potential role in oncogene regulation. An iM is stabilized by acidic pH that allows hemiprotonated cytidines to form a C·C⁺ base pair. Fundamental studies that aim to understand how the lengths of loops connecting the protonated C·C⁺ pairs affect intramolecular iM physical properties are described here. We characterized both the thermal stability and the pK_a of intramolecular iMs with differing loop lengths, in both dilute solutions and solutions containing molecular crowding agents. Our results showed that intramolecular iMs with longer central loops form at pHs and temperatures higher than those of iMs with longer outer loops. Our studies also showed that increases in thermal stability of iMs when molecular crowding agents are present are dependent on the loop that is lengthened. However, the increase in pK_a for iMs when molecular crowding agents are present is insensitive to loop length. Importantly, we also determined the proton activity of solutions containing high concentrations of molecular crowding agents to ascertain whether the increase in pK_a of an iM is caused by alteration of this activity in buffered solutions. We determined that crowding agents alone increase the apparent pK_a of a number of small molecules as well as iMs but that increases to iM pK_a were greater than that expected from a shift in proton activity.



	T_m (°C)		
No PEG	54.9	29.2	15.0
40% PEG	58.5	37.5	21.0

G4s and iMs are formed from guanine-rich and cytosine-rich sequences, respectively. Gellert et al.¹ in 1962 showed that four guanine bases can hydrogen bond with each other to form G-tetrads. Later, Arnott et al.² in 1974 showed that the array of G-tetrads stack to form G4s. The foundation of discovery of iMs was laid in 1963 when Langridge and Rich³ proposed that cytosines in poly-dC could hydrogen bond to form an ordered, hemiprotonated structure at low pH. However, it was only in 1993 that the four-stranded structure of iMs was determined (Figure 1).⁴ Curiously, the preponderance of genomic DNA with potential G4/iM sequences is found in genes that regulate cell growth.⁵ G4s and iMs are formed in G/C-rich DNA from promoters of genes involved in cell proliferation, including VEGF, KRAS, Bcl-2, pRb, PDGF-A, and MYC.^{6,7} Ever since these discoveries, the existence of G4s and iMs in living cells has remained a topic of great controversy. In 2013, Balasubramanian et al.⁸ published one of the first papers that confirmed G4s exist *in vivo*, which has been further supported by several additional reports on G4s *in vivo* in both DNA and RNA.^{9–13} No similar studies for iMs have been performed, so their existence *in vivo* is still yet to be determined. Unlike G4s, iMs have not yet been considered widely as stable *in vivo* because the pH in the nucleus is believed to be not much different from cytosolic pH (~ 7.3)^{14,15} while the pK_a of a typical iM is < 7.0 . However, recent studies from our lab and others^{16–18} have shown that iMs indeed form at neutral pH when molecular crowding agents are present. Because cells and cell nuclei are crowded with biopolymers, this finding suggests

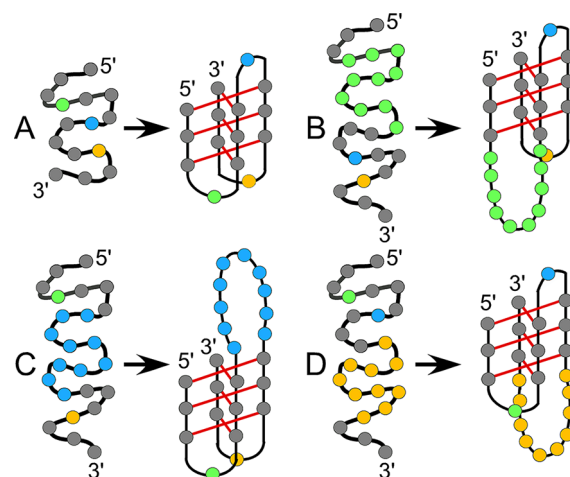


Figure 1. Representative folding of a random coil into an iM for oligos differing by loop length and position: (A) a reference structure (T1) in which all loops contain one thymine (colored spheres) and (B) first (green, Mod1 oligos), (C) central (blue, Mod2 oligos), and (D) third (yellow, Mod3 oligos) loops with 3–20 thymines.

Received: December 1, 2014

Revised: January 21, 2015

Published: January 26, 2015



iMs may be viable *in vivo* and deserve further study for comparison with the better-characterized G4 and Watson–Crick B-DNA structures.

In addition to their possible biological relevance, iMs are currently being used as a functional agent in biomaterials. Prior examples of iM use include biocompatible pH sensors and drug delivery vehicles. These applications take advantage of the iM's unique topology and proton dependence.^{19–23} For example, Song et al.²² reported an iM-based gold nanomaterial that could deliver the anticancer drug doxorubicin specifically to human cervical adenocarcinoma cells cultured *in vitro*. The drug is intercalated into duplex DNA containing mismatching pairs that are attached to a gold nanomaterial. Once the nanomaterial reaches the hypoxic cancer cells, the microenvironment's low pH causes the DNA to fold into an iM, releasing the intercalated doxorubicin. The properties of these and other nanomaterials are directly linked to iM biophysical properties, such as their thermal stability and pK_a . Thus, an improved understanding of the properties of iMs with varying characteristics will allow better control over the physical properties of iM-based nanomaterials.

