Structural Studies of a Trinucleotide Repeat Sequence Using 2-Aminopurine[†]

Natalya N. Degtyareva, Michael J. Reddish, Bidisha Sengupta, and Jeffrey T. Petty*

Department of Chemistry, Furman University, Greenville, South Carolina 29613

Received December 5, 2008; Revised Manuscript Received January 22, 2009

ABSTRACT: The secondary structure of repeated trinucleotide sequences results in the development of several neurodegenerative diseases, and these studies consider the $(CAG)_8$ sequence that forms a stem−loop hairpin. The structural and thermodynamic properties of this hairpin are assessed using 2-aminopurine substitutions for adenine at six positions in this repeated sequence. Circular dichroism spectra and thermal denaturation experiments show that the secondary structure is not disturbed by the modifications. The local structure of the hairpin was monitored using the fluorescence intensities of 2-aminopurines, the changes in the intensity relative to the denatured state, and the sensitivity of the fluorescence to quenching by acrylamide. To establish the stem and loop characteristics in $(CAG)_8$, known reference points for stem, loop, and exposed base motifs were used. In the vicinity of the loop, the bases become more solvent exposed, which suggests that the instability associated with this repeated hairpin influences the global secondary structure. These results provide the basis to interpret the structures adopted by other repeated (CAG) structures.

Triplet repeat sequences comprise a significant proportion of the human genome and are the molecular basis of several debilitating neurological diseases (*1*). An important factor in the development and severity of this class of diseases is the length of the repeated sequence. To illustrate, individuals afflicted with Huntington's disease have a median of 45 (CAG) repeats in the coding region of the Huntington's gene, in contrast with the normal population with a median of 19 repeats (*2*). The diseases are linked with the propensity of repeated sequences to adopt hairpin-based structures (*1*−*6*). Intramolecular folding that occurs during replication can lead to repetitive incorporation of bases into the hairpin, thus expanding the number of repeated triplet sequences (*7*). (CNG) repeats (where N = A, T, G, or C) are linked with a number of neurological diseases, and their structures have received significant attention (*4*). The stem of the hairpin has two C/G complementary base pairs and one mismatched self-pair. Spectroscopic studies support the formation of a stem with a conformation that is typical of duplex G/C-rich sequences (*8, 9*). Cross-linked strands with an antiparallel orientation have been demonstrated by NMR spectra, and these studies have also considered the structural features of the mismatched base pair (*2, 10*). The formation of a compact, intramolecular structure is deduced from the enhanced electrophoretic mobility relative to duplex DNA and on the concentration independence of thermal denaturation (*8, 11*). Chemical and biochemical probes of exposed bases demonstrate the presence of a single-stranded loop that connects the stem (*7*). Thermodynamic studies provide the basis to understand the factors that dictate the stability and structure of the trinucleotide repeat hairpins. The canonical G/C pairs stabilize the hairpin, while the mismatched bases have a destabilizing effect that is tempered by stacking interaction with neighboring bases and by reorientation to form hydrogen bonds with the opposing base (*7, 8*). Over a large range of sizes from 4 to 25 repeats, the trend for the (CNG) repeats from most to least stable is (CGG) > (CTG) > (CAG) > (CCG) (*8*). Long-range structural effects are evident as the stability does not follow the expected increase with the length of the oligonucleotide, and this plateau in the stability of the hairpin structures has been interpreted from thermodynamic and structural perspectives (*8, 9*).

The goal of our studies is to use 2-aminopurine to understand the local structure and thermodynamics of repeated (CAG) sequences. For the present studies, the $(CAG)_8$ sequence was used because it adopts a structure with a single stem and loop (*2, 8, 9*). To provide the structural and thermodynamic information, single substitutions of 2-aminopurine were made because this adenine analogue has spectroscopic properties that are sensitive to the DNA conformation. Specifically, the fluorescence quantum yield is suppressed by base stacking, thus providing a means of characterizing the structural motifs with secondary structures via the extent of solvent exposure (*12, 13*). Thus, our purpose is to relate the range of emission intensities to the key structural and thermodynamic features of the $(CAG)_8$ sequence, and recent studies are relevant (*14*−*16*). First, the properties of the repeated sequences depend on the location in the primary sequence (*14*). Within a hairpin comprised of three (CAG) repeats, fluorescence quenching and anisotropy studies show variations in the environments of the substituted 2-aminopurines, which suggests that stacking and hydrogen-bonding interactions induce significant order within the loop (*14, 17*). Second, the key structural motifs of ribonucleic acids have been investigated using 2-aminopurine (*15, 16*). In these studies, extrahelical/bulged, base paired, and loop bases were identified

[†] Funding was provided by the National Institutes of Health [R15GM084442 and P 20 RR-016461 (from the National Center for Research Resource)], the National Science Foundation (0718588), and the Henry Dreyfus Teacher-Scholar Awards Program.

