Stability of the pH-Dependent Parallel-Stranded d(CGA) Motif

E. M. Luteran, J. D. Kahn, and P. J. Paukstelis

ABSTRACT

Non-canonical DNA structures that retain programmability and structural predictability are increasingly being used in DNA nanotechnology applications, where they offer versatility beyond traditional Watson-Crick interactions. The d(CGA) triplet repeat motif is structurally dynamic and can transition between parallel-stranded homo-base paired duplex and anti-parallel unimolecular hairpin in a pH-dependent manner. Here, we evaluate the thermodynamic stability and nuclease sensitivity of oligonucleotides composed of the d(CGA) motif and several structurally related sequence variants. These results show that the structural transition resulting from decreasing the pH is accompanied by both a significant energetic stabilization and decreased nuclease sensitivity as unimolecular hairpin structures are converted to parallel-stranded homo-base paired duplexes. Furthermore, the stability of the parallel-stranded duplex form can be altered by changing the 5′-nucleobase of the d(CGA) triplet and the frequency and position of the altered triplets within long stretches of d(CGA) triplets. This work offers insight into the stability and versatility of the d(CGA) triplet repeat motif and provides constraints for using this pH-adaptive structural motif for creating DNA-based nanomaterials.

STATEMENT OF SIGNIFICANCE

This article addresses the stability of the d(CGA) triplet motif and variants in solution. Our study reveals changes in thermodynamic stability and nuclease resistance in response to pH. The identity of the 5′-nucleobase within each triplet and the position and frequency of different triplets within stretches of d(CGA) triplets can tune parallel-stranded duplex stability. This tunability can be used for nanotechnological applications where the specificity of the 5′-

nucleobase pairing interaction is used to order of long stretches of d(CGA) triplets. These results can inform the rational design of pH-sensitive structurally switchable DNA-based nanomaterials.

INTRODUCTION

Nucleobase pairing and stacking interactions inherent in self-complementary DNA provide the structural programmability and predictability needed to create reliable nanoscale self-assembling materials (1-5). Non-complementary sequences have the potential to form alternative DNA structures stabilized by non-Watson Crick base pairing and other interactions. The formation of such structures depends on local sequence and environmental factors, but the rules for stability and structure are not as established as they are for canonical duplex DNA. Predictable non-canonical structures can expand DNA's structural versatility for nanotechnology applications (6-10). Previous work on the nanotechnological applications of non-canonical DNA have focused primarily on motifs known to be biologically relevant, including G-quadruplexes (11-13), i-motifs (14-16), and triplex-forming strands (17-19). One potential advantage of these and other non-canonical structures are their sensitivities to the local environment, including cations, salt concentration, or pH, which allow them to undergo predictable structural changes in response to environmental perturbations. (20-23).

DNA oligonucleotides containing d(CGA) triplet repeats undergo pH-dependent structural switching (Figure 1A) (24-27). Under acidic conditions, the d(CGA) motif forms a parallel-stranded duplex comprised completely of homo-base pairs (C-CH$^+$, G-G, and A-A, Figure 1B) (25,26). At neutral pH, however, oligonucleotides composed of d(CGA) repeats form anti-parallel interactions that are consistent with the formation of a unimolecular DNA hairpin stabilized by G-C pairs separated by potential noncanonical A-A base pairs (24-27). At high

DNA concentration and low temperature, changes in salt concentration and pH allow d(CGA)$_4$ to sample four unique conformations – parallel-stranded duplex, anti-parallel hairpin, anti-parallel B-form duplex, and Z-form duplex (27). The parallel-stranded duplex is primarily observed in acidic conditions, while at higher pH (7.0 and 8.0) and low ionic strength, the structure converts to an anti-parallel hairpin. When ionic strength is increased at pH 8.0, the DNA favors an anti-parallel duplex, while very high ionic strength (4 M NaCl) with trace amounts of Ni^{2+} causes a switch in helix handedness to form Z-DNA.

Though the d(CGA) repeat motif is capable of adopting multiple structural forms, only its parallel-stranded form has been extensively structurally characterized (23,25,28-31). Originally called Π-DNA, the parallel motif is stabilized by hemi-protonated C-CH$^+$, G-G minor groove edge, and A-A Hoogsteen-face homo-base pairs (Figure 1B) (26, 31). The interstrand stacking of the purine nucleobases in the GA dinucleotide steps (Figure 1C) is thought to provide a significant stabilizing force for the motif and has been observed in a number of different sequence contexts (28, 31-34). A crystal structure containing both d(CGA) and internal d(TGA) triplets revealed that interstrand GA stacking is identical in the d(TGA) triplet and the d(CGA) motif, but there is increased structural asymmetry and duplex bending associated with the C-CH$^+$ base pair (29). Additionally, the same cross-strand GA dinucleotide step seen in d(CGA) and d(TGA) was observed in parallel 5′-d(GGA) sequences both in solution and in the non-canonical motif of a continuously base-paired 3D DNA lattice (33,35). In these examples the G-G base pair requires one of the two Gs to adopt a *syn* glycosidic angle to allow Watson-Crick-to-Hoogsteen face base pairing. The overall structural similarity of the d(CGA), d(TGA), and d(GGA) motifs suggests that they could be used together in d(NGA) repeat sequences, where specificity of the

C-CH$^+$, T-T, or G-G base pairs provide the strand registration that would be necessary for designing nanoscale DNA constructs that adopt a single base pairing geometry.

In this work, we investigate the stability of the d(CGA) motif and related sequences in solution at near-physiological temperature and salt concentration. Thermodynamics derived from UV absorbance melting curves indicate that the parallel-stranded homo-duplex structure is energetically more favorable than the anti-parallel structures, and that the frequency and identity of the 5′-nucleotide in the triplet repeat can have dramatic impacts on the relative stability. Also, increased resistance against double-strand specific nucleases in the parallel duplex form relative to the anti-parallel duplex suggests that the parallel motifs may offer unique advantages for the design, assembly, and delivery of nanostructures in cellular environments.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification

DNA oligonucleotides (Table S1) were synthesized on a 1 μmol scale using standard phosphoramidite chemistry on an Expedite 8909 Nucleic Acid Synthesizer (PerSeptive Biosystems, Framingham, MA) with reagents from Glen Research (Sterling, VA). Oligonucleotides were purified by denaturing 20 % (19:1) acrylamide/bis-acrylamide, 7 M urea gel electrophoresis and dialyzed against deionized water, as previously described (35).