Sequences that form iMs are usually derived from their corresponding G4 sequences that were located using Quadfinder²⁴ and/or Quadparser.²⁵ The default search in both Quadfinder and Quadparser is the sequence of $G_{3+N_{1-7}G_{3+N_{1-7}G_{3+N_{1-7}G_{3+N_{1-7}}}}$, where N stands for the loop residues and G for guanines. This default sequence has been based on multiple systematic studies that have shown that lengthening any loop of the G4 will decrease the thermal stability of the G4s, influence the type of G4 formed, and increase the potassium dependence of G4 formation.^{26–30} This search algorithm has also been used to find potential iMs on the opposite strand; however, the most stable G4 and iM formed from the same DNA duplex sequence of the c-MYC gene do not share the same loop lengths with the iM having much longer loops than its G4 counterpart.³¹ Studies performed with DNA hairpins^{32–35} show a decreased thermal stability with an increasing loop length; however, hairpins can reach up to 20 bases in length and still be stable at physiological temperatures.^{32,34} The loop length effects in iMs have not been well characterized and are the subject of this report.

To determine the optimal loop length to use in finding potential iMs, we examined the effect of loop length on an iM's structural stability. We determined the effects of loop length in dilute and molecularly crowded buffers that more closely mimic intracellular conditions.^{18,36–40} We examined 16 different oligonucleotide sequences (oligos) differing in both loop length and position (Table 1 and Figure 1). Polyethylene glycol (PEG) was employed as the crowding agent because it is the most commonly used molecular crowding agents in the literature on both iMs and G4s.^{18,36–40} We used 40% PEG (w/w) to mimic cells because biomolecules can occupy 20–40% of

the cell volume.^{17,18,41} Using circular dichroism (CD) spectroscopy and ultraviolet (UV) absorbance spectroscopy, we determined the thermal stability and pK_a of each oligo under dilute and molecularly crowded conditions, where the size of PEGs used varied in average molecular weight from 300 to 3350. We did not go beyond 3350 because at least one previous report¹⁷ states that PEGs with sizes above 4000 can bind to DNA secondary structures. Earlier studies reported that a high percentage (>30%) of other hydrophilic molecules, including PEG, shifted the apparent pH of buffered solutions.^{42–46} To observe the effect of PEG on the apparent pH, we analyzed five pH indicators and cytidine monophosphate to determine their pK_a values with and without PEG. Any increase in the pK_a of iMs greater than those arising from a change in proton activity (as observed with the indicators and cytidine monophosphate) can be attributed to PEG's crowding effect. Via the collection of key details about iM sequence-dependent physical properties, our results may help guide the creation of novel iM-based nanomaterials with carefully tuned pH responses. Further, our results provide guidance for adjusting search algorithms to find potential iMs in genomic DNA sequences other than those that are complementary to G4 sequences.

MATERIALS AND METHODS

Oligos. All DNA oligos (Table 1) were synthesized through standard solid-phase chemistry by Midland Certified Reagent Co., Inc. (Midland, TX). All other reagents were purchased from Fisher Scientific (Waltham, MA). The oligos were stored in 10 mM Tris, 1 mM EDTA buffer (pH 8.0) until they were used. The oligo sequence chosen (5'-CCC T CCC T CCC T CCC-3') is a nonphysiological one chosen to minimize variables. The loop length substitutions were thymine repeats of varying length (1, 3, 5, 10, 15, and 20 thymines). Unless specifically stated otherwise, final solutions for all experiments included 4 μ M DNA in 30 mM sodium cacodylate buffers at varying pHs (5.4–8.0) either with or without 40% PEG of the sizes indicated in the text. Solutions were initially heated to ~80 °C for 10 min and then cooled to room temperature to remove any multimeric duplex DNA present.

pK_a Determination. To determine the pK_a for iM folding (i.e., the pH at which 50% of the oligo is folded into the iM), CD spectra of DNA solutions at 20 °C were collected from 250 to 320 nm on an Olis DSM 20 circular dichroism (CD) instrument fitted with a Peltier heat block (Olis, Inc., Bogart, GA). An integration time as a function of high voltage was used. The CD signals observed at 298 nm were then plotted against pH, and then eq 1 was applied to obtain pK_a and the cooperativity parameter.

$$\text{signal}_{\text{total}} = \frac{\text{signal}_{\text{folded}} - \text{signal}_{\text{unfolded}}}{1 + 10^{\text{cooperativity} \times (\text{pH} - \text{p}K_a)}} + \text{signal}_{\text{unfolded}} \quad (1)$$

Thermal Stability and Denaturation of iMs. Melting of iMs between 20 and 80 °C was assessed by monitoring CD signals and/or UV absorbance at 1 °C intervals for all oligos (4 μ M) at pH 5.4 with and without 40% PEG. After iM melting had been verified via CD, repeated thermal melts were also performed on a Cary 100 UV–visible spectrometer (Agilent Technologies, Santa Clara, CA) at 260 nm. The thermal denaturations were similar on either spectrometer. With an assumed two-state equilibrium model (as suggested by an isodichroic point at ~280 nm), we determined the melting

Table 1. Oligos Used and Their Respective Sequences^a

name	5'-sequence-3'
T1	CCC T CCC T CCC T CCC
Mod1TX	CCC X CCC T CCC T CCC
Mod2TX	CCC T CCC X CCC T CCC
Mod3TX	CCC T CCC T CCC X CCC

