Received: 20 March 2020 | Revised: 6 June 2020 | Accepted: 11 June 2020

DOI: 10.1002/bip.23389

FULL PAPER

Biopolymers WILEY

Hysteresis in poly-2′-deoxycytidine i-motif folding is impacted by the method of analysis as well as loop and stem lengths

R. Aaron Rogers | Madeline R. Meyer | Kayla M. Stewart | Gabriela M. Eyring | Aaron M. Fleming | Cynthia J. Burrows

Department of Chemistry, University of Utah, Salt Lake City, Utah, U.S.A.

Correspondence
Cynthia J. Burrows, Department of Chemistry, University of Utah, 315 S. 1400 East, Salt Lake City, UT 84112-0850, U.S.A.
Email: burrows@chem.utah.edu

Funding information
National Cancer Institute, Grant/Award Number: Cancer Center Support Grant P30 CA042014; U. S. National Science Foundation, Grant/Award Numbers: CHE1659579, CHE1808745

Abstract

In DNA, i-motif (iM) folds occur under slightly acidic conditions when sequences rich in 2′-deoxycytidine (dC) nucleotides adopt consecutive dC self base pairs. The pH stability of an iM is defined by the midpoint in the pH transition (pH_T) between the folded and unfolded states. Two different experiments to determine pH_T values via circular dichroism (CD) spectroscopy were performed on poly-dC iMs of length 15, 19, or 23 nucleotides. These experiments demonstrate two points: (1) pH_T values were dependent on the titration experiment performed, and (2) pH-induced denaturing or annealing processes produced isothermal hysteresis in the pH_T values. These results in tandem with model iMs with judicious mutations of dC to thymidine to favor particular folds found the hysteresis was maximal for the shorter poly-dC iMs and those with an even number of base pairs, while the hysteresis was minimal for longer poly-dC iMs and those with an odd number of base pairs. Experiments to follow the iM folding via thermal changes identified thermal hysteresis between the denaturing and annealing cycles. Similar trends were found to those observed in the CD experiments. The results demonstrate that the method of iM analysis can impact the pH_T parameter measured, and hysteresis was observed in the pH_T and T_m values.

KEYWORDS
circular dichroism, DNA folding, hysteresis, i-motif

1 | INTRODUCTION

Non-canonical DNA folds referred to as i-motifs (iMs) are pH-dependent structures found in sequences comprised of four proximal regions of 2′-deoxycytidine (dC) tracks of greater than three dC nucleotides per track within a short region of sequence space (Figure 1A).[1,2] The unique base-pairing element of an iM structure is the hemiprotonated dC base pair (dC•dC$^+$) that forms between parallel-oriented dC nucleotides (Figure 1B,C). Two dC•dC$^+$ base-paired strands intercalate to generate the iM structure with base pairs in the core and three loops (Figure 1C). As a consequence of the low pK_a (~4.3) for N3 of dC involved in base pairing, iMs typically fold under somewhat acidic conditions; however, iM folding at neutral pH has been observed when iM-specific ligands,[3–5] Ag^+ ions,[6] or crowding co-solvents are present.[7,8] In biologically relevant sequence contexts such as specific dC-homopolymer microsatellite sequences with length > 13 nt, other dC-rich genomic repeats, and the dC-rich human gene promoter sequences *RAD17*, *ATXN2L*, *RET*, and *DAP*, as examples, all have been found to adopt iMs at neutral pH conditions.[9–15] Over the last several years, studies have demonstrated iMs can fold in the cellular context,[4,16–18] which has led to a focus on the potential for biological function for these non-canonical folds. The first approach to address whether potential iM-forming sequences may fold in the cellular context is to first study them *in vitro* by biophysical methods.

We dedicate this work to Professor Eric Kool (Stanford Univ.) and congratulate him as the 2019 recipient of the Murray Goodman Memorial Prize, presented by *Biopolymers*.

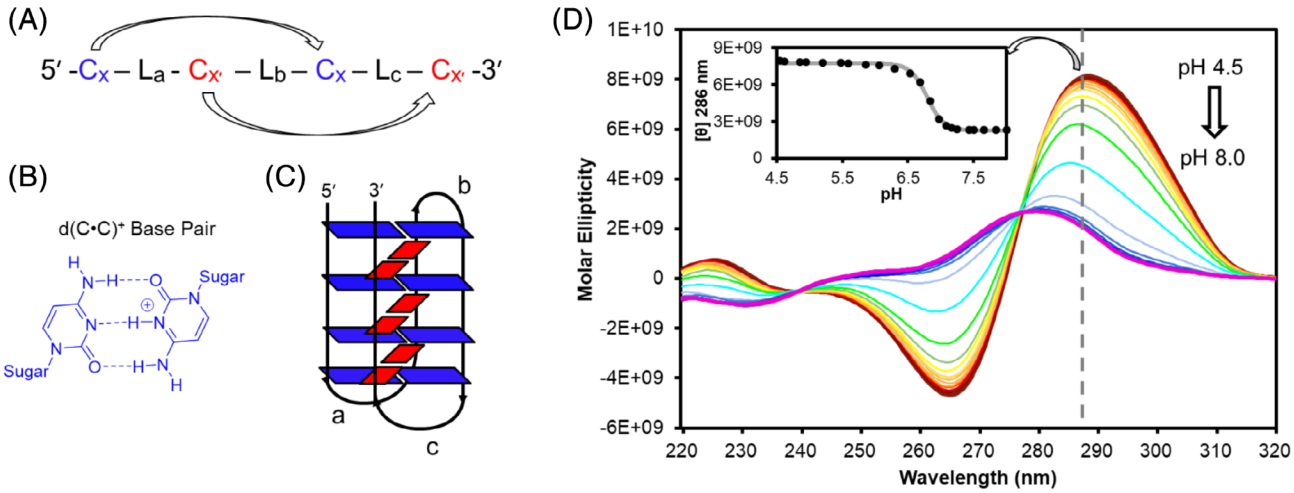


FIGURE 1 Sequences of DNA with four tracks of dC nucleotides >3 per track have the potential to adopt iM folds. A, The generalized consensus sequence for a possible iM-forming sequence. B, Structure of a hemi-protonated dC•dC$^+$ base pair. C, A cartoon representation of an iM fold. D, An example of a CD spectroscopic titration profile from pH 4.5 to 8.0 to determine the transition pH (pH_T). The inset to panel D is a plot of the molar ellipticity ([θ]) at 286 nm vs pH that was fit to determine the pH_T value. See the experimental section for a description of how the data were analyzed

A key quantitative measure for iM folding is the transition pH (pH_T) that describes the pH value of the inflection point in the transition from the folded to the unfolded state during a pH titration.[1,2] Experiments to determine this iM-specific value are typically conducted by monitoring the structural state via one of two different titration experiments.[1,2,4,16,17,19] Strands of DNA that can only adopt iM folds, and not any other completing secondary structures, are single stranded at pH 8 at which hemiprotonated dC•dC$^+$ base-pairs cannot form.[9,20–25] This unique feature of iM strands provides the opportunity to study their folding isothermally by dropping the pH from 8 to a lower value where dC•dC$^+$ base pairs can form; this folding approach differs from the traditional method for folding nucleic acids via heating to ~90 °C and slow cooling to obtain thermodynamically favorable folds. In the present study, isothermal refolding of the poly-dC iM strands was studied by using pH to drive the folding process to observe the impact on their physical parameters.

We used two different methods to study iM folding in the present work. *Method 1*: The unfolded DNA strands are directly annealed in a series of individual buffered solutions with pH values ranging between 4 and 8, and then after a set time to allow folding equilibrium to be reached (~30-60 minutes), the circular dichroism (CD) spectra were recorded. Folded iMs have a λ_{max} at ~286 nm that decays upon unfolding to produce spectra with λ_{max} values at ~280 nm (Figure 1D). The intensity of the molar ellipticity ([θ]) CD signal at 286 nm is plotted vs pH to yield a titration curve. When present, the inflection point of the titration curve obtained (i.e., pH_T) is used to determine whether the sequence can adopt an iM and provides a measure of its pH stability (Figure 1D inset). *Method 2*: The DNA strand can be folded at low pH (~4) and titrated with the addition of hydroxide up to pH 8 in a CD cuvette while monitoring changes in the spectrum at each point during the titration. This yields data similar to method 1 to derive pH_T values for a sequence of interest. This second approach can also be used to monitor the annealing process by titrating a pH 8 sample with HCl down to pH 4 while following the spectroscopic changes to determine the pH_T value. Method 2 enables one to follow a full cycle of the iM denaturing and annealing process that is not possible with method 1. In method 2, the pH_T values measured for the annealing and denaturing processes may differ displaying *isothermal* hysteresis.

In the second type of experiment to study iMs, the thermal stability of the fold is measured at specific pH values,[1,2,4,16,17,19] and this can be evaluated by following the thermally induced denaturing and annealing processes by CD or UV-vis spectroscopy at 286 or 295 nm, respectively, to follow the transition between folded and unfolded states. Monitoring the temperature dependency in folding yields a transition midpoint referred to as the thermal melting temperature (T_m), and T_m measurements can also reveal *thermal* hysteresis. A few studies on iM folds have identified and begun to understand folding hysteresis for specific sequences.[12,26] Studies to compare different analytical methods to yield iM-specific parameters and find experimental dependency have not been reported; inspection of iM hysteresis is, therefore, ripe for further study.

The poly-dC iMs found to fold in prior work from our laboratory provide a good test case to compare how the sample analysis impacts the isothermal hysteresis found in the pH_T values, as well as the thermal hysteresis in the T_m values.[10,14] Further, these homopolymer strands can be mutated with thymidine (T) nucleotides to favor preferred loop and stem lengths for determination of the pH_T values and isothermal hysteresis, as well as to measure the T_m values and thermal hysteresis.[10,14] The poly-dC iMs produced a chain-length pattern of stability that followed a $4n - 1$ repeat pattern from length = 12 to 30 nucleotides.[10,14] On the basis of our prior work, dC_{19} (n = 5) had the greatest pH_T value and pH 7 T_m value, while dC_{15} (n = 4) and dC_{23}

(n = 6) were other islands of high stability.[11] Deeper interrogation of the poly-dC iMs was conducted by locking in specific loop lengths via judicious mutation of dC to T nucleotides that do not participate in iM base pairing. These studies identified the $4n - 1$ repeat pattern results from the preferred stem and loop lengths.[10] In the present work, the poly-dC strands of length 15, 19, or 23 nt were interrogated by the two different methods described to measure the pH_T values that allowed demonstration of the dependency of the values obtained on the experimental approach used; additionally, thermal hysteresis in the T_m values was observed for the iMs studied. Additional studies on a series of iMs with loop and stem lengths selected via mutation of dC to T were explored to identify hysteresis in the pH_T and T_m values. These studies provide additional insight to the iM-sequence parameters that give rise to hysteresis. Lastly, this report illustrates the sensitivity of iM biophysical parameters to experimental design that is instructive to researchers.