Absorbance Melting Curve Procedures

UV melting curves were obtained on a Cary Bio 100 UV-visible spectrophotometer (Varian/Agilent, Palo Alto, CA) equipped with a 12-cell sample changer and Peltier heating/cooling system. DNA samples were diluted to working concentrations in either 20 mM

MES pH 5.5 or 20 mM sodium cacodylate pH 7.0, both supplemented with 100 mM sodium chloride. Sample absorbance data was obtained at 260 nm as the temperature was ramped from 4 ^{o}C to 95 ^{o}C at 1 ^{o}C/min. In addition, data were collected at 260 nm for DNA renaturation from 95 ^{o}C to 4 ^{o}C at 1 ^{o}C/min to check reversibility. Self-masking quartz cuvettes with 1 cm path lengths were used for 1.4 to 20 μM samples, while a 2 mm path length cuvette was used for 50 to 120 μM samples. Final working oligonucleotide concentrations were calculated from the absorbance at 85 ^{o}C.

Melting Curve Data Analysis

The measured absorbance values for samples at high concentration were corrected for deviations from linearity with respect to DNA concentration. A calibration curve was obtained and data falling above the linear range were corrected accordingly. We verified that this procedure yielded absorbance versus wavelength curves that were superimposable at all concentrations for samples that exhibited unimolecular behavior. Thermodynamic parameters for parallel-stranded duplex formation at pH 5.5 were obtained from the fit of each temperature versus absorbance melting curve to the bimolecular van't Hoff expression, as described by Petersheim and Turner (36). Similarly, the van't Hoff expression for unimolecular hairpin formation was used to fit the melt curves and extract thermodynamic parameters for data at pH 7.0 (37). The thermodynamic parameters obtained from the fits to the melting curves are reported as the average ± standard deviation of the results from independent fits from experiments with 4-6 independently prepared samples at concentrations ranging from 2-120 μM.

Circular Dichroism (CD) Spectroscopy

CD spectra were obtained using a Jasco J-810 spectropolarimeter fitted with a thermostatted cell holder (Jasco, Easton, MD). Samples were prepared using 10 μM DNA in the same buffers used

for UV melting experiments: 20 mM MES, 100 mM sodium chloride (pH 5.5), or 20 mM sodium cacodylate, 100 mM sodium chloride (pH 7.0). Samples were incubated at 4 ^{o}C overnight prior to data collection. Data were collected at room temperature at wavelengths from 220-320 nm.

S1 Nuclease Stability Assay

20 μM DNA was incubated with 200 U/mL S1 nuclease (Thermo Fisher Scientific, Waltham, MA) in 40 mM sodium acetate pH 4.5 or 40 mM Tris HCl pH 7.5, 300 mM sodium chloride, and 2 mM zinc sulfate at 22 ^{o}C for 2 hr. The reaction was quenched with 30 mM EDTA and incubated at 70 ^{o}C for 10 min. Samples were analyzed by denaturing 20 % (19:1) acrylamide/bis-acrylamide, 7 M urea polyacrylamide gel electrophoresis and stained with SYBR Gold (Thermo Fisher Scientific, Waltham, MA).

DNase1 Nuclease Stability Assay

DNA was incubated overnight at 4 ^{o}C with 50 mM magnesium formate to induce structure formation, as previously described (38). The DNA was diluted to 40 μM and was incubated with 7 U/mL DNase I (Thermo Fisher Scientific, Waltham, MA), 10 mM MES pH 5.5 or 10 mM Tris pH 7.5, 2.5 mM magnesium chloride, 0.5 mM calcium chloride at 37 ^{o}C for 10 min. The reaction was quenched by addition of 5 mM EDTA and incubated at 75 ^{o}C for 10 min. Samples were analyzed by 20 % denaturing polyacrylamide gel electrophoresis as above and stained with SYBR Gold (Thermo Fisher Scientific, Waltham, MA).

RESULTS AND DISCUSSION

$(CGA)_6$ adopts parallel-stranded duplex and anti-parallel hairpin forms in response to pH

Previous studies demonstrated that $d(CGA)_2$ and $d(CGA)_4$ repeat sequences could transition between anti-parallel hairpin and parallel-stranded duplex (27). To determine if this behavior is also found for $(CGA)_6$, we analyzed the oligonucleotide and related variants by circular dichroism and UV absorbance melting (Table 1). The CD spectrum for $(CGA)_6$ at pH 5.5 showed a prominent positive band at 265 nm and a negative band at 245 nm, consistent with parallel-stranded homo-duplex formation (Figure 2A) (26,27). A clear apparently two-state UV absorbance melting transition was observed for $(CGA)_6$ at pH 5.5, with a concentration-dependent melting temperature (T_m) indicating a reaction molecularity greater than one (Figure 2B), consistent with the parallel bimolecular complexes seen in crystal and solution structures containing the d(CGA) motif (25-29). At pH 7.0, the measured T_m is independent of concentration, indicating a unimolecular structure, and the CD spectrum is characteristic of anti-parallel strands (Figure 2A, B). These results are consistent with $(CGA)_6$ forming anti-parallel hairpin structures at neutral pH and parallel-stranded homo-duplexes under acidic conditions.

Parallel-stranded duplexes are more stable than anti-parallel hairpins

$(CGA)_6$ was analyzed by UV melting to establish thermodynamic parameters for the conformations observed at each pH (Table 2). Thermodynamic parameters for the parallel duplex or hairpin formation were extracted from UV melting curves fit to two-state bimolecular or unimolecular models, respectively. For each dataset, the entire absorbance vs temperature curve was fit with the van’t Hoff equation assuming constant enthalpy ($\Delta H°$) and entropy ($\Delta S°$), as

described in Materials and Methods. While $(CGA)_6$ exhibited apparent two-state melting profiles at both pH 5.5 and 7.0 and had similar melting temperatures in this concentration range, detailed examination of the pH 7.0 melting curves (Figure 2B) shows that there are significant underlying differences in the stabilities of the two structures. The higher baseline slope of the anti-parallel form at pH 7.0 suggest that its structure may be changing as a function of temperature even at low temperature. The decrease in the slope at T_m as concentration increases (i.e. an apparent decrease in $\Delta H°$ and $\Delta S°$) suggests some bimolecular or multimolecular behavior. Comparing the thermodynamic parameters of $(CGA)_6$ determined at pH 5.5 and at 7.0, the standard state formation free energy ($\Delta G°_{37}$) and $\Delta H°$ were destabilized by 12.2 kcal/mol and 58.6 kcal/mol, respectively; the similar T_ms arise from the difference between unimolecular interaction at very high effective local concentration vs. bimolecular interactions at concentrations much less than the standard state 1 M. This suggests that the parallel-stranded duplex form containing intermolecular GA stacking and hydrogen bonding interactions provides substantial structural stabilization versus the anti-parallel hairpin.

$(CGA)_6$ variants exhibit structural switching in response to pH

We designed several sequence variants to investigate how the triplet identity and position in hexa-repeats impact structural switching and stability (Table 1). $(CGA)_6$, $TGA(CGA)_5$, and $GGA(CGA)_5$ were synthesized to probe the significance of the 5′-most capping nucleotides in parallel duplexes. Furthermore, sequences $(CGA)_5TGA$, $(CGATGA)_3$, and $(TGA)_6$ were designed to test the positional and frequency compatibility of d(CGA) and d(TGA) triplets. Finally, $(GAC)_6$ allowed a direct comparison with $(CGA)_6$ to examine how a subtle terminal sequence permutation impacts thermodynamic stability.