^aX is the location and number of the thymine repeats (3, 5, 10, 15, or 20) in the loop modification.

temperature (T_m). All fractions were normalized such that the data ranged from 0 and 1 prior to fitting. The change in free energy at 37 °C ($\Delta G^\circ_{37\text{ }^\circ\text{C}}$) was calculated by eq 2

$$K = \frac{[N]}{[U]} = e^{-\Delta G^\circ/RT} \quad (2)$$

where $[N]$ and $[U]$ are the concentrations of the iM and random coil forms of DNA, respectively, R is the gas constant, and T is the physiological temperature (310 K).⁴⁷ In addition, to ensure the iMs formed were indeed intramolecular, a range of DNA concentrations (0.6, 4, and 20 μM) were analyzed for all oligos. All oligos were found to have the same T_m and $\Delta G^\circ_{37\text{ }^\circ\text{C}}$ regardless of concentration, suggesting all oligos form an intramolecular iM structure.

Relation of Measured pH to Proton Activity. The method of Gaboriaud⁴⁵ was used to compare apparent pH with proton activity in PEG solutions. The pH was measured by a high-performance, micro glass-body, Tris-compatible, combination pH electrode with a platinum junction and 3 M potassium chloride filling solution (model pHE-11, GeneMate, BioExpress, Kaysville, UT). Five pH indicators were used: bromothymol blue, 4-nitrophenol, 3,4-dinitrophenol, methyl red, and phenol red. In addition, to directly compare to iMs, cytidine 5'-monophosphate was also used. The pH titration of each indicator was performed on a Cary 100 UV-vis spectrometer over a wavelength range of 230–800 nm. Each indicator solution was made with a 30 mM sodium cacodylate buffer with or without 40% PEG-300 or PEG-3350. The sample size was 3 mL, and 0.05 M NaOH was titrated into the sample at increments of 0.02–0.10 mL, depending on the magnitude of the previous pH change. Once the spectra no longer changed with pH, the maximal absorbance values of the acidic form of each indicator, corrected for dilution factors, were plotted against their respective measured pH to determine the pK_a .

RESULTS AND DISCUSSION

pK_a Determination. The pK_a of each oligo is given in Table 2 with representative graphs shown in Figure 2. The trend in pK_a values observed for Mod2 is distinctly different from that of the other loop modifications (Mod1 and Mod3). Mod2 oligos are able to maintain iM structure at higher pH and with longer loop lengths than the other modifications. For example, with 10 bases in their respective loops, Mod2T10 has a pK_a of 6.62 while Mod1T10 and Mod3T10 have a pK_a of 6.13. The pK_a remains the same for the Mod2 oligos until more than 10 bases have been added to the loop. These data suggest that the length of the central loop has a weaker effect on the pK_a than the outer loops. The cooperativity parameter of all oligos, determined by fitting to eq 1, did not change significantly (~ 3 for all iMs formed), indicating that extending the loops does not affect how the iM unfolds when shifting from an acidic to a more basic pH. These findings suggest that central loops (e.g., Mod2) shorter than 15 bases and outer loops (e.g., Mod1 and Mod3) shorter than 10 bases are potentially viable *in vivo*. This is remarkably different from the results for G4s where, in the absence of other structural features in the loop, loop lengths are usually limited to seven bases to be potentially viable *in vivo*.

Effects of PEG on the Apparent pH. Given our data¹⁶ and those of others^{17,18} on the effect of PEG on iM pK_a , we investigated the difference between the pH determined by electrode measurement and the activity of protons in the PEG

Table 2. pK_a Values of All Oligos in 30 mM Sodium Cacodylate with or without 40% PEG-300^a

	pK_a		ΔpK_a
	no PEG	40% PEG	
T1	6.67 \pm 0.01	7.03 \pm 0.05	0.36
Mod1T3	6.66 \pm 0.01	7.04 \pm 0.05	0.38
Mod1T5	6.25 \pm 0.04	6.71 \pm 0.12	0.46
Mod1T10	6.14 \pm 0.02	6.54 \pm 0.05	0.40
Mod1T15	nd ^b	nd ^b	nd ^b
Mod1T20	nd ^b	nd ^b	nd ^b
Mod2T3	6.66 \pm 0.03	7.16 \pm 0.05	0.50
Mod2T5	6.68 \pm 0.04	7.14 \pm 0.05	0.46
Mod2T10	6.62 \pm 0.02	6.98 \pm 0.07	0.36
Mod2T15	6.33 \pm 0.02	6.71 \pm 0.06	0.38
Mod2T20	5.98 \pm 0.05	6.48 \pm 0.06	0.50
Mod3T3	6.70 \pm 0.03	7.00 \pm 0.06	0.30
Mod3T5	6.27 \pm 0.02	6.63 \pm 0.08	0.36
Mod3T10	6.12 \pm 0.02	6.42 \pm 0.05	0.30
Mod3T15	nd ^b	nd ^b	nd ^b
Mod3T20	nd ^b	nd ^b	nd ^b

^aThe ΔpK_a values are also shown to emphasize the difference upon adding 40% PEG-300. ^bThe T_m was ≤ 20 °C; thus, the pK_a was not determined at 20 °C.