2 | MATERIALS AND METHODS

2.1 | Oligodeoxynucleotide synthesis and purification

All oligodeoxynucleotides were synthesized by the University of Utah DNA synthesis core facility. Each sequence was treated with 0.2 M piperidine at 90 °C for 20 minutes before purification to cleave any damaged strands generated during synthesis. The piperidine-treated strands were purified by anion-exchange HPLC using a method running A = 9:1 ddH_2O and MeCN and B = 1.5 M NaOAc pH 7 in 9:1 ddH_2O and MeCN in which the method was initiated with 15% B followed by a linear increase to 100% B over 30 minutes with a flow rate of 1 mL min^{-1} while monitoring the absorbance at 260 nm. The purified samples were subsequently dialyzed for 36 hours against ddH_2O at 4 °C after which the samples were lyophilized to dryness and resuspended in pH 8 buffer to furnish unfolded stock solutions of purified oligodeoxynucleotides that were frozen at −20 °C until needed. The concentrations of each stock solution (∼100 μM) were determined by measuring the UV-vis molar extinction coefficients (ε; L μmol^{-1} cm^{-1}) that were estimated using the nearest-neighbor approach based on the primary sequence of the oligodeoxynucleotide (Table 1).

2.2 | CD spectroscopy, pH titrations, and data analysis

Monitoring the folding process via CD analysis followed two different methods. In method 1 to evaluate pH_T values, pH-dependent CD studies were conducted from pH 4.00 to 8.00 by taking measurements at intervals of 0.25 pH units in discrete samples. The various pH values were established in 20 mM Britton-Robinson buffer that is comprised of 20 mM each of sodium salts of phosphate, acetate, and borate with 140 mM KCl added. The Britton-Robinson buffer has a broad range of capacity from pH 2 to 12 that is ideal for pH titrations of iM sequences from pH 4 to 8. The oligomers were diluted from ∼100-μM stock solutions at pH 8 down to 2 μM at each pH condition interrogated. The unfolded pH 8 iMs were folded when diluting to the lower concentration in lower pH buffer solutions, after which isothermal folding proceeded for 30 minutes at 20 °C before CD analysis. The CD spectra were recorded from 220 to 320 nm with a scan-interval of 0.1 nm. The differential absorbance units measured were converted to molar ellipticity values ([Θ]) and then plotted. Next, the pH-dependent spectra were stacked for presentation, followed by making a secondary plot of $[\Theta]_{286\ nm}$ vs pH yielding a titration curve that was fit with the Henderson-Hasselbalch equation to determine the transition midpoint (pH_T). The titrations were conducted in triplicate to obtain experimental error bars.

In method 2, the oligomer samples were prepared by direct addition of the unfolded sequence to a buffer of 20 mM KP_i (pH 4.0), 12 mM NaCl, and 120 mM KCl, equilibrated at 20 °C. Sample concentrations were selected by calculating the amount needed to give an absorbance of ∼0.5 at 260 nm in a 1-cm cuvette (∼2 μM). The CD spectra were recorded at pH increments of ∼0.1 to 0.2 using small aliquots of 1 M LiOH to adjust the pH at each reading. Each sample was allowed to equilibrate for ∼90 seconds before recording the spectrum for this pH value. The 90 seconds delay time was experimentally

TABLE 1 Oligodeoxynucleotide Sequences Studied

Name	Sequence	ε_{260} (L μmol^{-1} cm^{-1})
dC_{15}	5′-CCC CCC CCC CCC CCC	0.1058
dC_{19}	5′-CCC CCC CCC CCC CCC CCC C	0.1340
dC_{23}	5′-CCC CCC CCC CCC CCC CCC CCC CC	0.1622
iM−$4_1 3_1 4_1 3$	5′-T_5 CCCC T CCC T CCCC T CCC	0.1665
iM−$4_1 3_3 4_1 3$	5′-T_5 CCCC T CCC TTT CCCC T CCC	0.1827
iM−$4_1 3_5 4_1 3$	5′-T_5 CCCC T CCC TTTTT CCCC T CCC	0.1989
iM−$4_1 4_1 4_1 4$	5′-T_5 CCCC T CCCC T CCCC T CCCC	0.1809
iM−$4_1 4_3 4_1 4$	5′-T_5 CCCC T CCCC TTT CCCC T CCCC	0.1971
iM−$4_1 4_5 4_1 4$	5′-T_5 CCCC T CCCC TTTTT CCCC T CCCC	0.2133

validated by monitoring the time-dependent change in the CD spectra, verifying no changes occur after this equilibration time. Upon reaching pH 8.0, the titration was then reversed using 1 M HCl via the same procedure as described to reach a pH of ~4.0. We refer to each pH-directional titration as a "leg." Throughout the complete titration cycle, no more than 90 μL total of 1 M LiOH and 1 M HCl was added to the 2000-μL sample resulting in a volume change of less than 5% in all cases. All titrations were performed in triplicate. The approach for analysis of the data was previously reported by our laboratory.[12]

2.3 | Thermal melting (T_m) studies

The iM strands were prepared at 3 μM concentration in Britton-Robinson buffer with 140 mM KCl at pH 5.0 or 6.0. The T_m experiments were conducted by thermally equilibrating the strands at 5 °C for 20 minutes followed by heating the samples to 100 °C at a rate of 1, 0.5, or 0.1 °C min^{-1} while thermally equilibrating at each step for 60 seconds before measuring the UV absorbance. Absorbance readings were taken at 260 and 295 nm. The T_m values were determined by plotting the 260 or 295 nm reading as a function of temperature to obtain the thermal denaturation profile; the data at 295 nm were used to make the plots in this report. The transition point was determined using a two-point average method implemented with the instrument's software (Shimadzu UV-1800). All measurements were made in triplicate to obtain the experimental error bars.

3 | RESULTS AND DISCUSSION

3.1 | The method for analysis of pH_T impacts the value measured for poly-dC iMs

First, the poly-dC strands of length 15, 19, or 23 nucleotides were synthesized and HPLC purified for analysis. Before commencing the experiments, the dC_{19} strand was analyzed by ^{1}H-NMR at pH 6 or 8 (100 μM) to verify the pH dependency of the folding and unfolding of this strand. The pH 6 spectrum gave imino peaks around 15 to 16 ppm indicative of iM folding[9,20–25] while these signals disappeared in the pH 8 spectrum supporting the unfolded state at the higher pH (Figure S1). This observation underscores an interesting point about poly-dC strands. This polyanion does not self-associate at pH 8, that of our stock solutions and NMR studies. There is no mechanism for base pairing at pH 8, and base stacking is poor for pyrimidines. Therefore, we assume based on the dC_{19} result that all sequences studied herein were unfolded at pH 8 in the stock solutions before refolding them when they were diluted into lower pH buffers.

The strands were interrogated to determine their pH_T values using the two methods outlined in the introduction (Figure 1D and Figure S2). In method 1, the unfolded strand was added directly to different buffers with pH values ranging from 4.5 to 8.0 in 0.25 pH increments and then allowed to anneal for 30 minutes at 20 °C before CD analysis. In method 2, the unfolded strand was added directly to a pH 4.0 buffer and allowed to anneal for 30 minutes in a CD cuvette at 20 °C before monitoring the unfolding process. The iM unfolding process was driven by titration with LiOH into the sample to ultimately reach pH 8.0. The pH value obtained after each addition of LiOH was ~0.2 unit increase at each step that was determined using a pH meter. The samples were allowed to equilibrate at each pH step, which was determined when the CD spectrum showed no change (~90 seconds), and then the final spectrum was recorded and used in the analysis to determine the pH_T value (Figure S2). Method 2 enabled the analysis of the annealing process by then titrating the sample with HCl back to pH 4.0 following the same approach as just described. Method 2 can yield isothermal hysteresis between the two titration curves when they follow different paths from the folded to unfolded state as determined by the CD spectra.

For dC_{15}, method 1 for pH_T measurement produced an average value of 6.9 (Figure 2A, black). Isothermal hysteresis for dC_{15} was observed using method 2 for the determination of pH_T values upon monitoring denaturing vs annealing. The average denaturing pH_T value was 7.2 and the annealing pH_T value was 6.5 to give a 0.7 pH-unit hysteresis (Figure 2A, blue vs red). Overall a similar trend in the pH_T values was found for dC_{19}, in which method 1 gave a pH_T value of 7.3 (Figure 2B, black), and method 2 produced hysteresis in the values. Moreover, the dC_{19} sequence gave a biphasic transition during the denaturing cycle with pH_T values of 6.8 and 7.5 and a single transition upon annealing with a lower pH_T value of 6.7 (Figure 2B, blue vs red). The biphasic nature of the transition for dC_{19} in the denaturing cycle of the method 2 approach has been reported previously by our laboratory.[12] In the dC_{23} strand investigated, method 1 produced a pH_T value of 7.1, and both denaturing and annealing legs of method 2 also produced pH_T values of 7.1; thus, no hysteresis was observed with the longest dC_{23} iM studied (Figure 2C).