CD analysis was used to examine the structural transitions of the variant sequences in response to pH. All variants except $(TGA)_6$ and $(CGATGA)_3$ had CD spectra consistent with the pH-dependent structures seen in $(CGA)_6$ (Figure S1). Interestingly, CD spectra for $(TGA)_6$ had characteristics of anti-parallel strands at both pH 5.5 and 7.0, while the spectra for $(CGATGA)_3$ could represent an equilibrium between both forms. This suggests a decreased pH sensitivity and a preference for anti-parallel strands with increasing d(TGA) triplet content. All variants except $(TGA)_6$ had concentration-dependent, two-state melting curves at pH 5.5 and concentration-independent melting transitions at pH 7.0, indicating bimolecular and intramolecular interactions, respectively. Interestingly, $(CGA)_6$ was the only oligomer that exhibited clear two-state melting at pH 7.0, with all of the other sequence variants having an apparent minor lower-temperature melting transition between 20-40 ^{o}C. The multi-state transitions at pH 7.0 may reflect populations of hairpin and anti-parallel duplex structures (24).

Anti-parallel homogeneity correlates with d(CGA) repeat length

At pH 7.0, there are interesting differences in the melting profiles for $(CGA)_n$ ($n = 4 – 7$) (Figure S2,S3). When $n = 4$ or 5, the melting profile is broad with multiple transitions, whereas when $n \geq 6$ the melting transition is still broad but becomes more two-state. The multi-transition melting profile seen in shorter $(CGA)_n$ repeats ($n \leq 5$) suggests that a temperature dependent equilibrium between hairpin and duplex or higher-order forms may exist. In contrast, the two-state profile for longer repeat sequences ($n \geq 6$) suggest they exist predominantly in the hairpin conformation. This trend is similar to the length-dependent conformational equilibrium seen in other triplet repeats (24) and further supports the importance of oligo lengths in rational structure design.

Estimated $\Delta G°_{37}$ values can be used to predict the energetics of new triplet repeat sequences

To estimate the energetic contribution of each parallel-stranded triplet unit we compared thermodynamic parameters of oligonucleotides with differing number of repeat units. Strong apparent enthalpy-entropy compensation is observed for the set of $(CGA)_n$ variants tested (Figure S4). The positive correlation seen in the $\Delta H°$ vs $\Delta S°$ plot demonstrates the strong enthalpy-entropy compensation, while the $\Delta G°_{37}$ vs $\Delta H°$ plot confirms that the apparent compensation is not an artifact of experimental error. As a simple predictor of the $\Delta G°_{37}$ for $(CGA)_n$ sequences, the number of d(CGA) units were plotted against the respective experimental $\Delta G°_{37}$ value (Figure S5). Interestingly, the *y*-intercept (-0.250 kcal/mol) of the linear fit, which corresponds to the $\Delta G°_{37}$ for a n=1 sequence, could represent the gain in stability upon formation of the first triplet unit slightly overbalancing the entropic penalty of bringing the strands together. The linear relationship can be used to predict the $\Delta G°_{37}$ for a $(CGA)_n$ triplet of any length and can be combined with other estimates (5′-d(GGA), 5′-d(TGA), 3′-d(TGA), see below) to predict $\Delta G°_{37}$ of various sequences (Figure S5). Importantly, these results are internally consistent and provide a baseline for understanding the energetic contributions of these triplets.

The 5′-nucleotide influences parallel-duplex stability

We compared $\Delta G°_{37}$ for variants with different 5′-terminal capping triplets (d(CGA), d(GGA), or d(TGA)) to $(CGA)_5$ to understand how the identity of the 5′ triplet contributed to thermodynamic stability of the parallel-stranded duplex. The 5′-d(GGA) triplet ($\Delta\Delta G°_{37}$ = -1.4 kcal/mol with respect to $(CGA)_5$) and the 5′-d(TGA) triplet (-1.6 kcal/mol with respect to $(CGA)_5$) show similar contributions, both being less stabilizing than the 5′-d(CGA) triplet, estimated above at -

2.4 kcal/mol. If the 5′ triplet contains the anticipated GA interstrand stacking interactions, the differences in $\Delta G°_{37}$ between d(CGA), d(GGA), and d(TGA) triplets can be attributed primarily to the terminal 5′-nucleotide identity. Though the 5′-d(GGA) triplet appears compatible with 5′-d(CGA) in the parallel-stranded duplex, incurring only a slight thermodynamic destabilization, our preliminary modeling suggests that stretches of internal d(GGA) triplets would not be well accommodated due to structural clashes that would arise from an internal G_{syn} residue (33,35). CD data further suggest that stretches of d(GGA) triplets do not form the expected structure at either pH (Figure S6). The small change between the terminal d(GGA) and d(TGA) suggests that the 5′-G and 5′-T are thermodynamically comparable in the parallel duplex. Interestingly, the entropy of formation for $GGA(CGA)_5$ and $(CGA)_6$ is considerably lower than for $TGA(CGA)_5$, which could result from conformational variability introduced by a G residue in the *syn* conformation, as seen in crystal structures containing the parallel-stranded d(GGA) motif. The destabilization caused by the 5′-T can also be attributed to enthalpic destabilization (-92.1 to -85.6 kcal/mol) reflecting the loss of one hydrogen bond by replacing the hemi-protonated $C\text{-}CH^+$ with T-T base pair. Additional electronic effects from the cationic $C\text{-}CH^+$ base pair that are not found in the T-T or G-G base pairs may also contribute.

Notably, the $\Delta G°_{37}$ for $(GAC)_6$ (-13.3 kcal/mol) is different than $(CGA)_6$ (-14.2 kcal/mol), though d(GAC) is a simple permutation of d(CGA). The structural basis for this difference is not immediately clear, but it does suggest that the identity of the terminal homo base pairs – 5′ or 3′ – can influence parallel stranded duplex stability. The variation in thermodynamic stability suggests that the strength of interactions at the 5′-terminus could play an important role in parallel-stranded duplex nucleation and stabilization. Nucleotides providing strong

intermolecular interactions could help to lock the duplex in place to avoid premature end fraying, contributing to enhanced overall thermodynamic stability. Further studies will show whether the apparent non-nearest neighbor effects in d(GAC) vs d(CGA) act at the triplet level or the oligonucleotide terminus level.

d(TGA) triplet position and frequency impacts the thermodynamic stability of mixed $(YGA)_6$ oligomers

The identity of the 5′-pyrimidine in the d(CGA) triplet motif leads to subtle but significant structural changes in the parallel-stranded duplex, evident in crystal structures containing d(CGA) and d(TGA) repeats (29,30). C-CH^+ base pairs induce an asymmetry, likely as a result of a hydrogen bond between N4 of one cytosine to the non-bridging phosphate oxygen of the previous intrachain adenosine (29). In contrast, the T-T base pair retains duplex symmetry at the expense of stacking interactions. Our results show that the frequency of d(TGA) triplets within $(YGA)_6$ oligomers has a significant impact on thermodynamic stability of the parallel-stranded duplex.