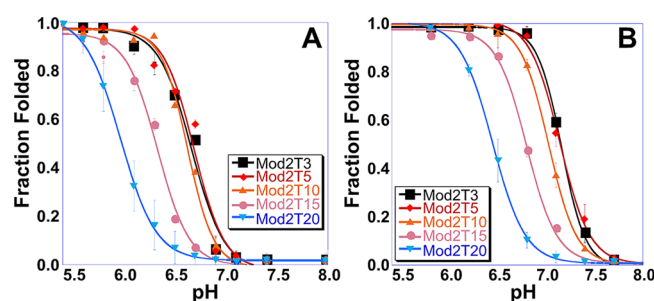


Figure 2. Determination of pK_a values of Mod2 oligos by monitoring the CD signal at 298 nm under (A) dilute buffer conditions and (B) molecularly crowded conditions (40% PEG-300). The pK_a values for these and the Mod1 and Mod3 oligos are listed in Table 2.

solutions as detected by small molecules. We examined five well-known pH indicators and cytidine monophosphate by monitoring the change in UV-vis absorbance with the change in pH in 30 mM sodium cacodylate buffer with and without 40% PEG-300. Changes in the pK_a observed for these indicators in PEG solutions were attributed to a change in proton activity, because molecular crowding should not significantly affect the pK_a of small molecules such as the indicators. As shown in Table 3, the observed pK_a of indicators consistently increased by 0.25 upon addition of 40% PEG-300 to the buffer. We also tested the indicators with 40% PEG-3350 to see if the pK_a is affected by the size of PEG; no additional differences in pK_a were noted, suggesting the change in activity is linked to the percentage of PEG (w/w) in solution and not PEG size.

Effects of PEG on iM pK_a . The pK_a values of all oligos show increase upon addition of PEG (Table 2); however, the increase in pK_a is not different between oligos of different loop lengths or positions. The lack of a difference in pK_a between these oligos suggests that the increase in pK_a caused by molecular crowding conditions is not dependent on loop length. On the basis of our oligos, the increase in pK_a is likely associated with the number of C–C⁺ bonds, which all these

Table 3. Determination of pK_a Values for Five pH Indicators and Cytidine Monophosphate with Two Different Sized PEGs at 40% (w/w) in 30 mM Sodium Cacodylate Buffer

	pK_a					
	4-nitrophenol	bromo-thymol blue	3,4-dinitrophenol	methyl red	phenol red	cytidine monophosphate
literature value ^a	7.1	7.1	5.4	5.0	7.9	4.3
no PEG	7.10 ± 0.04	7.07 ± 0.03	5.31 ± 0.02	5.04 ± 0.04	7.98 ± 0.05	4.35 ± 0.06
PEG-300	7.33 ± 0.04	7.35 ± 0.04	5.52 ± 0.04	5.32 ± 0.03	8.27 ± 0.02	4.57 ± 0.04
PEG-3350	7.35 ± 0.06	7.36 ± 0.04	5.53 ± 0.02	5.28 ± 0.03	8.25 ± 0.02	4.61 ± 0.02
ΔpK_a	0.24	0.28	0.22	0.26	0.28	0.24
ΔpK_a (avg)	0.25 ± 0.03					

^aFrom refs 50–55 from left to right, respectively.

Table 4. T_m and $\Delta G^\circ_{37^\circ\text{C}}$ Values of All Oligos in 30 mM Sodium Cacodylate with or without 40% PEG-300^a

	T_m (°C)		ΔT_m (°C)	$\Delta G^\circ_{37^\circ\text{C}}$ (kcal/mol)		$\Delta\Delta G^\circ_{37^\circ\text{C}}$ (kcal/mol)
	no PEG	40% PEG		no PEG	40% PEG	
T1	54.9 ± 0.4	58.5 ± 0.5	3.6	−2.2	−2.9	−0.7
Mod1T3	47.7 ± 0.4	53.3 ± 0.5	5.6	−0.9	−2.0	−1.1
Mod1T5	41.9 ± 0.3	48.8 ± 0.4	6.9	−0.6	−1.3	−0.7
Mod1T10	29.2 ± 0.5	37.5 ± 0.5	6.2	0.7	0.0	−0.7
Mod1T15	19.5 ± 0.4	25.8 ± 1.0	6.3	3.4	1.6	−1.9
Mod1T20	15.0 ± 0.5	21.0 ± 0.5	6.0	4.5	2.3	−2.2
Mod2T3	54.7 ± 0.3	66.7 ± 0.7	12.0	−2.3	−3.3	−1.1
Mod2T5	54.5 ± 0.6	64.0 ± 0.6	9.5	−2.7	−4.0	−1.3
Mod2T10	47.1 ± 0.2	55.8 ± 0.8	8.7	−1.6	−2.8	−1.2
Mod2T15	40.6 ± 0.1	49.8 ± 0.7	9.2	−0.7	−1.7	−0.9
Mod2T20	36.3 ± 0.6	46.2 ± 0.8	9.9	−0.6	−1.5	−0.9
Mod3T3	47.0 ± 0.4	55.4 ± 0.3	8.4	−0.8	−2.0	−1.3
Mod3T5	38.9 ± 0.6	45.5 ± 0.9	6.6	−0.3	−0.9	−0.7
Mod3T10	26.9 ± 0.1	34.3 ± 1.3	7.4	1.1	0.3	−0.8
Mod3T15	13.2 ± 0.3	21.2 ± 1.4	8.0	5.6	3.2	−2.4
Mod3T20	10.2 ± 0.2	17.9 ± 0.3	7.7	5.5	3.2	−2.2