These initial studies indicate that the poly-dC iM chain length influences whether isothermal hysteresis was observed by method 2 for pH_T value determination. The shorter poly-dC iMs with lengths 15 or 19 nucleotides displayed isothermal hysteresis; in contrast, the 23-nucleotide iM did not. This study suggests iM length is one parameter that determines whether hysteresis in the pH_T values occurs. Consistent with this finding is prior work from our laboratory and others that found other iMs produced isothermal hysteresis in the pH_T values.[12,19,26]

The more interesting finding in the present studies is that the method (1 vs 2) for titrating iM-forming sequences to measure pH_T values can yield different results. The difference was maximally displayed in the dC_{15} and dC_{19} sequences (Figure 2A,B). Method 1 is commonly used for pH_T value measurement, and it gives a value similar to the denaturing leg of method 2; however, the shapes of the method 1 vs 2 denaturing curves are different. The method 1 curves show steep transitions for the two shorter sequences indicative of a highly cooperative process,[27] while the method 2 denaturing curves were comparatively broader. One explanation for this difference is that the method 2 samples were not at equilibrium before analysis,

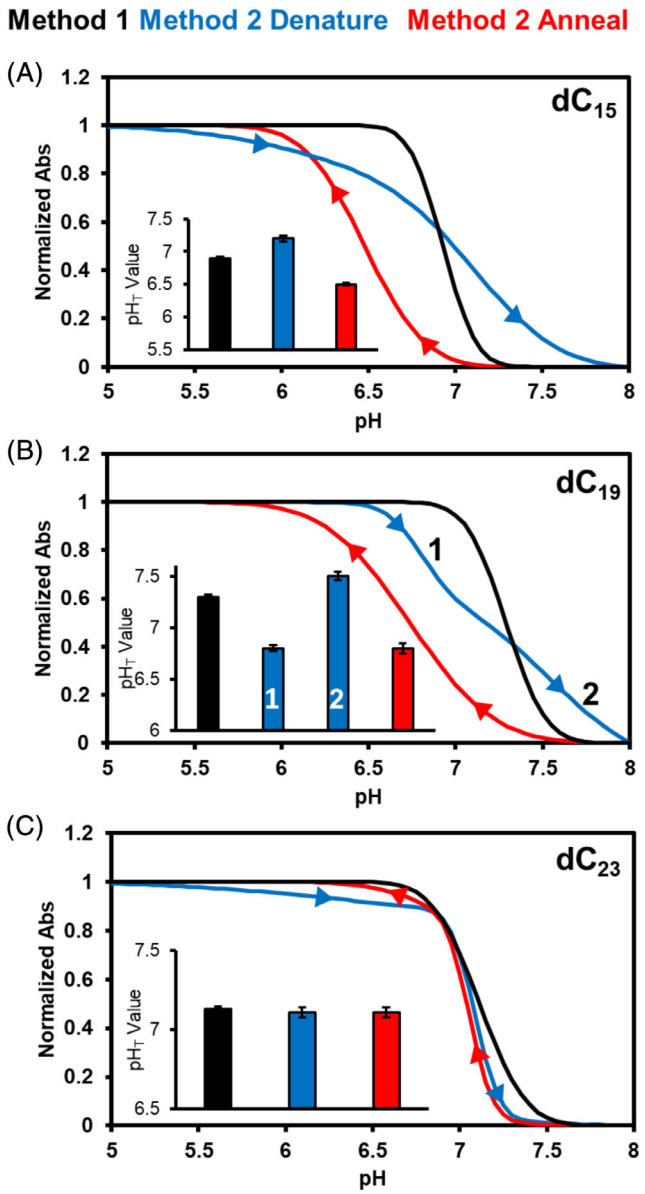


FIGURE 2 Analysis of the poly-dC iM-forming sequences with lengths of 15, 19, or 23 nt by two different CD spectroscopic methods identify dependency of the pH_T value reported on the method used. Analysis of the, A, dC_{15}, B, dC_{19}, and C, dC_{23} iMs by method 1 (black curves) or method 2 that allowed following the denaturing (blue) and annealing (red) processes for the three strands studied. The inset for each panel is a bar chart of the pH_T values found from the different methods. For dC_{19} (B), the denaturing process was biphasic yielding two pH_T values labeled 1 and 2

although the spectrum at each pH increment was monitored until no change was observed before taking the final spectrum used in the analysis, suggesting all measurements were made at folding equilibrium. This point is discussed in more detail below. Method 2 when used to follow annealing found the shorter poly-dC iMs (i.e., dC_{15} and dC_{19}) had lower pH_T values that could be as much as 0.7-pH units less than the denaturing value, resulting in isothermal hysteresis. Prior iM analysis from our laboratory found when the iM-forming sequence first prepared by heat denaturing and slow cooling at pH 4 results in loss of the hysteresis.[12] Lastly, the longest dC_{23} iM gave nearly identical titration curves for both method 1 and the denaturing and annealing legs of method 2 (Figure 2C). This final observation indicates that not all poly-dC iMs display hysteresis and the two methods for pH_T evaluation can give the same values. In summary, the method of analysis to measure the pH_T value for an iM-forming sequence can impact the results obtained.

In the annealing leg of the method 2 analysis, the curves for the shorter homopolymer iMs were different than the denaturing leg of this method giving rise to the hysteresis observed; further, this method of following the pH dependency of folding produced a different result than method 1. A key difference in the annealing leg of method 2 compared to the other ways of assessing pH_T values (i.e., method 1 and method 2 denaturing leg) is the rate of the initial annealing. In method 1, the unfolded iM oligomers are directly added to buffers of known pH and fold rapidly. This is also true of the method 2 denaturing leg, in which the strand was directly added to a pH 4.0 buffer before the commencement of the analysis. In contrast, monitoring annealing by method 2 was achieved via a slower folding process that occurred when the pH was slowly adjusted downward. The experiment took ~3 hours to perform. Prior work by Skolakova *et al*. found folding iMs slowly produced bimolecular structures in preference to intramolecular folds for long iMs and those with short loops[19]; this provides a possible explanation for the differences observed between the two different approaches to monitoring annealing. In our work, hysteresis was only observed for dC_{15} or dC_{19} and not the longer dC_{23} iM, the present trend observed here is opposite of the chain-length trend reported.[19] We note that there are other differences between the studies, such as analysis temperature and sample preparation that could easily lead to different findings.[28] Thus, any further comparison of these data sets is not possible or warranted, especially in light of the present observation that the method for pH_T value determination impacts the results obtained. Additional studies to better understand the molecularity of folding of the poly-dC strands under the conditions of the present analysis conditions were not pursued; however, we did previously report that dC_{15} and dC_{19} do not adopt a bimolecular structure when directly folded at pH 6.[14] The most important finding herein is that the method of analysis for pH_T value assessment can impact the value found. To reiterate, the present findings are important when researchers try to compare pH_T values for a given sequence studied by different laboratories.[29,30]

3.2 | Impact of dC•dC$^+$ base pair count and central loop length on isothermal hysteresis

Previously our laboratory studied iM strands with judiciously mutated dC to T nucleotides for evaluation of preferred core base pair counts and loop lengths to address the $4n - 1$ repeat pattern observed in the poly-dC iMs.[10,14] A model addressing dC_{19} as the most stable iM was proposed because it could adopt two different structures with

similar high pH and thermal stabilities. The first structure had an even number of base pairs in the core—four base pairs per intercalated strand—and single nucleotides in the three loops (Figure 1A,C). The second structure of nearly equal stability had an odd number of base pairs in the core—four in one strand and three in the other strand—with loop lengths of one in each exterior loop and three in the central loop (Figure 1A,C). For iMs with an even number of core base pairs, two different conformations can exist, referred to as 3′E when the 3′-most dC is in an exterior base pair or 5′E when the 5′-most dC is in an exterior base pair (Figure 3).[23] This information was used to study how the experimental approach in pH_T value determination (method 1 vs 2) produced different values for iMs with different base pair counts and central loop lengths with a focus on the strands from the prior study.[10] The nomenclature used for each sequence variant has the number of dC nucleotides in each of the four tracks interspersed with the number of T nucleotides in each of the three loops subscripted, with the numbers describing the sequence from the 5′ to 3′ ends. For example, iM-$4_14_14_14$ has four dC tracks of length 4 nucleotides each with 1 T nucleotide in each of the three loops.

For the iMs with an even number of base pairs in the core, the following method-dependent and loop length-dependent results were obtained. Two noteworthy points regarding these data must be made: (1) the method 1 values for the core-length and loop-length studies were previously reported by our laboratory using strands that did not contain tails (Table 1),[10] and (2) the sequences used in method 2 were comprised of the same iM-forming sequence but had a 5′-poly-T tail of five nucleotides (Table 1). Using method 2 to evaluate pH_T values, two dC_{19} strands with or without the 5′-T_5 tail were analyzed to find that the T_5 tail decreased the transition values by ~0.3 pH units (Figure S3), thus the values are not directly comparable. Nonetheless, these data do allow further demonstration of the influence of method choice in pH_T value measurement and show that loop and core lengths impact isothermal hysteresis observed in method 2. For the iM-$4_14_14_14$ strand, method 2 produced a biphasic denaturing curve and a monophasic annealing curve similar to dC_{19} (Figure 2B), with denaturing pH_T values of 6.5 and 7.3, and an annealing pH_T value of 6.5. The method 1 pH_T value for the iM-$4_14_14_14$ iM was 7.1 that is similar to the highest transition seen in the biphasic denaturing curve of method 2 analysis. When the core base pair count remained the same and the central loop was increased to 3 or 5 nucleotides, that is, iM-$4_14_34_14$ and iM-$4_14_54_14$, respectively, the method 2 denaturing curves were also biphasic and the annealing curves were monophasic at a lower value (Figure 3A), thus showing isothermal hysteresis. The method 1 values for these strands, in general, were similar to the lower pH_T values found in the denaturing curves that were also similar to the annealing pH_T values found with method 2. The key finding was that iMs with eight-core base pairs always showed hysteresis by method 2 analysis, regardless of the central loop length when studied with 1, 3, or 5 T nucleotides.