* To whom correspondence should be addressed. E-mail: jeff.petty@furman.edu. Phone: (864) 294-2689. Fax: (864) 294-3559.

10.1021/bi802225y CCC: $40.75
Published on Web 01/26/2009

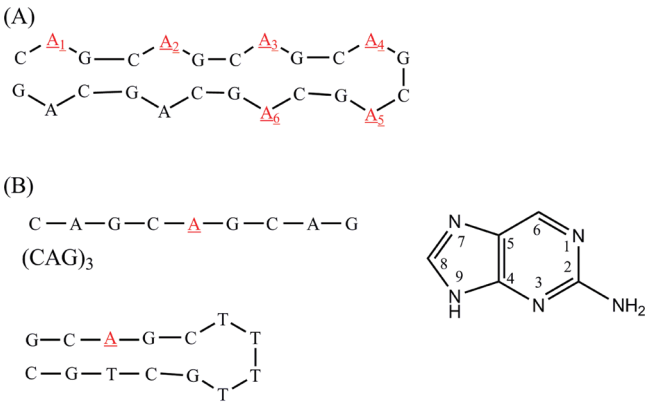

FIGURE 1: (A) Model structure for the blunt-ended hairpin formed by $(CAG)_8$. (B) Structures of the references $(CAG)_3$, CAG-HP, and 2-aminopurine. For the oligonucleotides, the positions of the 2-aminopurine substitutions are indicated in red and are underlined.

based on the fluorescence intensities, fluorescence lifetimes, and the sensitivities of the fluorescence intensities to solutes. Detailed structural insight was obtained using an unstructured control sequence to isolate the local structural properties of the RNA sequence. In our studies, six oligonucleotides with single substitutions of 2-aminopurine are investigated. Through the fluorescence intensities, the changes in the fluorescence intensities with temperature, and the acrylamide quenching constants, the structure and thermodynamics of the $(CAG)_8$ sequence were studied. Three references were used to compare the structure of $(CAG)_8$ to common structural features adopted by nucleic acids. The results support a hairpin structure in which long-range effects influence the distinction between the loop and stem motifs.