^aThe ΔT_m and $\Delta\Delta G_{37^\circ\text{C}}$ are also shown to emphasize the difference upon addition of 40% PEG-300.

iMs have in common. In addition, there was no significant change in the cooperativity of the iMs with addition of PEG-300 (~3 in both dilute and crowded buffers). The increase in pK_a with 40% PEG-300 ($+0.40 \pm 0.07$) is larger than the change observed with small molecules [$+0.25 \pm 0.03$ (Table 3)], suggesting that a change in proton activity alone does not account for the change in iM pK_a . We suggest that the additional increase in pK_a corresponds to the physical effects of molecular crowding, inhibiting the unfolding of the structure.

Next, we examined the effect of varying PEG size on pK_a with T1, Mod1T3, and Mod1T10. Mod1 oligos were chosen, as they are symmetric with Mod3 oligos, and they have distinct differences in thermal stability and pK_a values among short (1 base), medium (3–5 bases), and long (10 bases) loops. The pK_a increased by ~0.4 upon addition of PEG-300 and by ~0.7 upon addition of PEG-3350, showing that PEG size has a direct influence on the pK_a of iMs. This increase is constant across all loop lengths tested, further suggesting that the pK_a increase from crowding agents is not dependent on loop length. We conclude that the changes in pK_a we observed are due to molecular crowding alone because the change in proton activity (as determined from the indicators to be $+0.25$) is the same for both PEG-300 and PEG-3350 (Table 3). We also calculated the excluded volume for each PEG–DNA complex between 0 and 40% PEG (Table S1 of the Supporting Information), using the approach of Spink and Chaires.⁴⁸ The T_m increase correlates well with excluded volume over the PEG size and loop lengths used, but the pK_a does not, suggesting that excluded volume

alone cannot explain the effects that PEG and loop length have on the pK_a of an iM.

Thermal Denaturation of Folded i-Motifs. Thermal denaturations were performed on all folded oligos from Table 1. From the recorded spectra, the melting temperature (T_m) of each oligo was determined (Table 4 and Figure 3). All oligos were found to have the same T_m and $\Delta G^\circ_{37^\circ\text{C}}$ values across a range of DNA concentrations (0.6–20 μM), suggesting all oligos form an intramolecular iM structure. The trends observed for the T_m values of Mod1 and Mod3 are similar as loop length increases; however, these T_m values are distinctly lower than those of Mod2 oligos of the same loop length.

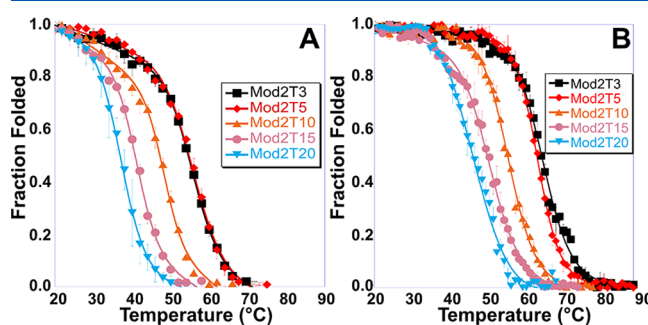


Figure 3. Thermal denaturations of Mod2 oligos in (A) dilute buffer and (B) molecularly crowded conditions (40% PEG-300). Both Mod1 and Mod3 showed similar trends. The T_m data are listed in Table 4.

Mod2 oligos can incorporate loops much longer than Mod1 and Mod3. For example, the Mod2 oligos can have loops up to five bases long and not incur a change in thermal stability. The trends in T_m found for these oligos suggest that longer central loops are more stable in iMs than in G4 oligos, displaying a trend more similar to those observed by DNA hairpins. In the Mod1 and Mod3 oligos, which behave similar to G4 oligos, the loop (Figure 1B,D) may be sterically hindered by the other loop, leading to destabilization. On the other hand, the central loop (Figure 1C) does not encounter such steric hindrance in our iM models and would not in other iM systems unless 5' and/or 3' overhangs are present. All enthalpies of folding determined by van't Hoff analysis of T_m data were approximately -5.0 kcal/mol per C-C⁺ bond, which corresponds well with previously published melting enthalpies of iMs⁴⁹ and suggests little or no bonding interactions arise because of the loops. When loops were lengthened, the entropy of folding decreased from -0.11 kcal mol⁻¹ K⁻¹ for a loop length of 1 to -0.20 kcal mol⁻¹ K⁻¹ for a loop length of 20, regardless of loop location. As a better indicator of biological viability, we calculated the standard state ΔG° of folding for each oligo at the physiological temperature of 37 °C ($\Delta G^\circ_{37\text{ °C}}$ in Table 4). Of all iMs examined, Mod2 iMs were most biologically viable, showing a negative $\Delta G^\circ_{37\text{ °C}}$ up to a loop length of 20 thymines (Mod2T20). The $\Delta G^\circ_{37\text{ °C}}$ for Mod2 oligos becomes less negative when the central loop is longer than 5 bases; however, the $\Delta G^\circ_{37\text{ °C}}$ values still suggest that long central loops could exist *in vivo*. In contrast, iMs formed by the Mod1 and Mod3 oligos become unfavorable at 37 °C when there are more than 5 thymines in the loop, suggesting long outer loops are not biologically viable.