In contrast, when the iM had an odd number of core base pairs hysteresis was minimal or not observed by method 2. For example, the iM-$4_13_14_13$ sequence gave a ~0.1 pH unit difference in the pH_T values between denaturing and annealing when measured via method 2 (6.4 vs 6.3; Figure 3B); furthermore, the denaturing leg of the method 2 analysis for these strands was monophasic. The pH_T value observed for the same sequence using method 1 was slightly higher at 6.7, as expected because a tail was not present (Figure 3B). Next, the iM-$4_13_34_13$ and iM-$4_13_54_13$ sequences both produced nearly the

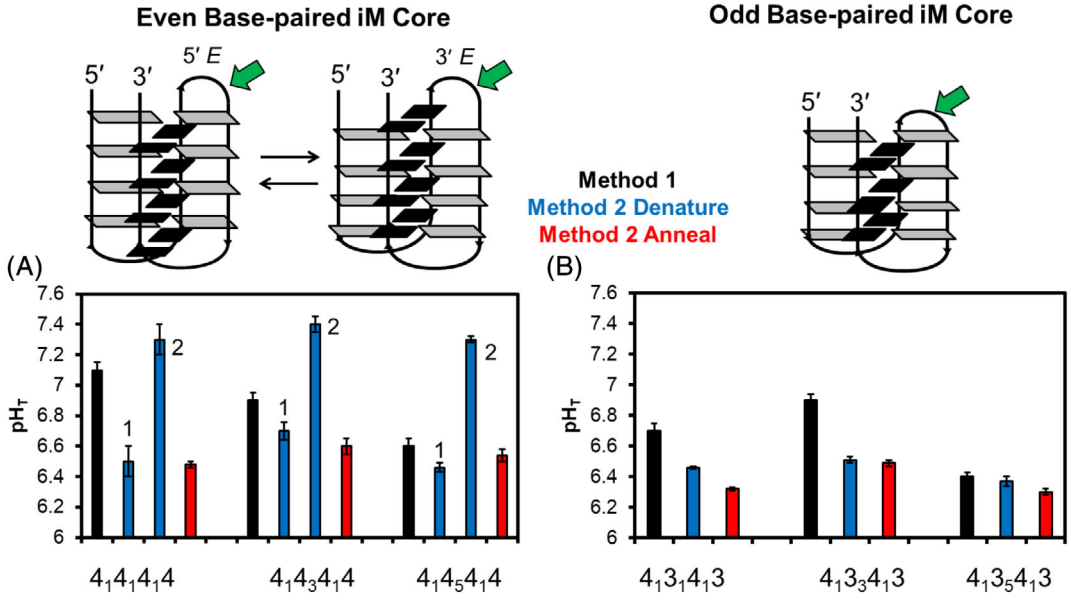


FIGURE 3 Analysis of iMs mutated to prefer specific dC•dC^+ base pair counts and loop lengths to evaluate the method dependency in pH_T value determination. Plots of pH_T values for iMs with an even number (i.e., 8) of base pairs in the core (A), and an odd number of base pairs in the core (B) (i.e., 7). Both systems also maintained exterior loops of 1 T nucleotide each and increased the central loop length from 1, 3, or 5 T nucleotides. Biphasic curves for method 2 denaturation are labeled 1 and 2 for the two pH_T values

same pH_T values by the denaturing and annealing legs of method 2 analysis and did not show hysteresis. From these studies, we conclude that the iMs with seven base pairs in the core that possess single-nucleotide exterior loops and a central loop comprised of 1, 3, or 5 T nucleotides did not display isothermal hysteresis when monitoring denaturation and annealing. These data in their entirety suggest isothermal hysteresis is maximal for iMs that contain an even number of base pairs in the core, and the central loop length with T nucleotides has minimal impact, while iMs with an odd number of core base pairs show minimal or non-existent isothermal hysteresis with no additional impact of the central loop length.

3.3 | Studies of iM thermal hysteresis

In the final set of studies to inspect for hysteresis, T_m curves were monitored during the denaturing and annealing phases to see the difference in the inflection points (i.e., T_m values) between the two curves (Figure 4A: blue = denaturing, red = annealing), indicating thermal hysteresis. To ensure all measurements were made at thermal equilibrium, the ramp rate when studying dC_{19} was decreased from 1, 0.5, or 0.1 °C min^{-1} to find the 0.5 and 0.1 °C min^{-1} ramp rates gave similar curves. This suggests that the curves analyzed at the 0.5 °C min^{-1} ramp rate were at thermal equilibrium, and the hysteresis observed results from the iM structure and not the experimental setup. Lastly, the T_m curves were analyzed via a two-point analysis using the instrument's software that defined the folded and unfolded states by a leveling off of the change in intensity of the curve to identify the midpoint in the transition, defined as the T_m value.

The first iMs inspected for thermal hysteresis were the poly-dC strands of length 15, 19, and 23 nucleotides that were analyzed at either pH 5 or 6 (Figure 4B). Overall, the thermal hysteresis for these strands was greatest at pH 6 (>15 °C) and minimal at pH 5 (~5 °C). Further, at both pH values studied, an inverse relationship with chain length and thermal hysteresis was observed (i.e., longer chains gave less hysteresis in the T_m curves). Interestingly, the trend in chain-length dependency in thermal hysteresis for the poly-dC iMs was similar to that obtained in the isothermal CD studies (Figures 2A-C and 4B). These trends suggest that the greater flexibility for longer iM strands to fold may provide a similar folding and unfolding pathway that is not observed for the shorter strands. This hypothesis needs to be verified by future modeling and experimental work for which techniques exist to address this question.[31–36]

Next, the model iMs with eight $dC•dC^+$ base pairs in the core with single T nucleotides in the exterior loops and the interior loop systematically increased from 1 to 3 to 5 T nucleotides (i.e., $4_14_14_14$, $4_14_34_14$, or $4_14_54_14$) were interrogated for thermal hysteresis. Trends were again observed in the thermal hysteresis measured. In all three sequences studied, the hysteresis was greater by ~5 °C when measured at pH 6 than at pH 5 (Figure 4C green vs purple). Furthermore, as the central loop length increased in length, the thermal hysteresis decreased at both pH 5 and 6. For example, at pH 6 with all single-

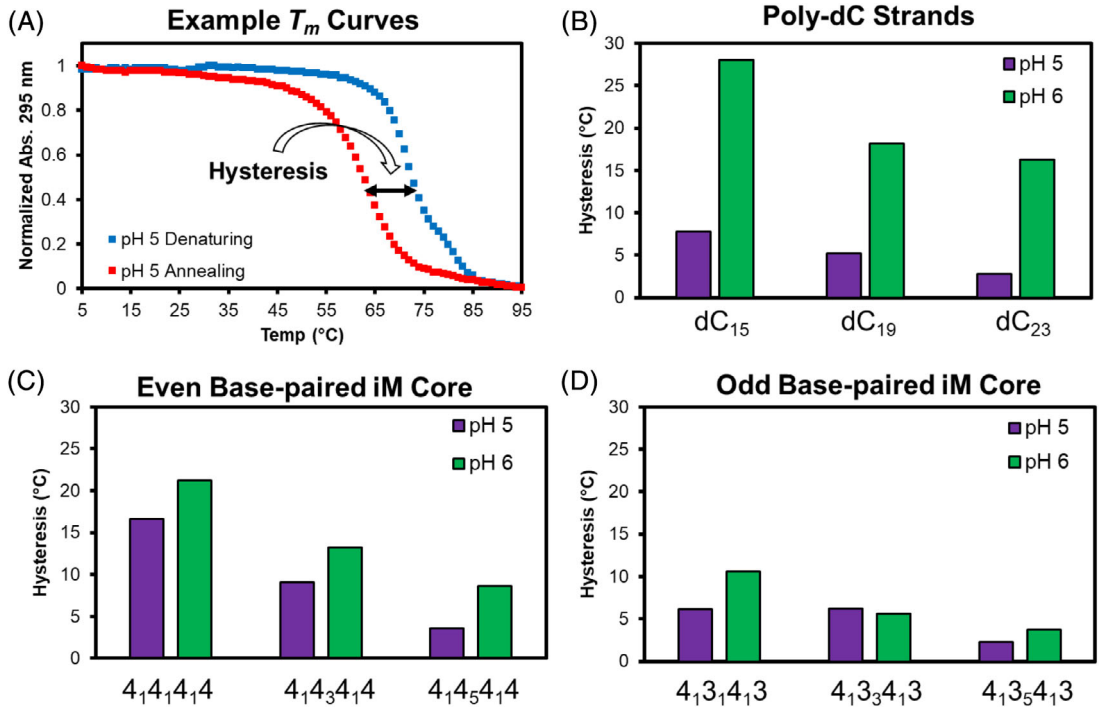


FIGURE 4 Analysis of the thermal hysteresis for the iM-forming sequences. A, Examples of denaturing (blue) and annealing (red) T_m curves that display thermal hysteresis. Hysteresis values found for the, B, poly-dC strands, C, even number of base pairs in the iM core, and D, odd number of base pairs in the iM core. In panels B, C, and D the hysteresis values found at pH 5 are the purple bars and those at pH 6 are the green bars

nucleotide loops, the hysteresis was 21.2 °C which decreased to 8.6 °C when the central loop was increased to five T nucleotides. From these data, both isothermal and thermal hysteresis studies (Figures 3 and 4), the iMs with eight dC•dC$^+$ base pairs in the core and loop lengths of 1 to 5 nucleotides always yield hysteresis between the denaturing and annealing profiles. This finding further supports iMs with an even number of core base pairs can give rise to hysteresis; however, whether this results from structural plasticity[23] is not known at present.

The final model iMs studied for thermal hysteresis had seven dC•dC$^+$ base pairs in the core with single nucleotide exterior loops comprised of T nucleotides and the central loop increased from 1, 3, or 5 T nucleotides. The denaturing and annealing T_m curves for these iMs displayed the least hysteresis at both pH values studied (<10 °C; Figure 4D). Additionally, as the central loop increased in length from 1 to 5 Ts, the hysteresis decreased by ~5 °C. In summary, the iMs with an odd number of base pairs in the core were found to have the least thermal hysteresis in the T_m studies and isothermal hysteresis when evaluating the pH_T values by method 2 (Figures 3B and 4D).