MATERIALS AND METHODS

The buffer components (Sigma-Aldrich, St. Louis, MO) and acrylamide (Acros Organics, Belgium) were used as received. All measurements were conducted in a buffer with 10 mM $H_2PO_4^-/HPO_4^{2-}$ (pH = 7) and 50 mM NaCl. The oligonucleotides (Integrated DNA Technologies, Coralville, IA, and Operon Biotechnologies, Huntsville, AL) are shown in Figure 1A,B. The nomenclature for the modified $(CAG)_8$ sequences indicates the position of the 2-aminopurine substitution for adenine. For example, in 4-AP, the replacement occurs in the fourth repeat from the 5′ end of the sequence. The oligonucleotides were purified by HPLC and annealed by fast cooling of the solution from 95 to 4 °C on ice. The oligonucleotide concentration was determined spectrophotometrically by extrapolation to 25 °C of the absorbance at 260 nm of the denatured single-stranded oligonucleotide using the following extinction coefficients: 237300 M^{-1} cm^{-1} for $(CAG)_8$, 225500 M^{-1} cm^{-1} for 1-AP, 2-AP, 3-AP, 4-AP, 5-AP, and 6-AP, and 108268 M^{-1} cm^{-1} for CAG-HP.[1] The extinction coefficients of the oligonucleotides at 260 nm were derived from the nearest neighbor approximation, and the small contribution from 2-aminopurine (1000 M^{-1} cm^{-1} at 260 nm) was accounted for in the calculation. (*18*) Absorbance was collected as a function temperature on a Cary 300 spectrometer equipped with the multicell holder (Varian, Palo Alto, CA). The experiments were started at 20 °C, and the temperature was increased to 95 °C at a rate of 1 °C/min. The absorbance of the buffer solution was used to correct the absorbance associated with the DNA. Fluorescence spectra were collected on a Fluorolog spectrometer (Jobin-Yvon Horiba, Edison, NJ) at 15 °C, equipped with DataMax 3.4 software and Neslab RTE-7 circulating bath. The excitation wavelength was set at 307 nm and the emission wavelength was 370 nm, and the samples were equilibrated for 5 min at each temperature before recording the signal. For the melting curves collected via emission and absorbance, the melting temperatures were determined as the intersection point between experimental melting curve and the median between lower and upper baselines of the melting curve. The upper and lower baselines correspond to the folded and denatured forms of the hairpins. The enthalpy of melting, ΔH, was obtained from the van't Hoff analysis of the linear plot of ln K vs T^{-1}, where K is the equilibrium constant between double stranded and melted hairpin at the temperature T and R is the universal gas constant (*19*). For normalized melting profiles, the fluorescence intensities were acquired for samples of the modified $(CAG)_8$ sequences and $(CAG)_3$ at the same concentrations. Repeated experiments showed the reproducibility of the fluorescence ratio was ≈10%. Circular dichroism spectra were recorded on a JASCO spectrometer in a quartz cell with 1 cm path at room temperature. Differential scanning calorimetry scans were collected using a VP-DSC instrument (Microcal, Northampton, MA). The sample concentration was 200 μM in buffer. The sample was heated from 10 to 95 °C at 1 °C/min. Initially, the instrument was equilibrated by running buffer in the reference and sample cell, and after reaching the reproducible traces the sample was loaded into the cell during the cooling cycle and run five times. The buffer traces were subtracted from the sample traces and analyzed using Origin 7.0 program software. The enthalpy change for melting was calculated for each scan, and then the average was calculated. For the quenching studies, acrylamide solutions were prepared daily. The typical concentration range of the oligonucleotide was 0.5−1.5 μM. The standard Stern−Volmer analysis was an adequate model to determine the static quenching constants (*20*).

RESULTS

$(CAG)_8$ Characterization. $(CAG)_8$ was chosen as a model trinucleotide repeat sequence that favors hairpin formation (Figure 1) (*8, 9*). In support of the folded hairpin, nondenaturing gel electrophoresis showed that the mobility of this 24 base oligonucleotide is comparable to that of an ∼12 base pair duplex (Supporting Information Figure 1S). The formation of an intramolecular complex is demonstrated by the independence of the melting temperature (57 ± 1 °C) for thermal denaturation over the 100-fold concentration range of 0.5−50 μM (*19*). Base pairing interactions in the stem are indicated by the 15% hyperchromicity at 260 nm in the thermal denaturation profile (Supporting Information Figure 2S). Furthermore, the base pairing interactions produce a structure with a B-like conformation, as indicated by the similarity of the circular dichroism spectra of $(CAG)_8$ with

[1] Abbreviations: CAG-HP, GC/2-aminopurine/GC TTTT GCTGC; $(CAG)_3$, CAG C/2-aminopurine/G CAG.

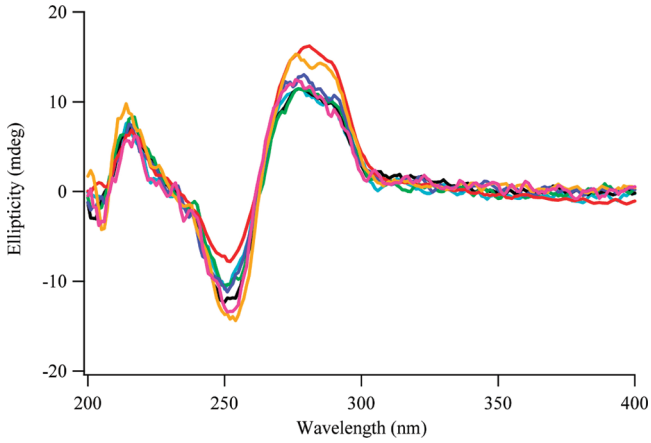

FIGURE 2: Circular dichroism spectra of the $(CAG)_8$ sequences at 20 °C: 1-AP (yellow), 2-AP (pink), 3-AP (cyan), 4-AP (black), 5-AP (green), 6-AP (blue), and unmodified $(CAG)_8$ (red).