Multiple studies^{16–18} have shown that adding a molecular crowding agent increases the T_m values of select iMs. We studied the effects of molecular crowding on the T_m and folding as a function of loop length (Table 4 and Figure 3). The melting temperatures of all oligos increase in the presence of 40% PEG-300 as shown in Table 4; however, the magnitude of this increase in T_m is dependent on which loop is elongated. For example, in 40% PEG-300, when the loops of the Mod1 or Mod3 oligos are elongated, the T_m increases by 6–8 °C, but if the loops of the Mod2 oligos are elongated, the T_m increases by 9–12 °C. These different trends in PEG solutions suggest that longer central loops are more stable in molecularly crowded conditions than their outer loop counterparts. We again interpret this as the central loop not being sterically hindered by the other loops (Figure 1B–D), which allows for Mod2 oligos to fold into and retain the iM structure more readily than their Mod1 or Mod3 counterparts. The enthalpy of folding did not change significantly between the dilute buffer and 40% PEG-300 solutions. The $\Delta G^\circ_{37\text{ °C}}$ observed under crowding conditions shows a trend similar to that of T_m values, with the $\Delta G^\circ_{37\text{ °C}}$ values of the Mod1 and Mod3 oligos ranging from -2 to 3 kcal/mol and the $\Delta G^\circ_{37\text{ °C}}$ values of the Mod2 oligos ranging from -4 to -1 kcal/mol.

We used oligos T1, Mod1T3, and Mod1T10 to further examine the changes in melting temperature as the size of PEG increased (Table 5). The thermal stability of all oligos showed an increase in T_m in different PEG solutions until the PEG size was greater than a molecular weight of 600, which is in agreement with a previous study.¹⁷ At PEG sizes above 600, the T_m remains constant. We surmise that PEGs with sizes of <600 can readily fit into the space the iM needs to fold, preventing the iM from folding. However, the physical crowding of PEG

Table 5. T_m Values of Three iMs across Different Sized PEGs at 40% (w/w) in 30 mM Sodium Cacodylate at pH 5.4

	T_m (°C)		
	T1	Mod1T3	Mod1T10
no PEG	54.9 ± 0.4	47.7 ± 0.4	29.2 ± 0.5
PEG-300	58.5 ± 0.5	53.3 ± 0.5	37.5 ± 0.5
PEG-600	59.1 ± 1.4	56.7 ± 0.9	44.0 ± 1.2
PEG-1500	65.3 ± 0.3	55.3 ± 1.0	47.4 ± 0.5
PEG-3350	65.0 ± 0.5	58.0 ± 0.5	48.8 ± 1.5

outside of the space the iM folds still forces the iM together; thus, the increase in T_m is a combination of these two effects. The PEGs with sizes of >600 cannot fit easily into the folding space and thus do not hinder iM folding. In addition, unlike pK_a , the excluded volume (Table S1 of the Supporting Information) correlated very well with the observed increase in T_m as PEG size and loop length increased. Consequently, the increase in T_m comes purely from the physical crowding of the space around the iM by PEG.

CONCLUSION

In this paper, we demonstrate that iMs with five or more thymidines in the central loop can form at pH values higher than the pH values those with loops near the termini of the iM-forming regions. We also establish that the increase in the pK_a values of iMs (+0.40) upon addition of PEG is independent of loop length and position. From our indicator studies, we found that a portion of this increase in pK_a upon addition of PEG is from a change in proton activity, but crowding appears to contribute additional iM stability at higher pH. Our studies also demonstrated that the thermal stabilities of the iMs with three or more bases in the central loop are much higher (~ 15 °C) than those of their outer loop counterparts when no 5' or 3' overhangs are present, displaying behavior more similar to that of DNA hairpins than that of G4s. When the iMs are in molecularly crowded conditions, we found those with outer loops exhibit a T_m increase of 6–8 °C while those with central loops exhibit a T_m increase of 9–12 °C. Together, these trends show that not all loops can be considered equal in iM formation and indicate that iMs with long central loops have a melting temperature and a pK_a higher than previously thought. Adjusting loop length may allow for better control of the physical properties of iM-based nanomaterials because we have demonstrated that small changes in the loop lengths greatly affect the pH and temperatures at which iM formation occurs. In addition, we also have shown that iMs with long central loops are potentially viable *in vivo* at physiological temperature, even with loop lengths of 20 bases; however, similar studies with 5' and 3' overhangs present are needed before this can be readily assumed for iMs in genomic DNA. This finding suggests that the optimal search algorithm for locating potential iM-forming sequences in genomic databases may need to be adjusted to be different from algorithms used for finding potential G4 sequences.

ASSOCIATED CONTENT

Supporting Information

Additional method of calculation and calculated excluded volume of PEG–iM DNA complexes (Table S1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rwadkins@olemiss.edu.

Funding

The National Science Foundation under Mississippi EPSCoR Grant EPS-0903787 and National Institutes of Health Grant R15 CA173667-01A1 (Tracy Brooks, PI; Randy Wadkins, co-PI) supported this research.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

iM, i-motif; C, cytosine; G4, G-quadruplex; G, guanine; T, thymine; VEGF, vascular endothelial growth factor; kRAS, Kirsten rat sarcoma viral oncogene homologue; Bcl-2, B-cell lymphoma 2; pRb, retinoblastoma protein; PDGF-A, platelet-derived growth factor subunit A; MYC, V-Myc avian myelocytomatosis viral oncogene homologue; PEG, polyethylene glycol; CD, circular dichroism.