4 | CONCLUSION

The non-canonical DNA folds referred to as iMs are pH-dependent tetraplex structures comprised of hemiprotonated dC·dC$^+$ base pairs in the core of the structure with three loops (Figure 1A-C). Recent studies have found iMs can fold in the cellular context,[16,17] and there exist human sequences that fold at neutral pH, such as the poly-dC microsatellite sequences.[9,14,15,19] In the present study, we used CD spectroscopy to measure the folding transition pH value (i.e., pH_T value) via two different titration methods. First, we employed both titration methods to find that the pH_T values for the poly-dC strands of length 15, 19, or 23 nucleotides were dependent on the method of analysis (Figure 2). In method 2, the denaturing and annealing cycles were monitored to find strong isothermal hysteresis when monitoring the two different cycles for dC_{15} and dC_{19}, but hysteresis was not observed for dC_{23} (Figure 2). Next, a set of model oligomers that locked specific base pair counts and loop lengths for the iMs were studied to find that eight-core base pairs with single nucleotide exterior loops and a central loop of up to five nucleotides always displayed isothermal hysteresis (Figure 3A). In contrast, when the model iMs have seven core base pairs and the same combinations of loop lengths, isothermal hysteresis in the pH_T values was either minimal or not observed (Figure 3B). The present study in its entirety identifies two critical considerations when analyzing the iM folds.[1] The method to evaluate pH_T values impacts the value measured (Figure 2).[2] The number of central dC•dC$^+$ base pairs impacts the isothermal and thermal hysteresis observed (Figures 3 and 4). When the core has an even base pair count maximal hysteresis was measured, and in the odd base pair count the least hysteresis was measured. Future work is needed to address whether these trends hold for other iMs, particularly biologically relevant sequences. This information is noteworthy for laboratories that try to make comparisons between reported pH_T and T_m values on the same sequence that were prepared differently or studied by different methods. Our observations build off of a recent report that found sample handling of the human telomere iM can impact the results of biophysical studies.[28]

In the second set of studies, denaturing and annealing T_m curves were obtained on the poly-dC strands (15, 19, or 23 nt) to find *thermal* hysteresis (Figure 4B), showing values with the same trend in core base pair count and central loop length as seen in the *isothermal* studies (Figures 3A,B and 4C,D). The present work has identified some features of iMs that appear to cause increased hysteresis when conducting biophysical studies on these DNA strands; however, a few mysteries regarding this phenomenon associated with iMs remain. For instance, are different iM folds responsible for the hysteresis observed? Second, what are the structures that lead to biphasic pH_T curves with the poly-dC strands and those with an even number of core base pairs? Future structural studies are needed to address these fascinating questions. Nonetheless, researchers studying iM folding must be aware that the method of choice for analyzing the pH stability can impact the values found, and hysteresis may exist when monitoring the denaturing vs annealing process of iM folds.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation grant (CHE1808475). M.R.M. was supported by an NSF REU grant (CHE1659579). The oligonucleotides were prepared by the University of Utah Health Sciences Core facilities that are supported in part by the National Cancer Institute Cancer Center Support Grant P30 CA042014.

CONFLICT OF INTEREST

The authors declare no competing interests.

REFERENCES

[1] H. Abou Assi, M. Garavis, C. Gonzalez, M. J. Damha, *Nucleic Acids Res.* **2018**, *46*, 8038.
[2] H. A. Day, P. Pavlou, Z. A. Waller, *Bioorg. Med. Chem.* **2014**, *22*, 4407.
[3] S. Sedghi Masoud, K. Nagasawa, *Chem. Pharm. Bull.* **2018**, *66*, 1091.
[4] H. J. Kang, S. Kendrick, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4172.
[5] C. Cristofari, R. Rigo, M. L. Greco, M. Ghezzo, C. Sissi, *Sci. Rep.* **2019**, *9*, 1210.
[6] H. A. Day, C. Huguin, Z. A. Waller, *Chem. Commun.* **2013**, *49*, 7696.
[7] A. Rajendran, S.-i. Nakano, N. Sugimoto, *Chem. Commun.* **2010**, *46*, 1299.
[8] Y. P. Bhavsar-Jog, E. Van Dornshuld, T. A. Brooks, G. S. Tschumper, R. M. Wadkins, *Biochemistry* **2014**, *53*, 1586.
[9] E. P. Wright, J. L. Huppert, Z. A. Waller, *Nucleic Acids Res.* **2017**, *45*, 2951.
[10] A. M. Fleming, K. M. Stewart, G. M. Eyring, T. E. Ball, C. J. Burrows, *Org. Biomol. Chem.* **2018**, *16*, 4537.
[11] R. A. Rogers, A. M. Fleming, C. J. Burrows, *ACS Omega* **2018**, *3*, 9630.
[12] R. A. Rogers, A. M. Fleming, C. J. Burrows, *Biophys. J.* **2018**, *114*, 1804.
[13] J. Zhou, C. Wei, G. Jia, X. Wang, Z. Feng, C. Li, *Mol. BioSyst.* **2010**, *6*, 580.
[14] A. M. Fleming, Y. Ding, R. A. Rogers, J. Zhu, J. Zhu, A. D. Burton, C. B. Carlisle, C. J. Burrows, *J. Am. Chem. Soc.* **2017**, *139*, 4682.

[15] B. Mir, I. Serrano, D. Buitrago, M. Orozco, N. Escaja, C. Gonzalez, *J. Am. Chem. Soc.* **2017**, *139*, 13985.
[16] M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger, D. Christ, *Nat. Chem.* **2018**, *10*, 631.
[17] S. Dzatko, M. Krafcikova, R. Hänsel-Hertsch, T. Fessl, R. Fiala, T. Loja, D. Krafcik, J.-L. Mergny, S. Foldynova-Trantirkova, L. Trantirek, *Ang. Chem. Int. Ed. Engl.* **2018**, *57*, 2165.
[18] A. M. Fleming, J. Zhu, Y. Ding, C. J. Burrows, *ACS Chem. Biol.* **2017**, *12*, 2417.
[19] P. Školáková, D. Renčiuk, J. Palacký, D. Krafčík, Z. Dvořáková, I. Kejnovská, K. Bednářová, M. Vorlíčková, *Nucleic Acids Res.* **2019**, *47*, 2177.
[20] J. Zhou, S. Amrane, D. N. Korkut, A. Bourdoncle, H. Z. He, D. L. Ma, J. L. Mergny, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 7742.
[21] S. Kendrick, Y. Akiyama, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2009**, *131*, 17667.
[22] L. Lannes, S. Halder, Y. Krishnan, H. Schwalbe, *ChemBioChem* **2015**, *16*, 1647.
[23] A. L. Lieblein, J. Buck, K. Schlepckow, B. Furtig, H. Schwalbe, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 250.
[24] A. T. Phan, J. L. Mergny, *Nucleic Acids Res.* **2002**, *30*, 4618.
[25] J. Dai, E. Hatzakis, L. H. Hurley, D. Yang, *PLoS One* **2010**, *5*, e11647.
[26] J. L. Mergny, L. Lacroix, *Nucleic Acids Res.* **1998**, *26*, 4797.
[27] I. V. Nesterova, E. E. Nesterov, *J. Am. Chem. Soc.* **2014**, *136*, 8843.
[28] M. A. S. Abdelhamid, Z. A. E. Waller, *Front. Chem.* **2020**, *8*, 40.
[29] J. L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* **2005**, *33*, e138.
[30] C. Ma, R. C.-T. Chan, C. T.-L. Chan, A. K.-W. Wong, B. P.-Y. Chung, W.-M. Kwok, *Chem. Asian J.* **2018**, *13*, 3706.
[31] Y. Ding, A. M. Fleming, L. He, C. J. Burrows, *J. Am. Chem. Soc.* **2015**, *137*, 9053.
[32] S. Jonchhe, P. Shrestha, K. Ascencio, H. Mao, *Anal. Chem.* **2018**, *90*, 3205.
[33] A. L. Lieblein, B. Furtig, H. Schwalbe, *ChemBioChem* **2013**, *14*, 1226.
[34] H. A. Assi, R. W. Harkness, N. V; Martin-Pintado, C. J. Wilds, R. Campos-Olivas, A. K. Mittermaier, C. González, M. J. Damha, *Nucleic Acids Res.* **2016**, *44*, 4998.
[35] M. Endo, X. Xing, H. Sugiyama, *Methods Mol. Biol.* **2019**, *2035*, 299.
[36] T. Panczyk, P. Wolski, *Biophys. Chem.* **2018**, *237*, 22.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: RA Rogers, MR Meyer, KM Stewart, GM Eyring, AM Fleming, CJ Burrows. Hysteresis in poly-2′-deoxycytidine i-motif folding is impacted by the method of analysis as well as loop and stem lengths. *Biopolymers*. 2020; e23389. https://doi.org/10.1002/bip.23389

Supporting Information

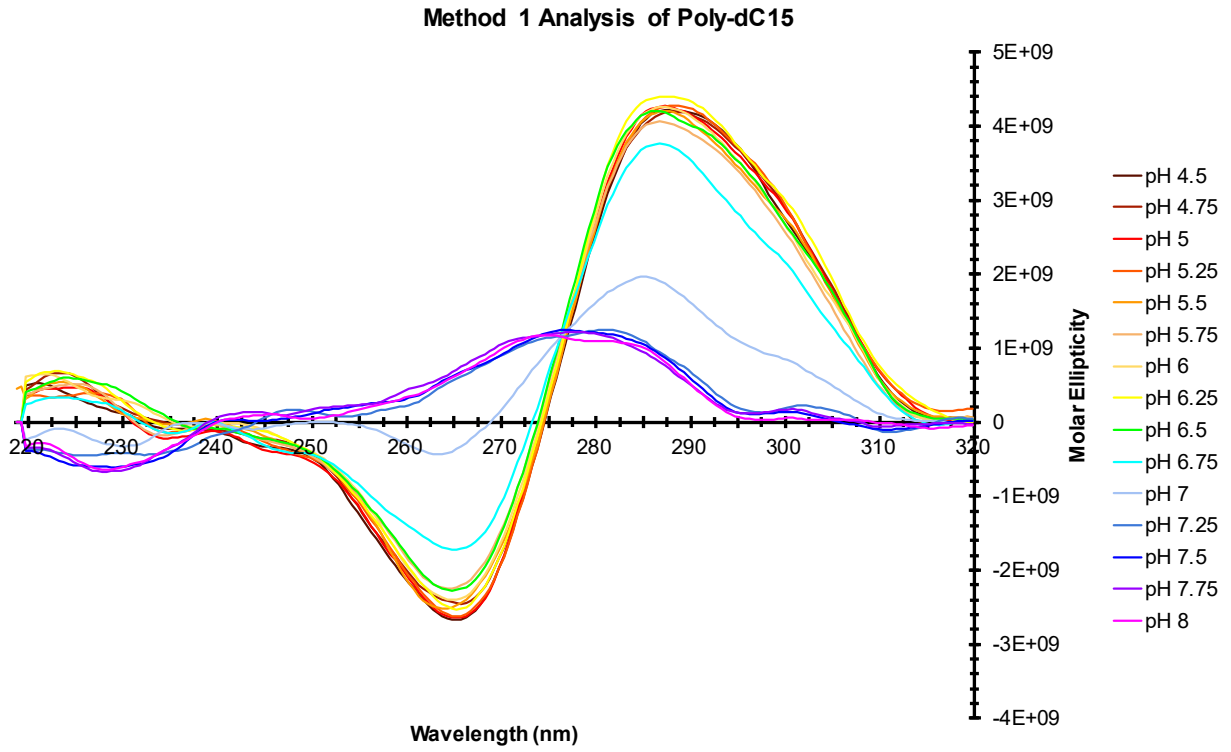
Poly-2`-deoxycytidine i-motif folding hysteresis is impacted by the method of analysis as well as loop and stem lengths