G/C-rich duplex DNA (Figure 2) (*8, 9, 21*). This latter result suggests that the two G/C base pairs in each repeat dictate the overall conformation of the oligonucleotide. The formation of a single type of hairpin structure is indicated by the comparison of the spectroscopically and thermodynamically derived enthalpy changes for thermal denaturation (*9*). To extract the enthalpy change using the absorbance changes with temperature, the conversion from an annealed to a denatured state is assumed. The two-state hypothesis is validated, as the van't Hoff derived enthalpy change of 29.0 (±1.5) kcal/mol of oligonucleotide agrees with the 28.0 (±1.0) kcal/mol of oligonucleotide determined from the model-independent measurements using differential scanning calorimetry. Thermodynamic analysis affords insight into the effect of the mismatch on the stability of the hairpin. Based on a 10 base pair stem (Figure 1A), the average enthalpy change is 2.9 kcal/mol of base pair, which is lower than the 8 kcal/mol of base pair for duplex B-form DNA composed of canonical base pairs (*22, 23*). A statistical mechanical model that describes DNA unfolding predicts that the melting temperature is ~12 °C higher and the enthalpy change is ~20 kcal/mol higher than is experimentally measured. (*24*) Discrepancies of this magnitude have also been observed for duplex melting and could arise from a variety of factors. (*25*) Collectively, the experimental results indicate that $(CAG)_8$ adopts an intramolecularly folded structure with a loop connecting a double-stranded stem. The B-form conformation is dictated by the G/C base pairs, but the overall structure is destabilized by the A/A mismatches.

Influence of Modifications on Structure and Stability of $(CAG)_8$. Six oligonucleotides were used with adenine substituted with 2-aminopurine at six positions in the $(CAG)_8$ sequence (Figure 1). Prior studies have established that this modification has a minor effect on DNA structure, conformation, and reactivity (*14, 26, 27*). To evaluate the effect of 2-aminopurine on the global structure of the $(CAG)_8$ sequence, circular dichroism and thermal denaturation experiments were conducted. The similarities of the spectral shapes and of the ellipticities in the circular dichroism spectra indicate that the overall B-form conformation of $(CAG)_8$ is not altered by the single substitutions of 2-aminopurine (Figure 2). The thermodynamic effects of 2-aminopurine substitution were examined using thermal denaturation (Table 1). The global unfolding was monitored at 260 or 270 nm, where the molar absorptivity of the oligonucleotide is largest. The similarities of the melting temperatures indicate that the stabilities of the unmodified and modified oligonucleotides are comparable. In some cases, the modified oligonucleotides have melting temperatures that are 1−3 °C higher than the unmodified oligonucleotide, which has also been observed when 2-aminopurine substitutions are made in the canonical adenine/thymine base pair (*26*). The enthalpy changes derived from the van't Hoff analysis of the absorbance changes are within 15% of the value for the unsubstituted oligonucleotide. Thus, the circular dichroism and thermal denaturation experiments indicate that the global structure and stability of the $(CAG)_8$ hairpin is not significantly perturbed by the single base substitutions with 2-aminopurine.

Fluorescence Data. To illuminate local structural information about the hairpin formed by $(CAG)_8$, fluorescence spectra of the modified oligonucleotides were collected at 15 °C, a temperature that favors the hairpin state (Figure 3 and Supporting Information Figure 2S). The hairpin structure in Figure 1A provides a context for interpreting the fluorescence intensities. The relatively high emission intensity for 1-AP is attributed to local denaturation at the terminus of the stem (i.e., "fraying"), which is characteristic of hairpin structures (*28*). The duplex stem is established by the second repeat, as the relatively low fluorescence intensity of 2-AP indicates that this substitution has minimal solvent exposure. The fluorescence intensities increase moving toward the loop, with the substitution at the fifth repeat having the highest emission intensity. The equivalent intensities from the 3-AP and 6-AP support the symmetry of these positions in the blunt-ended hairpin (Figure 1A). Thus, the results support a stem and loop structure with sequestered and solvent-exposed bases, respectively, but the degree of solvent exposure varies with the location within the hairpin.

To provide further structural interpretation of the fluorescence intensities, three references were used: a substituted nine base oligonucleotide ($(CAG)_3$ in Figure 1B), a substituted hairpin oligonucleotide with canonical base pairs (CAG-HP in Figure 1B), and the unincorporated nucleobase (Figure 1B). The short length of the nine base $(CAG)_3$ precludes hairpin formation, as evidenced by the variation of the fluorescence intensity with temperature (Supporting Information Figure 3S). Thus, this oligonucleotide provides a structural reference that exhibits loop-like behavior, as the 2-aminopurine is embedded in an unstructured sequence. The extent of base incorporation into the helix was evaluated using the hairpin sequence CAG-HP, which has canonical base pairs in the stem connected by a four base loop of thymines. Because 2-aminopurine forms a complementary base pair with thymine, this oligonucleotide serves as a reference point for the fully sequestered form of 2-aminopurine (*29*). Finally, the upper limit of solvent exposure was assessed using the 2-aminopurine nucleobase.