At pH 5.5, $(CGA)_6$ has a clear two-state melting transition, indicating structure formation when the sequence is completely comprised of d(CGA) triplets (Figure 3). In contrast, $(TGA)_6$ had no evident melting transition at pH 5.5 or 7.0 (Figure 3), suggesting that the d(TGA) triplet is unable to adopt a stable structure at either pH. This could indicate that the interactions formed by the T-T base pair are either not strong enough to nucleate structure formation or is transient.

We tested the destabilizing effects associated with position and number of d(TGA) triplets in a given $(YGA)_6$ sequence using the sequence variants $TGA(CGA)_5$, $(CGA)_5TGA$, $(CGATGA)_3$, and $(TGA)_6$. As described above, in the parallel duplex form, the addition of a d(TGA) triplet to the 5′-end stabilized the free energy of formation by -1.6 ± 0.7 kcal/mol, compared to $(CGA)_5$ (Table 1). $(CGA)_5TGA$ was tested to observe the impact of the d(TGA) triplet at the 3′-end. This resulted in a destabilization of the free energy of duplex formation by 1.4 ± 0.7 kcal/mol, compared to $(CGA)_5$, suggesting that the position of the d(TGA) triplet significantly impacts the thermodynamic stability of the parallel-stranded duplex. When the d(TGA) is at the 5′-end, the T-T homo-base pair interaction induces a slight thermodynamic destabilization, compared to a 5′-C-CH^+ homo-base pair. In contrast, when the T-T homo-base is pair positioned internally within the 3′ triplet, there is a significantly greater impact on duplex stability, presumably through disruption of the downstream G-G and A-A base pairs. Correspondingly, more substantial destabilization (7.5 kcal/mol) was observed when multiple d(TGA) triplets were added as in $(CGATGA)_3$.

Although the d(TGA) triplets decrease thermodynamic stability, they have the potential to modulate pH sensitivity. CD results for $(TGA)_6$ (Fig. S1) suggest that the presence of multiple d(TGA) triplets correspond to decreased pH sensitivity and preference for the anti-parallel form. d(TGA) triplets can be combined with other strong parallel-strand inducting triplets, such as d(GGA) or d(CGA), potentially providing a unique system for tuning structure formation over specific pH ranges. Information on the thermodynamic stability of sequences containing tandem d(YGA) triplets can be combined with structural data to provide valuable insight in optimizing parallel-stranded duplexes for 3D DNA crystal design and nanoscale applications.

Parallel-stranded duplexes resist nuclease digestion

With increasing interest in using DNA nanotechnology in cellular applications, we assessed the stability of DNA structures in conditions that mimic cellular environments. To gain insight on the behavior of the d(YGA) triplet motif in cellular conditions, we qualitatively examined resistance to nucleases with preferences for double stranded B-DNA (DNase I) or unpaired, single-stranded DNA (S1 nuclease). Control experiments confirmed the activity of both enzymes for different substrates at the pHs tested: DNase I functioned optimally on double-stranded substrates at low pH, while S1 nuclease showed decreased activity at pH 7.5 for its single-stranded substrate (Figure S7).

Each variant tested was incubated with DNase I at pH 5.5 and 7.5 (Figure 4A). At pH 7.5, $(CGA)_6$, $TGA(CGA)_5$, and $(CGA)_5TGA$ show clear cleavage patterns, suggesting that the structure formed at pH 7.5 has DNase I-recognizable regions. This supports previous analysis indicating the formation of an anti-parallel hairpin at near-neutral pH. Importantly, these same variants showed clear DNase I resistance when the reaction was repeated at pH 5.5. This suggests that the parallel-stranded duplex formed by $(CGA)_6$, $TGA(CGA)_5$, $(CGA)_5TGA$, and $(CGATGA)_3$ is not recognized by DNase I. $(CGATGA)_3$ appears to have slight protection at the lower pH, while the cleavage of $(TGA)_6$ is identical at both pHs. Overall, these results are consistent with our UV melting observations showing no stable structure formation for either of these last two oligos at pH 7.0 and some weakly forming structure for $(CGATGA)_3$ at the lower pH.

A similar experiment was performed to test d(YGA) repeat susceptibility to the single-strand specific S1 nuclease. $(CGA)_6$, $TGA(CGA)_5$, $(CGA)_5TGA$, and $(CGATGA)_3$ S1 nuclease cleavage products at pH 7.5 suggest single-stranded regions are primarily associated with hairpin ends (Figure 4B). Intense bands visible at ≥14 nt are consistent with the degradation of terminal non-base paired nucleotides and or single-stranded triplet overhangs at the hairpin terminus. The difference in degradation pattern between all variants suggests that the hairpins formed are not structurally identical. All oligonucleotides were fully digested at pH 7.5 when exposed to S1 nuclease for 2 days, suggesting that the anti-parallel hairpin is dynamic enough to be susceptible to S1 digestion over extended time periods (Figure S8).

$(CGA)_6$, $TGA(CGA)_5$, $(CGA)_5TGA$, and $(CGATGA)_3$ were only partially digested by S1 nuclease at its optimal pH of 4.5 (Figure 4B). UV melting and CD experiments indicate that these oligonucleotides form parallel stranded duplexes in conditions similar to S1 nuclease reaction conditions, explaining the lack of complete S1 digestion. The intense 15-nt band present in the degradation of $(CGA)_6$ suggests that a small population of the parallel-stranded duplexes are frame-shifted by one d(CGA) triplet, creating 3-nt S1 nuclease-accessible single-stranded overhangs. The intense unreacted band remaining at 18-nt suggests that most oligomers in the reaction populated the perfectly aligned parallel-stranded duplex which resisted S1 nuclease digestion. The low molecular weight degradation products could result from frame-shifted duplexes containing more than one triplet overhang. Apparently, once duplexes are formed frame-shifting is extremely slow at room temperature, else equilibration among frames would have exposed all of the sequence to S1 activity.