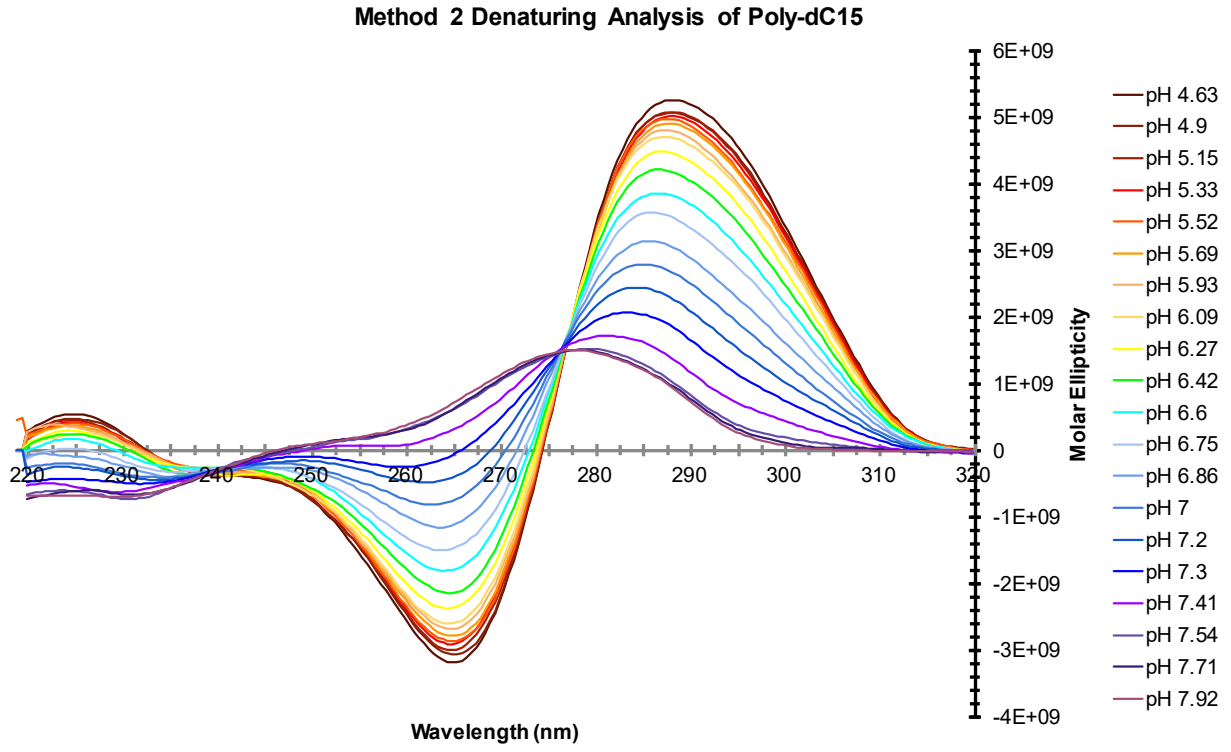
R. Aaron Rogers, Madeline R. Meyer, Kayla M. Stewart, Gabriela M. Eyring, Aaron M. Fleming, and Cynthia J. Burrows*

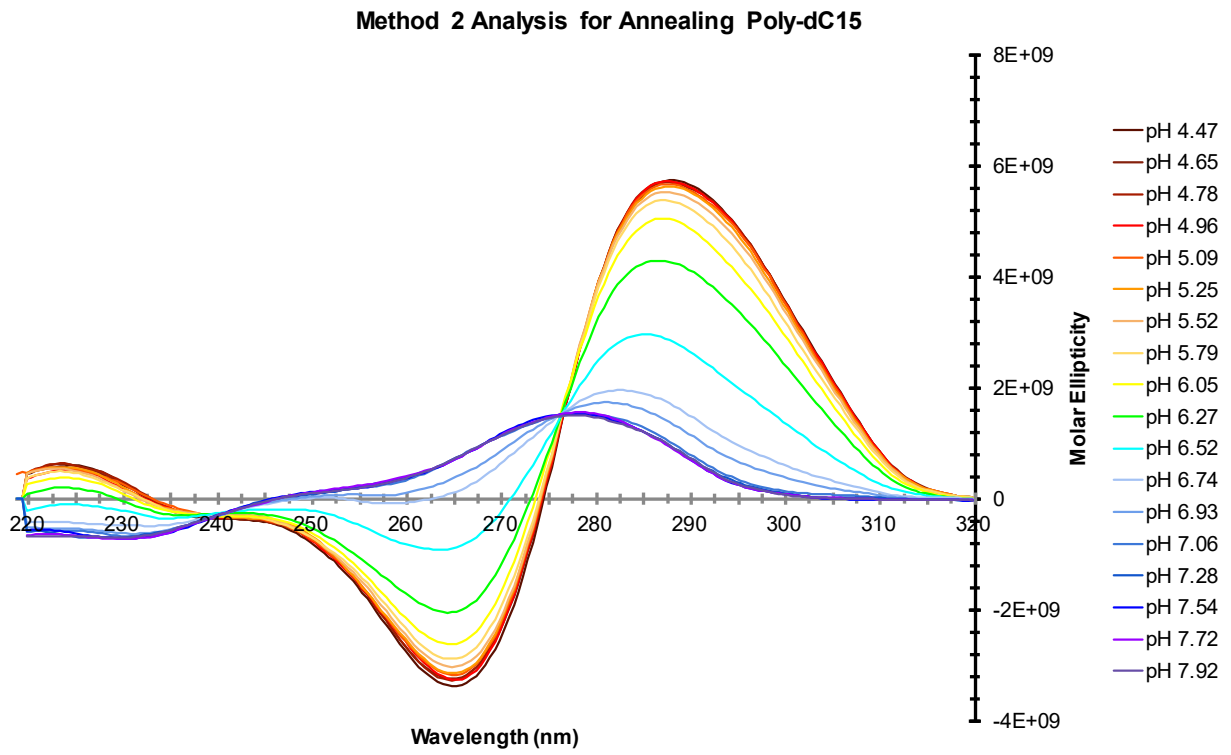
Department of Chemistry, University of Utah, 315 S. 1400 East, Salt Lake City, Utah 84112-0850, United States

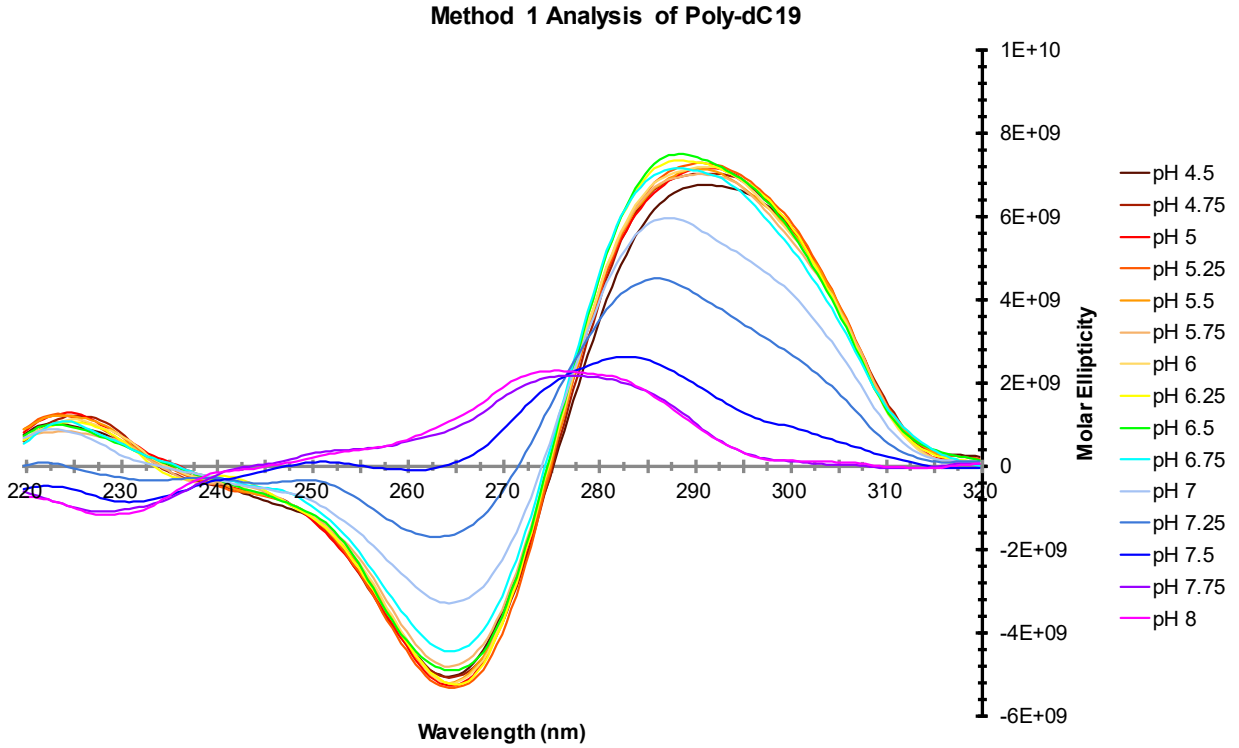
*To whom correspondence should be addressed. E-mail: burrows@chem.utah.edu