REFERENCES

- (1) Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.* 48, 2013–2018.
- (2) Arnott, S., Chandrasekaran, R., and Martilla, C. M. (1974) Structures for polyinosinic acid and polyguanylic acid. *Biochem. J.* 141, 537–543.
- (3) Langridge, R., and Rich, A. (1963) Molecular structure of helical polycytidylic acid. *Nature* 198, 725–728.
- (4) Gehring, K., Leroy, J. L., and Guéron, M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363, 561–565.
- (5) Eddy, J., and Maizels, N. (2006) Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* 34, 3887–3896.
- (6) Brooks, T. A., and Hurley, L. H. (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer* 9, 849–861.
- (7) Brooks, T. A., Kendrick, S., and Hurley, L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* 277, 3459–3469.
- (8) Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186.
- (9) Biffi, G., Di Antonio, M., Tannahill, D., and Balasubramanian, S. (2014) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* 6, 75–80.
- (10) Biffi, G., Tannahill, D., Miller, J., Howat, W. J., and Balasubramanian, S. (2014) Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. *PLoS One* 9, e102711.
- (11) Henderson, A., Wu, Y., Huang, Y. C., Chavez, E. A., Platt, J., Johnson, F. B., Brosh, R. M., Sen, D., and Lansdorp, P. M. (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* 42, 860–869.
- (12) Lam, E. Y. N., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* 4, 1796.
- (13) Yang, T. L., Lin, L., Lou, P. J., Chang, T. C., and Young, T. H. (2014) Detection of cell carcinogenic transformation by a quadruplex DNA binding fluorescent probe. *PLoS One* 9, e86143.
- (14) Bright, G. R., Fisher, G. W., Rogowska, J., and Taylor, D. L. (1987) Fluorescence ratio imaging microscopy: Temporal and spatial measurements of cytoplasmic pH. *J. Cell Biol.* 104, 1019–1033.
- (15) Llopis, J., McCafferty, J. M., Miyawaki, A., Farquhar, M. G., and Tsien, R. Y. (1998) Measurement of cytosolic, mitochondrial, and

Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6803–6808.

(16) Bhavsar-Jog, Y. P., Van Dornshuld, E., Brooks, T. A., Tschumper, G. S., and Wadkins, R. M. (2014) Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry* 53, 1586–1594.

(17) Cui, J., Waltman, P., Le, V. H., and Lewis, E. A. (2013) The effect of molecular crowding on the stability of human c-MYC promoter sequence i-motif at neutral pH. *Molecules* 18, 12751–12767.

(18) Rajendran, A., Nakano, S., and Sugimoto, N. (2010) Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* 46, 1299–1301.

(19) Liu, Z., Li, Y., Tian, C., and Mao, C. (2013) A smart DNA tetrahedron that isothermally assembles or dissociates in response to the solution pH value changes. *Biomacromolecules* 14, 1711–1714.

(20) Chen, L., Di, J., Cao, C., Zhao, Y., Ma, Y., Luo, J., Wen, Y., Song, W., Song, Y., and Jiang, L. (2011) A pH-driven DNA nanoswitch for responsive controlled release. *Chem. Commun.* 47, 2850–2852.

(21) Modi, S., Swetha, M. G., Goswami, D., Gupta, G. D., Mayor, S., and Krishnan, Y. (2009) A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.* 4, 325–330.

(22) Song, L., Ho, V. H. B., Chen, C., Yang, Z., Liu, D., Chen, R., and Zhou, D. (2013) Efficient, pH-triggered drug delivery using a pH-responsive DNA-conjugated gold nanoparticle. *Adv. Healthcare Mater.* 2, 275–280.

(23) Wang, C., Du, Y., Wu, Q., Xuan, S., Zhou, J., Song, J., Shao, F., and Duan, H. (2013) Stimuli-responsive plasmonic core–satellite assemblies: i-motif DNA linker enabled intracellular pH sensing. *Chem. Commun.* 49, 5739–5741.

(24) Scaria, V., Hariharan, M., Arora, A., and Maiti, S. (2006) Quadfinder: Server for identification and analysis of quadruplex-forming motifs in nucleotide sequences. *Nucleic Acids Res.* 34, 683–685.

(25) Huppert, J. L., and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* 33, 2908–2916.

(26) Guedin, A., Gros, J., Alberti, P., and Mergny, J. (2010) How long is too long? Effect of loop size on G-quadruplex stability. *Nucleic Acids Res.* 38, 7858–7868.

(27) Hazel, P., Huppert, J., Balasubramanian, S., and Neidle, S. (2004) Loop-length dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* 126, 16405–16415.

(28) 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.

(29) Risitano, A., and Fox, K. R. (2004) Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res.* 32, 2598–2606.

(30) Smargiasso, N., Rosu, F., Hsia, W., Colson, P., Baker, E. S., Bowers, M. T., De Pauw, E., and Gabelica, V. (2008) G-quadruplex DNA assemblies: Loop length, cation identity, and multimer formation. *J. Am. Chem. Soc.* 130, 10208–10216.