TGA$(CGA)_5$ and $(CGA)_5$TGA show S1 degradation patterns similar to each other, consistent with triplet cleavage, as intense bands persist at 15, 12, and 9-nt. Interestingly, each intense band is coupled with a faint band, where the position and intensity can be used to gain insight into the S1 cleavage site. The cleavage of a 5′-triplet will produce a fast migrating 15-nt product containing a terminal phosphate, whereas, the cleavage of a 3′-triplet will produce a slower migrating 15-nt product that does not have a 5′-phosphate. The faster migrating 15-nt band is very intense in TGA$(CGA)_5$, suggesting that the 5′- d(TGA) triplet was preferentially digested. The opposite is seen in $(CGA)_5$TGA, where the intense 15-nt band migrates slower, suggesting that the 3′-d(TGA) was cleaved (Figure 4B). The bands at 18, 15, and 12-nt persist when exposed to S1 nuclease for up to 2 days, indicating that the parallel-stranded duplex is stable at room temperature (Figure S8). $(TGA)_6$ was almost completely digested by S1 nuclease by the end of the reaction time, indicating the lack of structure formation at each pH (Figure 4B). This further confirms results from CD and UV melting experiments that suggest $(TGA)_6$ does not form stable structure at acidic or neutral pH.

Stable high molecular weight structures form in all cytosine containing variants

Unexpected high molecular weight (HMW) products form over time for $(CGA)_6$ and all variants, except $(TGA)_6$, and persist in 8 M urea denaturing gels (Figure S9). The formation of the stable HMW structures likely rely on cytosine reactivity, as HMW bands are never observed for $(TGA)_6$ or $(TGA)_5$. We cannot confirm the identity or mechanism of HMW band formation, but we suspect that stable aggregates or crosslinks are forming from degradation products (depurination) of the d(CGA) triplet repeats over time.

CONCLUSION

This work uses a systematic approach to determine the stability of the d(CGA) triplet repeat motif and variants in solution. It is evident that $(CGA)_6$ and variants undergo pH-induced structural switching that coincides with a thermodynamic destabilization. This trend can be linked to the loss of strong interstrand GA base stacking and homo-base pair interactions as pH is increased. Previous studies have investigated the formation of the parallel-stranded duplex of short $(CGA)_n$ repeats ($n \leq 4$) in solution (27) or within a complex crystal structure (29,31). Additional structural studies on sequences containing $(CGA)_n$ triplets ($n > 4$) will be useful in providing in-depth insight into the behavior of long stretches of tandem d(CGA) repeats to further confirm the specific interactions responsible for the difference in thermodynamic stability at pH 5.5 and 7.0. It is clear that the addition of d(TGA) repeats decreases thermodynamic stability, but structure is not completely destroyed if d(TGA) units are in tandem with strong parallel-strand inducing triplets. $(CGA)_6$ and related variants have been proposed to be useful in designing new parallel-stranded DNA architectures. Specifically, particles decorated with oligonucleotides containing $d(CGA)_n$ triplets could be localized or separated based on structural changes caused by pH (Figure 5). Considering slight structural and thermodynamic deviations between each terminal 5′-residue, this motif could also serve as a discriminator in programmable pairing of long stretches of d(YGA) triplets. Additionally, the parallel-stranded duplex formed in acidic pH offers increased nuclease resistance, furthering the potential for this material to be used in cellular applications. The solution stability data presented in this work can be used to strategically design and optimize sequences for use in the rational design of 3D DNA crystals or nanotechnology applications that could benefit from pH-triggered structural switching.

AUTHOR CONTRIBUTIONS

E.M.L. designed and performed experiments, analyzed the data, and wrote the manuscript. J.D.K. contributed to data analysis, discussion, and writing. P.J.P. designed experiments, contributed to data analysis and discussion, and wrote the manuscript.

REFERENCES

1. Seeman, N. C. 1982. Nucleic Acid Junctions and Lattices. J. Theor. Biol. 99:237-247.
2. Rothemund, P. W. K. 2006. Folding DNA to create nanoscale shapes and patterns. Nature. 440:297-302.
3. Jones, M. R., Seeman, N. C., and Mirkin, C. A. 2015. Programmable materials and the nature of the DNA bond. Science. 347:1260901.
4. Seeman, N. C. 2010. Nanomaterials based on DNA. Annu. Rev. Biochem. 79:65-87.
5. Wu, D., Wang, L., Li, W., Xu, X., and Jiang, W. 2017. DNA nanostructure-based drug delivery nanosystems in cancer therapy. Int. J. Pharm. 533:169-178.
6. Rich, A. 1993. DNA comes in many forms. Gene. 135:99-109.
7. Chakraborty, S., Sharma, S., Maiti, P. K., and Krishnan, Y. 2009. The poly dA helix: a new structural motif for high performance DNA-based molecular switches. Nucleic Acids Res. 37:2810-2817.
8. Szabat, M., and Kierzek, R. 2017. Parallel-stranded DNA and RNA duplexes – structural features and potential applications. FEBS J. 284:3986-3998.
9. Mergny, J. L., and Sen D. 2019. DNA Quadruple Helices in Nanotechnology. Chem. Rev. 119:6290-6325.

10. Dong, Y., Yang, Z., and Liu, D. 2014. DNA nanotechnology based on i-motif structures. Acc. Chem. Res. 47:1853-1860.
11. Sha, R., Xiang, L., Liu, C., Balaeff, A., Zhang, Y., Zhang, P., Li, Y., Beratan, D. N., Tao, N., and Seeman, N. C. 2018. Charge splitters and charge transport junctions based on guanine quadruplexes. Nat. Nanotechnol. 13:316-321.
12. Kotlyar, A. B., Borovok, N., Molotsky, T., Cohen, H., Shapir, E., and Porath, D. 2005. Long, Monomolecular Guanine-Based Nanowires. Adv. Mater. 17:1901-1905.
13. Li, Z., and Mirkin, C. A. 2005. G-quartet-induced nanoparticle assembly. J. Am. Chem. Soc. 127:11568-11569.
14. Nesterova, I. V., and Nesterov, E. E. 2014. Rational Design of Highly Responsive pH Sensors Based on DNA i-Motif. J. Am. Chem. Soc. 136:8843-8846.
15. Song, L., Ho, V. H., 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. Healthc. Mater. 2:275-280.
16. Keum, J. W., and Bermudez, H. 2012. DNA-based delivery vehicles: pH-controlled disassembly and cargo release. Chem. Commun. 48:12118-12120.
17. Chandrasekaran, A. R., and Rusling, D. A. 2018. Triplex-forming oligonucleotides: a third strand for DNA nanotechnology. Nucleic Acids Res. 46:1021-1037.
18. Liao, W. C., Riutin, M., Parak, W. J., and Willner, I. 2016. Programmed pH-Responsive Microcapsules for the Controlled Release of CdSe/ZnS Quantum Dots. ACS Nano. 10:8683-8689.
19. Chen, Y., and Mao, C. 2004. Reprogramming DNA-Directed Reactions on the Basis of a DNA Conformational Change. J. Am. Chem. Soc. 126:13240-13241.