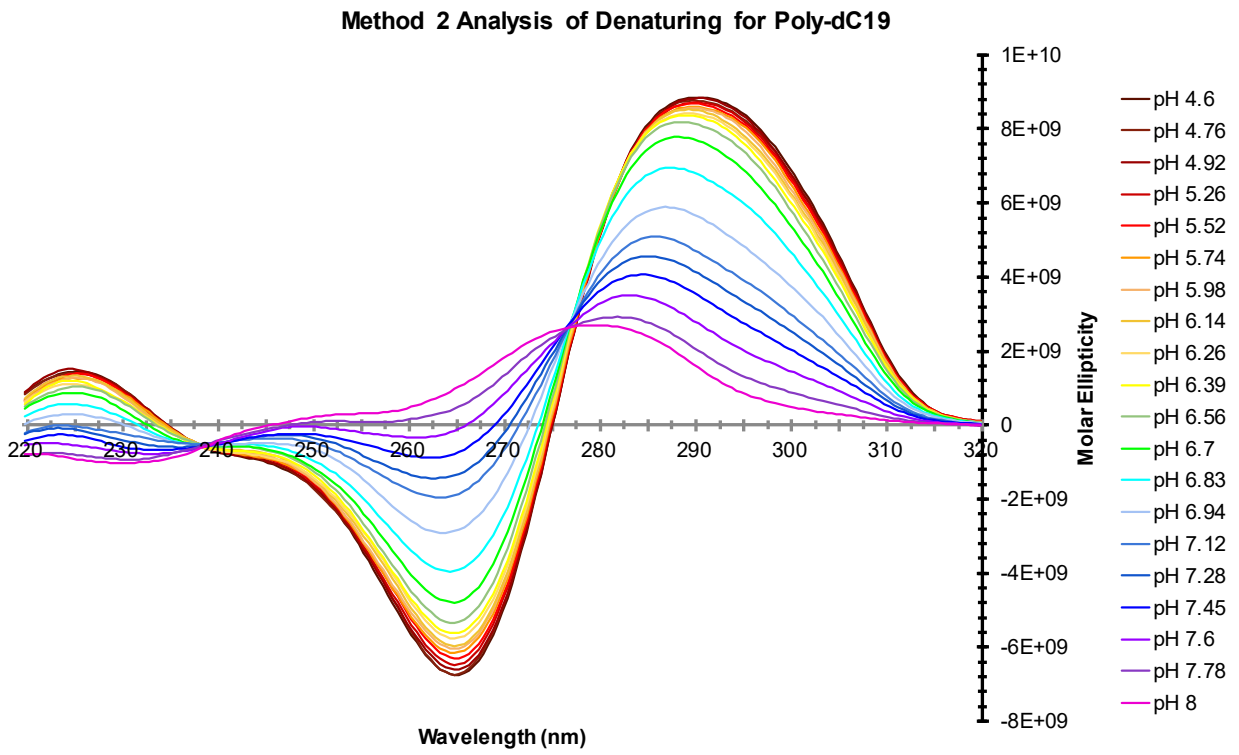
Figure S1. Method 1 and 2 analysis of the poly-dC iMs.

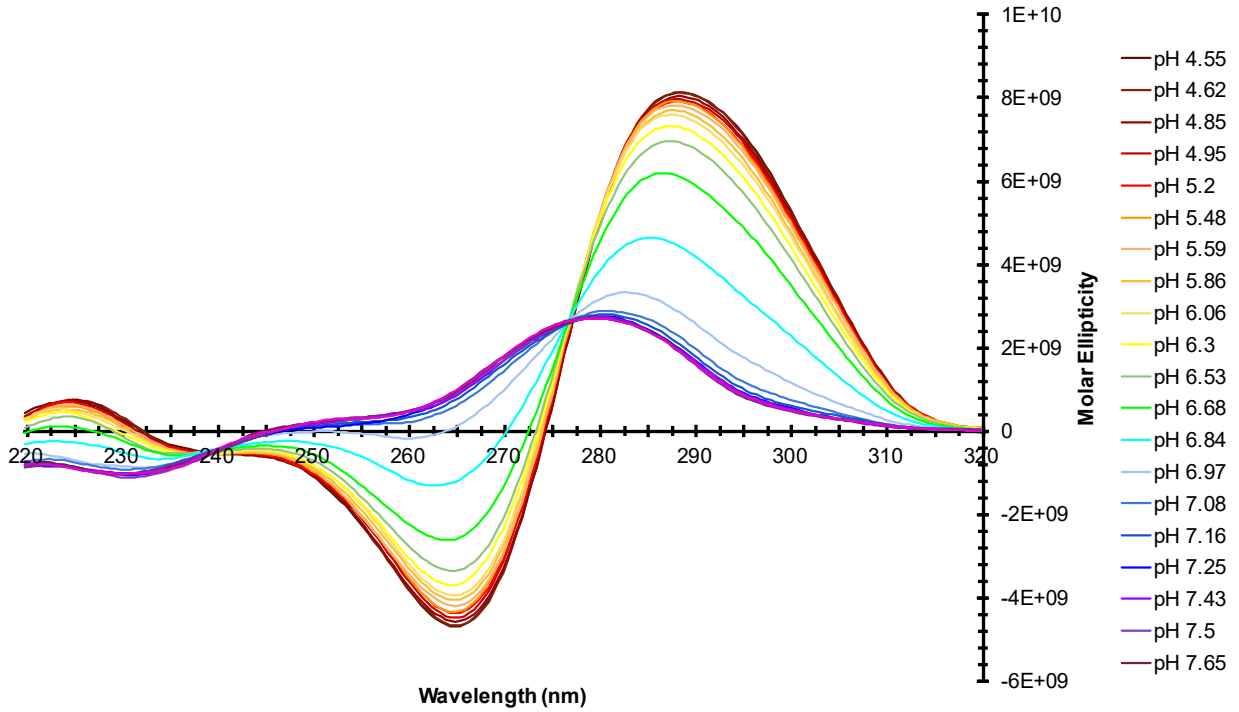


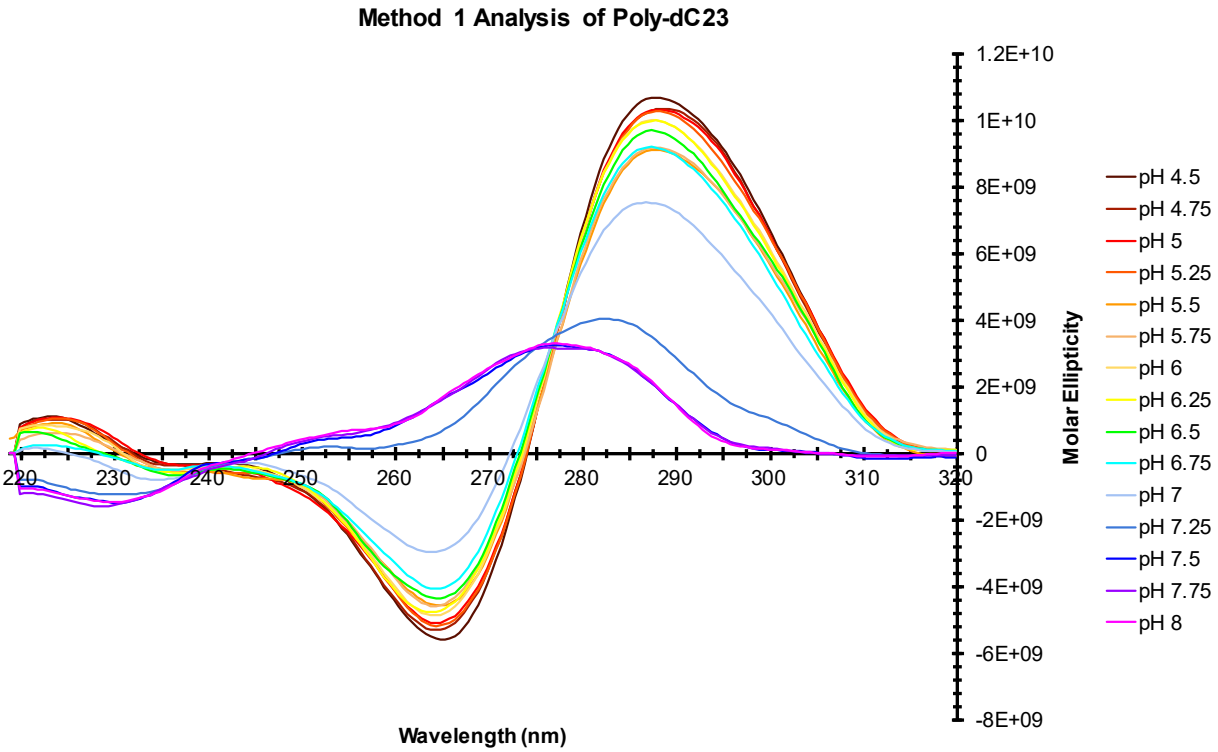


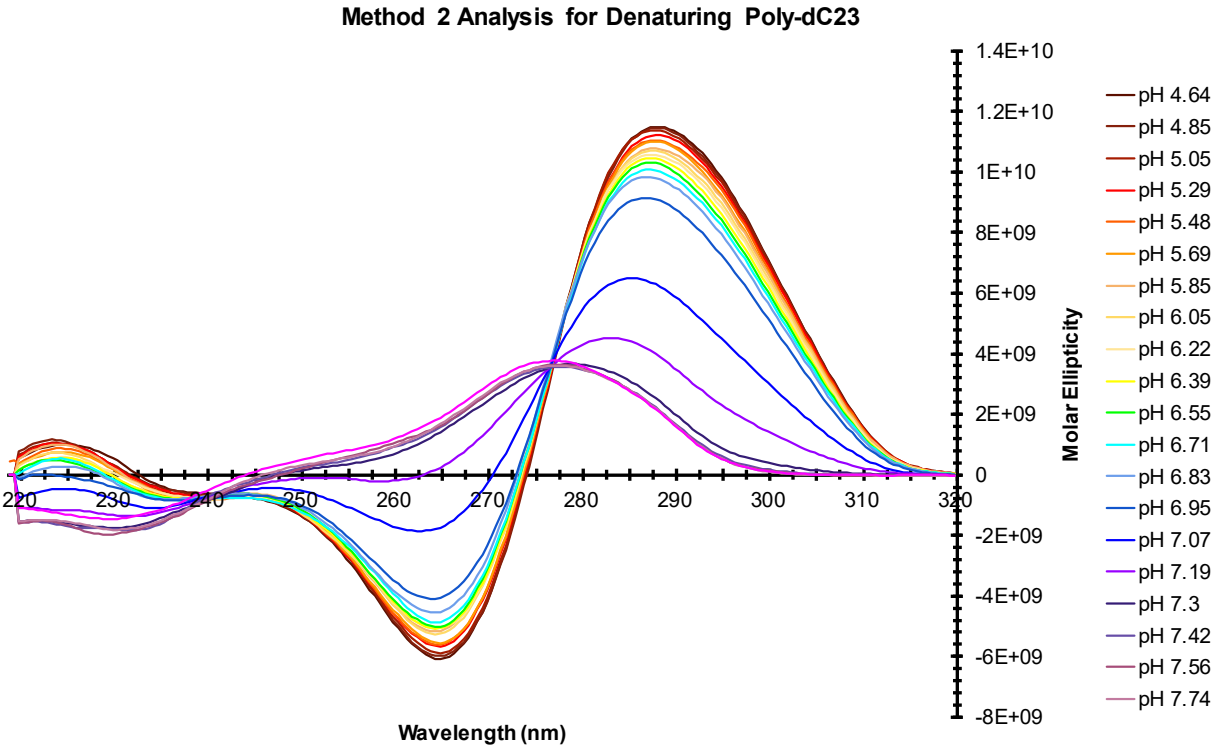

Method 2 Analysis for Annealing Poly-dC15
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.47
pH 4.65
pH 4.78
pH 4.96
pH 5.09
pH 5.25
pH 5.52
pH 5.79
pH 6.05
pH 6.27
pH 6.52
pH 6.74
pH 6.93
pH 7.06
pH 7.28
pH 7.54
pH 7.72
pH 7.92