(31) Sun, D., and Hurley, L. H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression. *J. Med. Chem.* 52, 2863–2874.

(32) Goddard, N. L., Bonnet, G., Krichesky, O., and Libchaber, A. (2000) Sequence dependent rigidity of single stranded DNA. *Phys. Rev. Lett.* 85, 2400–2403.

(33) Kuznetsov, S. V., Ren, C. C., Woodson, S. A., and Ansari, A. (2008) Loop dependence of the stability and dynamics of nucleic acid hairpins. *Nucleic Acids Res.* 36, 1098–1112.

(34) Nayak, R. K., and Van Orden, A. (2013) Counterion and Polythymidine Loop-Length-Dependent Folding and Thermodynamic Stability of DNA Hairpins Reveal the Unusual Counterion-Dependent Stability of Tetraloop Hairpins. *J. Phys. Chem. B* 117, 13956–13966.

- (35) Rentzeperis, D., Alessi, K., and Marky, L. A. (1993) Thermodynamics of DNA hairpins: Contribution of loop size to hairpin stability and ethidium binding. *Nucleic Acids Res.* 21, 2683–2689.
- (36) Kan, Z. Y., Lin, Y., Wang, F., Zhuang, X. Y., Zhao, Y., Pang, D. W., Hao, Y. H., and Tan, Z. (2007) G-quadruplex formation in human telomeric (TTAGGG)₄ sequence with complementary strand in close vicinity under molecularly crowded condition. *Nucleic Acids Res.* 35, 3646–3653.
- (37) Kan, Z. Y., Yao, Y., Wang, P., Li, X. h., Hao, Y. h., and Tan, Z. (2006) Molecular crowding induces telomere G-quadruplex formation under salt-deficient conditions and enhances its competition with duplex formation. *Angew. Chem., Int. Ed.* 45, 1629–1632.
- (38) Miyoshi, D., Matsumura, S., Nakano, S.-i., and Sugimoto, N. (2004) Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc.* 126, 165–169.
- (39) Zheng, K.-w., Chen, Z., Hao, Y.-h., and Tan, Z. (2010) Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic Acids Res.* 38, 327–338.
- (40) Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z., and Li, C. (2009) Human telomeric G-quadruplex formed from duplex under near physiological conditions: Spectroscopic evidence and kinetics. *Biochimie* 91, 1104–1111.
- (41) Miyoshi, D., Nakao, A., and Sugimoto, N. (2002) Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry* 41, 15017–15024.
- (42) Travers, F., Douzou, P., Pederson, T., and Gunsalus, I. C. (1975) Ternary solvents to investigate proteins at sub-zero temperatures. *Biochimie* 57, 43–48.
- (43) Arakawa, T., and Timasheff, S. N. (1985) Mechanism of polyethylene glycol interaction with proteins. *Biochemistry* 24, 6756–6762.
- (44) Chuy, S., and Bell, L. N. (2006) Buffer pH and pK_a values as affected by added glycerol and sucrose. *Food Res. Int.* 39, 342–348.
- (45) Gaboriaud, R. (1966) Sur le comportement des acides non chargés dans les milieux eau-méthanol. *C. R. Acad. Sci.* 263, 911–914.
- (46) Hui-Bon-Hoa, G., and Douzou, P. (1973) Ionic strength and protonic activity of supercooled solutions used in experiments with enzyme systems. *J. Biol. Chem.* 248, 4649–4654.
- (47) Bruylants, G., Wouters, J., and Michaux, C. (2005) Differential scanning calorimetry in life science: Thermodynamics, stability, molecular recognition and application in drug design. *Curr. Med. Chem.* 12, 2011–2020.
- (48) Spink, C. H., and Chaires, J. B. (1999) Effects of hydration, ion release, and excluded volume on the melting of triplex and duplex DNA. *Biochemistry* 38, 496–508.
- (49) Kaushik, M., Suehl, N., and Marky, L. A. (2007) Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, d(C₃TA₂)₄. *Biophys. Chem.* 126, 154–164.
- (50) Schwarzenbach, R. P., Stierli, R., Folsom, B. R., and Zeyer, J. (1988) Compound properties relevant for assessing the environmental partitioning of nitrophenols. *Environ. Sci. Technol.* 22, 83–92.
- (51) Crofts, A., and Jackson, J. (1969) Bromothymol blue and bromocresol purple as indicators of pH changes in chromatophores of *Rhodospirillum rubrum*. *Eur. J. Biochem.* 10, 226–237.
- (52) Espinosa, S., Bosch, E., and Rosés, M. (2002) Retention of ionizable compounds in high-performance liquid chromatography: 14. Acid–base pK values in acetonitrile–water mobile phases. *J. Chromatogr. A* 964, 55–66.
- (53) Tobey, S. W. (1958) The acid dissociation constant of methyl red. A spectrophotometric measurement. *J. Chem. Educ.* 35, 514.
- (54) Berthois, Y., Katzenellenbogen, J. A., and Katzenellenbogen, B. S. (1986) Phenol red in tissue culture media is a weak estrogen: Implications concerning the study of estrogen-responsive cells in culture. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2496–2500.
- (55) Flogel, M., and Biltonen, R. L. (1975) The pH dependence of the thermodynamics of the interaction of 3'-cytidine monophosphate with ribonuclease A. *Biochemistry* 14, 2610–2615.