20. Largy, E., Marchand, A., Amrane, S., Gabelica, V., and Mergny, J. L. 2016. Quadruplex Turncoats: Cation-Dependent Folding and Stability of Quadruplex-DNA Double Switches. J. Am. Chem. Soc. 138:2780-2792.
21. Engelhard, D. M., Nowack, J., and Clever, G. H. 2017. Copper-Induced Topology Switching and Thrombin Inhibition with Telomeric DNA G-Quadruplexes. Angew. Chem. Int. Ed. Engl. 56:11640-11644.
22. Liu, D., and Balasubramanian, S. 2003. A proton-fuelled DNA nanomachine. Angew. Chem. Int. Ed. Engl. 42:5734-5736.
23. Muser, S. E., and Paukstelis, P. J. 2012. Three-Dimensional DNA Crystals with pH-Responsive Noncanonical Junctions. J. Am. Chem. Soc. 134:12557-12564.
24. Zhang, M., Huang, X., Smith, G. K., Yang, X., and Gao, X. 1996. Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. J. Mol. Biol. 264:323-336.
25. Robinson, H., van der Marel, G. A., van Boom, J. H., and Wang, A. H. 1992. Unusual DNA conformation at low pH revealed by NMR: parallel-stranded DNA duplex with homo base pairs. Biochemistry. 31:10510-10517.
26. Robinson, H., and Wang, A. H. 1993. 5′-CGA sequence is a strong motif for homo base-paired parallel-stranded DNA duplex as revealed by NMR analysis. Proc. Natl. Acad. Sci. U.S.A. 90:5224-5228.
27. Kejnovska, I., Tumova, M., and Vorlickova, M. 2001. $(CGA)_4$: parallel, anti-parallel, right-handed and left-handed homoduplexes of a trinucleotide repeat DNA. Biochim. Biophys. Acta. 1527:73-80.

28. Sunami, T., Kondo, J., Kobuna, T., Hirao, I., Watanabe, K., Miura, K., and Takenaka, A. 2002. Crystal structure of d(GCGAAAGCT) containing a parallel-stranded duplex with homo base pairs and an anti-parallel duplex with Watson-Crick base pairs. Nucleic Acids Res. 30:5253-60.
29. Tripathi, S., and Paukstelis, P. J. 2016. Structural Implications of Homopyrimidine Base Pairs in the Parallel-Stranded d(YGA) Motif. ChemBioChem. 177:1177-1183.
30. Tripathi, S., Zhang, D., and Paukstelis, P. J. 2015. An intercalation-locked parallel-stranded DNA tetraplex. Nucleic Acids Res. 43:1937-1944.
31. Wang, Y., and Patel, D. J. 1994. Solution structure of the d(T-C-G-A) duplex at acidic pH. A parallel-stranded helix containing C+.C, G.G and A.A pairs. J. Mol. Biol. 242:508-526.
32. Robinson, H., van Boom, J. H., and Wang, A. H. 1994. 5′-CGA Motif Induces Other Sequences To Form Homo Base-Paired Parallel-Stranded DNA Duplex: The Structure of $(G\text{-}A)_n$ Derived from Four DNA Oligomers Containing $(G\text{-}A)_3$ Sequence. J. Am. Chem. Soc. 116:1565-1566.
33. Kettani, A., Bouaziz, S., Skripkin, E., Majumdar, A., Wang, W., Jones, R. A., and Patel, D. J. 1999. Interlocked mismatch-aligned arrowhead DNA motifs. Structure. 7:803-815.
34. Rippe, K., Fritsch, V., Westhof, E., and Jovin, T. M. 1992. Alternating d(G-A) sequences form a parallel-stranded DNA homoduplex. EMBO J. 11:3777-3786.
35. Paukstelis, P. J., Nowakowski, J., Birktoft, J. J., and Seeman, N. C. 2004. Crystal structure of a continuous three-dimensional DNA lattice. Chem. Biol. 11:1119-1126.

36. Petersheim, M., and Turner, D. H. 1983. Base-Stacking and Base-Pairing Contributions to Helix Stability: Thermodynamics of Double-Helix Formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. Biochemistry. 22:256-263.
37. Siegfried, N. A., and Bevilacqua, P. C. 2009. Thinking inside the box: designing, implementing, and interpreting thermodynamic cycles to dissect cooperativity in RNA and DNA folding. Methods Enzymol. 455:365-393.
38. Saoji, M., and Paukstelis, P. J. 2015. Sequence-dependent structural changes in a self-assembling DNA oligonucleotide. Acta Cryst. D71:2471-2478.
39. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. 1991. Functional group substitutions as probes of hydrogen bonding between GA mismatches in RNA internal loops. J. Am. Chem. Soc. 113:4313-4322.
40. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. 1998. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs. Biochemistry. 37:14719-14735.
41. Allawi, H. T., and SantaLucia, J., Jr., 1997. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry. 36:10581-10594.

DNA Name	Sequence (5′→3′)
$(CGA)_7$	d(CGACGACGACGACGACGACGA)
$(CGA)_6$	d(CGACGACGACGACGACGA)
$(CGA)_5$	d(CGACGACGACGACGA)
$(CGA)_4$	d(CGACGACGACGA)
$TGA(CGA)_5$	d(TGACGACGACGACGACGA)
$(CGA)_5TGA$	d(CGACGACGACGACGATGA)
$(CGATGA)_3$	d(CGATGACGATGACGATGA)
$(TGA)_6$	d(TGATGATGATGATGATGA)
$GGA(CGA)_5$	d(GGACGACGACGACGACGA)
$(GAC)_6$	d(GACGACGACGACGACGAC)

TABLE 1. Sequences of DNA oligonucleotides. Underlined nucleotides differ from the $(CGA)_n$ pattern.

Thermodynamic Parameters, pH 5.5[A]					
Sequence	**T_m (°C)[B]**	**ΔG^o_{37} (kcal/mol)**	**ΔH^o (kcal/mol)**	**ΔS^o (e.u.)**	**$\Delta\Delta G^o_{37}$ (kcal/mol)[C]**
$(CGA)_7$	61	-15.6 ± 0.9	-107.9 ± 7.2	-298 ± 21	-3.8 ± 1.0
$(CGA)_6$	60	-14.2 ± 0.7	-92.1 ± 3.1	-251 ± 9	-2.4 ± 0.8
$(CGA)_5$	54	-11.8 ± 0.4	-75.8 ± 1.4	-207 ± 4	0
$(CGA)_4$	43	-8.9 ± 0.2	-60.9 ± 3.5	-168 ± 11	+2.9 ± 0.4
$TGA(CGA)_5$	58	-13.4 ± 0.5	-85.6 ± 2.4	-233 ± 7	-1.6 ± 0.7
$(CGA)_5TGA$	49	-10.3 ± 0.3	-65.8 ± 2.2	-179 ± 7	+1.4 ± 0.5
$(CGATGA)_3$	32	-6.9 ± 0.3	-59.7 ± 3.2	-171 ± 10	+4.9 ± 0.5
$(TGA)_6$	-	-	-	-	-
$GGA(CGA)_5$	56	-13.2 ± 0.7	-91.2 ± 2.6	-252 ± 8	-1.4 ± 0.8
$(GAC)_6$	58	-13.3 ± 0.4	-85.8 ± 2.3	-234 ± 7	-1.5 ± 0.6
Thermodynamic Parameters, pH 7.0[D]					
Sequence	**T_m (°C)[B]**	**ΔG^o_{37} (kcal/mol)**	**ΔH^o (kcal/mol)**	**ΔS^o (e.u.)**	**$\Delta\Delta G^o_{37}$ (kcal/mol)**
$(CGA)_6$	55	-2.0 ± 0.2	-33.5 ± 1.1	-102 ± 3	+9.8 ± 0.4