Method 1 Analysis of Poly-dC19
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.5
pH 4.75
pH 5
pH 5.25
pH 5.5
pH 5.75
pH 6
pH 6.25
pH 6.5
pH 6.75
pH 7
pH 7.25
pH 7.5
pH 7.75
pH 8


Method 2 Analysis of Denaturing for Poly-dC19
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
-8E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.6
pH 4.76
pH 4.92
pH 5.26
pH 5.52
pH 5.74
pH 5.98
pH 6.14
pH 6.26
pH 6.39
pH 6.56
pH 6.7
pH 6.83
pH 6.94
pH 7.12
pH 7.28
pH 7.45
pH 7.6
pH 7.78
pH 8


Method 2 Analysis for Annealing of Poly-dC19
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.55
pH 4.62
pH 4.85
pH 4.95
pH 5.2
pH 5.48
pH 5.59
pH 5.86
pH 6.06
pH 6.3
pH 6.53
pH 6.68
pH 6.84
pH 6.97
pH 7.08
pH 7.16
pH 7.25
pH 7.43
pH 7.5
pH 7.65


Method 1 Analysis of Poly-dC23
1.2E+10
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
-8E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.5
pH 4.75
pH 5
pH 5.25
pH 5.5
pH 5.75
pH 6
pH 6.25
pH 6.5
pH 6.75
pH 7
pH 7.25
pH 7.5
pH 7.75
pH 8


Method 2 Analysis for Denaturing Poly-dC23
1.4E+10
1.2E+10
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
-8E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.64
pH 4.85
pH 5.05
pH 5.29
pH 5.48
pH 5.69
pH 5.85
pH 6.05
pH 6.22
pH 6.39
pH 6.55
pH 6.71
pH 6.83
pH 6.95
pH 7.07
pH 7.19
pH 7.3
pH 7.42
pH 7.56
pH 7.74

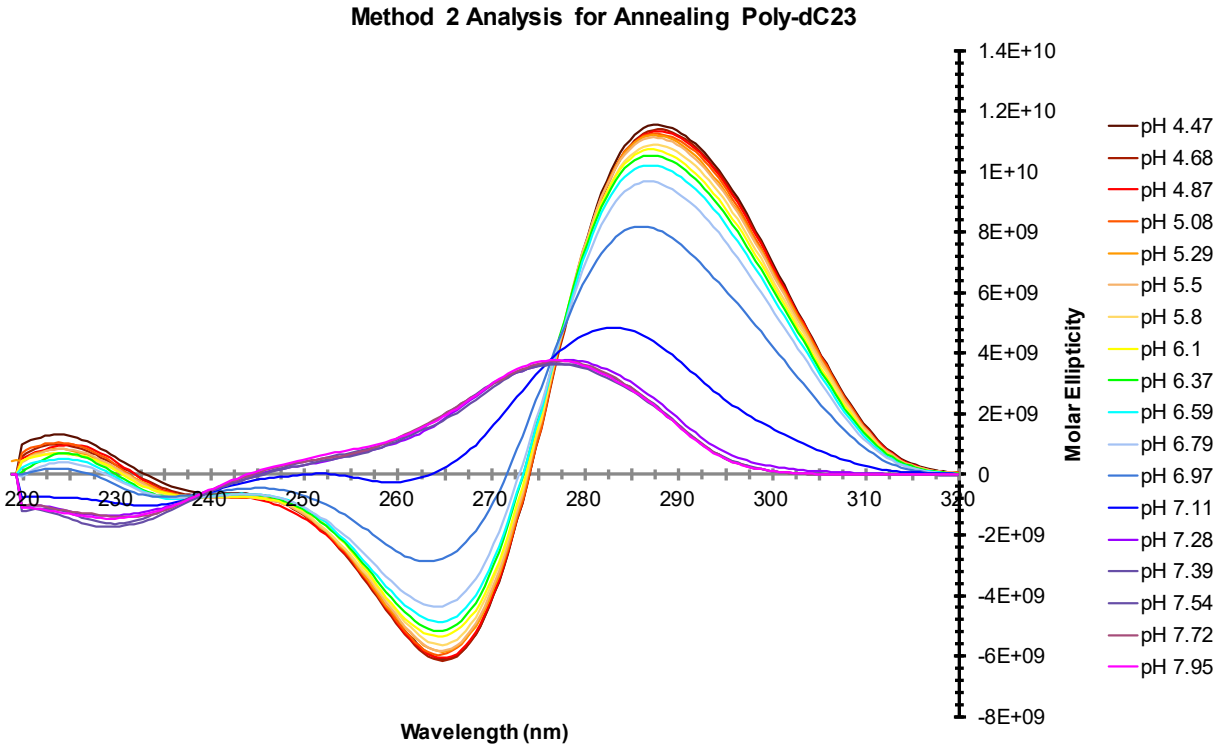
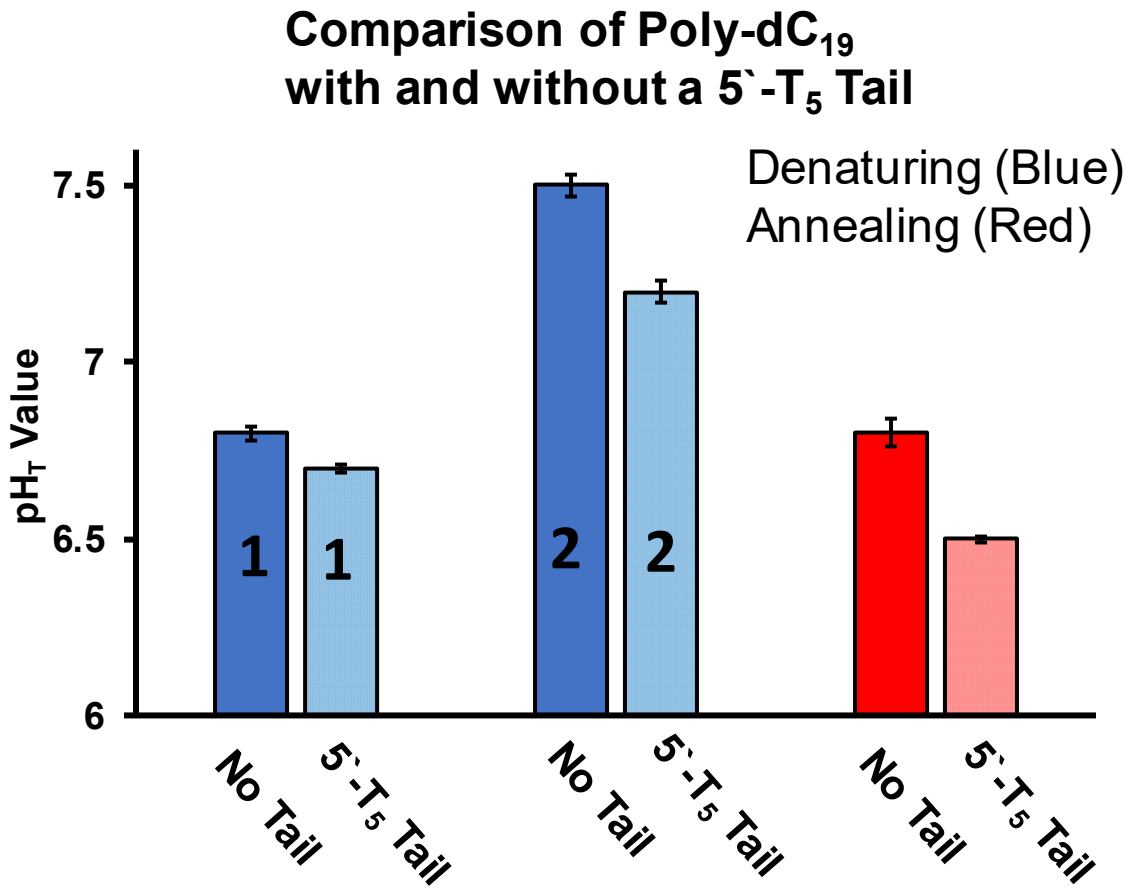

Method 2 Analysis for Annealing Poly-dC23
1.4E+10
1.2E+10
1E+10
8E+09
6E+09
4E+09
2E+09
0
-2E+09
-4E+09
-6E+09
-8E+09
220
230
240
250
260
270
280
290
300
310
320
Molar Ellipticity
Wavelength (nm)
pH 4.47
pH 4.68
pH 4.87
pH 5.08
pH 5.29
pH 5.5
pH 5.8
pH 6.1
pH 6.37
pH 6.59
pH 6.79
pH 6.97
pH 7.11
pH 7.28
pH 7.39
pH 7.54
pH 7.72
pH 7.95

Figure S2. Comparison of pH_T values obtained from method 2 analysis of dC_{19} with and without a 5`-T_5 tail.



The denaturing leg (blue) of the method 2 analysis was biphasic for both strands, and therefore, two pH_T values are shown labeled 1 or 2.