TABLE 2 Thermodynamic parameters for $(CGA)_6$ and variants at pH 5.5 and 7.0. (A) Thermodynamic parameters obtained at pH 5.5 are reported as averages of van't Hoff analysis curve fitting results assuming bimolecular duplex formation, with oligomer concentrations between 1-120 x 10^{-6} M. Errors in ΔH^o and ΔS^o are reported as one standard deviation from at least 4 concentrations, and error in ΔG^o_{37} is propagated as previously described (39-41). (B) T_m values are calculated for oligomer concentration of 3 x 10^{-6} M. (C) $\Delta\Delta G^o_{37}$ for each sequence calculated with respect to $(CGA)_5$. Negative values represent a stabilized free energy of formation, whereas positive values represent destabilized free energy of formation. (D) Thermodynamic parameters and T_m for pH 7.0 are reported as the average of 4 van't Hoff curve fit datasets, assuming unimolecular hairpin formations, with oligomer concentrations between 1-10 x 10^{-6} M.

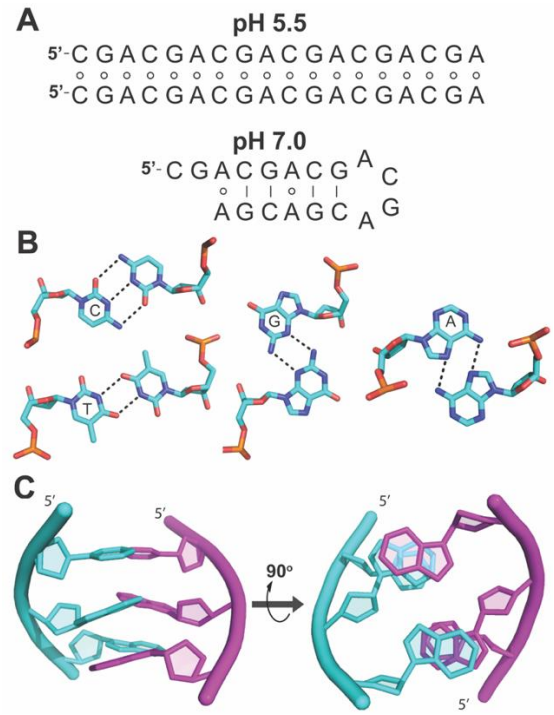

FIGURE 1 d(CGA) triplet motif. (A) $(CGA)_6$ is predicted to form a parallel-stranded homo-duplex at pH 5.5 and an anti-parallel hairpin at pH 7.0, based on NMR and circular dichroism analysis (24-27). Circles indicate non-canonical base pairing and solid lines represent canonical base pairing. It is unclear whether the anti-parallel hairpin anticipated at pH 7.0 forms with (shown) or without (not shown) A-A homo-base pairs (24). (B) Homo-base pairing interactions in the parallel-stranded form of the d(YGA) motif: protonation of cytosine N3 allows formation of a hemiprotonated C-CH^+ homo-base pair; T-T homo-base pair; G-G minor groove edge homo-base pair; and A-A Hoogsteen-face homo-base pair. (C) A d(CGA) triplet forming a parallel-stranded homo-duplex, shown from the side and also rotated 90° so the 5′ ends go into the page, to show interstrand GA stacking interactions. Schematics in (B) and (C) used PDB structures 4RIM and 5CV2.

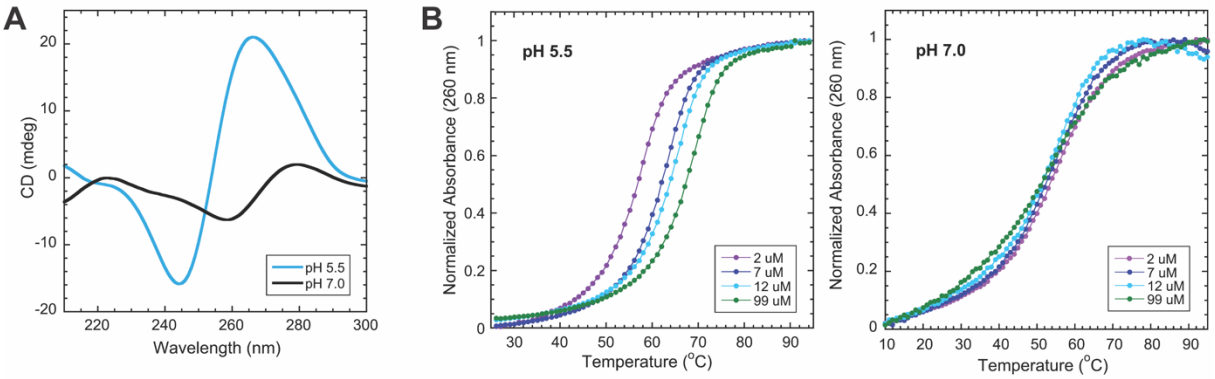

FIGURE 2 $(CGA)_6$ adopts a parallel-stranded duplex or an anti-parallel hairpin in response to pH. (A) Circular dichroism spectra of 10 μM $(CGA)_6$ at pH 5.5 and pH 7.0. At pH 5.5, the prominent positive band at 265 nm and negative band at 245 nm are consistent with parallel-stranded duplex formation. At pH 7.0, the positive band at 275 nm and negative band at 258 nm are characteristic of anti-parallel strands (26). (B) Normalized temperature vs absorbance curves for $(CGA)_6$ show concentration dependent T_m at pH 5.5 and concentration independent T_m at pH 7.0, suggesting bi-/multi-molecular and largely unimolecular structure formation, respectively.

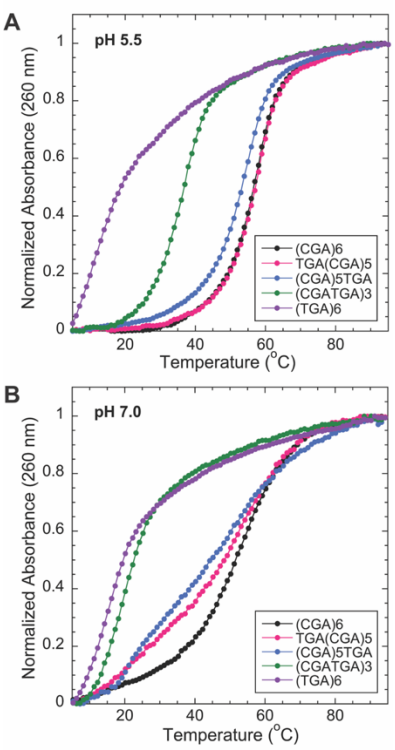

FIGURE 3 The addition of d(TGA) triplets decreases melting temperature of $(YGA)_6$ oligomers. Normalized temperature vs absorbance curves for $(YGA)_6$ oligomers with d(TGA) triplets (n = 6, 3, 1, 0). (A) At pH 5.5, all variants except $(TGA)_6$ exhibit two-state melting. (B) At pH 7.0, $(CGA)_6$ is the only variant to retain two-state melting. $(CGA)_5TGA$ and $TGA(CGA)_5$ have broad melting curves with multiple transitions, while $(TGA)_6$ and $(CGATGA)_3$ do not have clear melting transitions, suggesting a lack of stable structure formation.

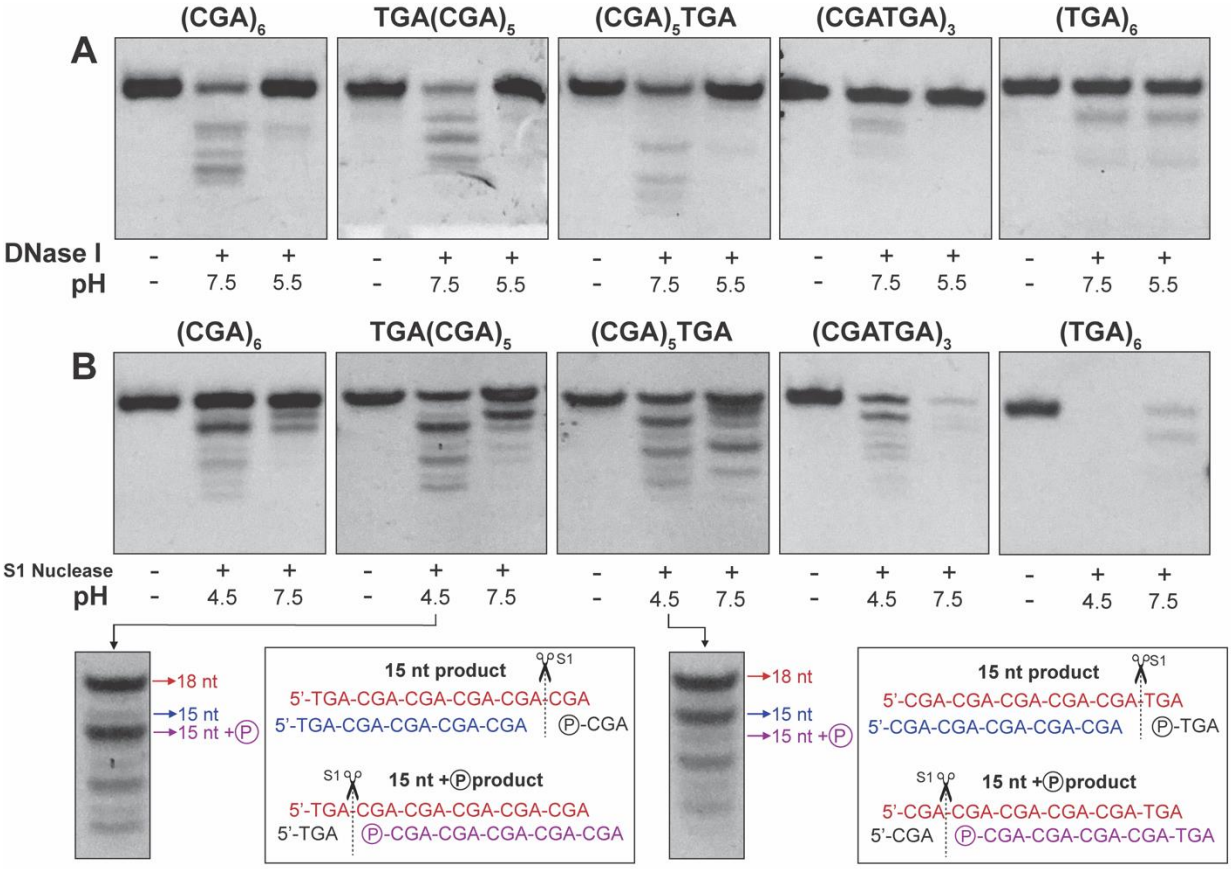

FIGURE 4 Parallel-stranded duplex enhances nuclease resistance. (A) $(YGA)_6$ variants digested with DNase I at pH 7.5 or 5.5. DNase I preferentially cuts double stranded DNA, suggesting that the parallel duplexes formed at pH 5.5 are resistant to DNase I degradation. (B) $(YGA)_6$ variants digested with S1 nuclease at pH 7.5 or 4.5. Red labeled sequences correspond to full length DNA (18-nt), without any nuclease modification. Blue sequences represent the nuclease cleavage product without a terminal phosphate group. The magenta sequence corresponds to the nuclease cleavage product containing a terminal phosphate group, leading to increased mobility.

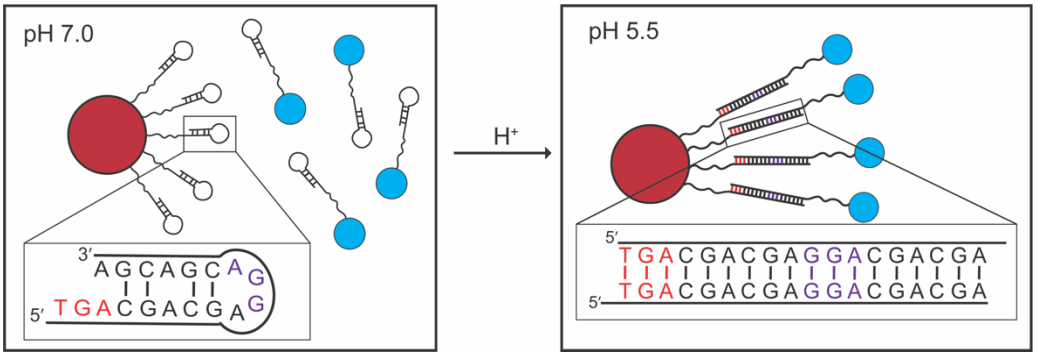

FIGURE 5 Potential use of $(CGA)_6$ and variants in DNA nanotechnology applications where pH changes could localize particles together. At pH 7.0, d(CGA) forms hairpin structures, preventing duplex formation and particles remain separate. At pH 5.5, parallel-stranded duplexes can form and localize particles of interest together. d(TGA) and d(GGA) triplets can be used to ensure desired registration and length of the linking duplex. Linker stretches solely comprised of d(CGA) triplets could be used, but distance between particles of interest could vary.