These references allow interpretation of the fluorescence intensities within the context of important structural motifs adopted by nucleic acids (Figure 3). The comparable fluorescence intensities for $(CAG)_3$ and for 1-AP support the end-fraying of the hairpin. The substituted oligonucleotides 2-AP, 3-AP, 4-AP, and 6-AP have lower fluorescence than $(CAG)_3$, indicating that these bases are incorporated to varying degrees into the double helical stem. Relative to the

Table 1: Thermodynamic Data for Thermal Denaturation and Acrylamide Quenching Constants for $(CAG)_8$ Sequences[a]

	$T_{m,abs}$ (°C)[b]	ΔH_{abs} (kcal/mol)[c]	$T_{m,\ fluor}$ (°C)[b]	ΔH_{fluor} (kcal/mol)[c]	K_q (M^{-1})[d]
1-AP	58.8 ± 1.0	32.0 ± 1.9	ND[e]	ND	9.8 ± 0.5
2-AP	60.1 ± 1.4	28.5 ± 2.4	52.8 ± 0.3	31.2 ± 0.9	5.0 ± 0.5
3-AP	60.3 ± 1.0	29.9 ± 4.7	54.4 ± 0.9	32.7 ± 1.3	9.1 ± 0.8
4-AP	61.1 ± 1.0	32.1 ± 1.4	50.4 ± 0.6	30.9 ± 2.9	11.8 ± 1.0
5-AP	59.2 ± 1.4	28.9 ± 1.5	47.0 ± 1.0	20.0 ± 2.0	14.8 ± 1.2
6-AP	60.2 ± 1.0	29.5 ± 3.7	51.8 ± 0.3	32.7 ± 1.1	10.6 ± 0.4
$(CAG)_3$	ND	ND	ND	ND	10.0 ± 0.8
CAG-HP	69.9 ± 0.5	46.3 ± 3.6	69.4 ± 0.5	42.2 ± 1.3	4.2 ± 1.7
$(CAG)_8$	57.4 ± 0.6	29.0 ± 1.5	ND	ND	ND

[a] The standard deviations are derived from a minimum of three measurements. [b] The melting temperatures (T_m) were derived from the absorbance (second column) and the fluorescence (fourth column) changes. [c] The enthalpy changes (ΔH) were derived from the absorbance (third column) and the fluorescence (fifth column) changes using the van't Hoff model. [d] The quenching constants (K_q) for acylamide quenching of the fluorescence from 2-aminopurine. [e] ND = not determined.

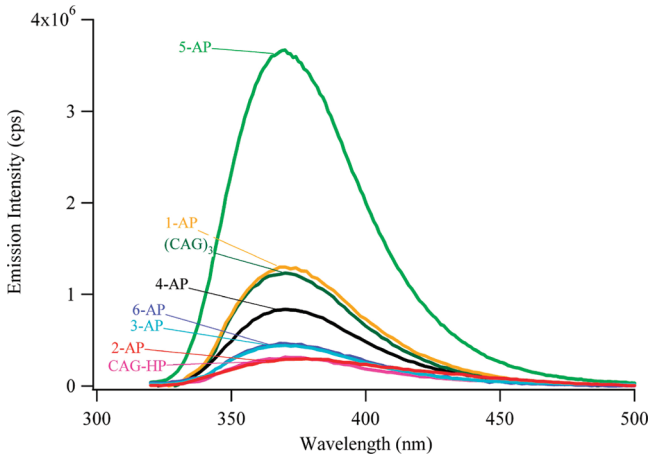

FIGURE 3: Fluorescence spectra of the $(CAG)_8$ sequences at a concentration of 1 μM and recorded at 15 °C (λ_{ex} = 307 nm): 1-AP (yellow), 2-AP (pink), 3-AP (cyan), 4-AP (black), 5-AP (green), 6-AP (blue), CAG-HP (red), and $(CAG)_3$ (dark green). The emission maxima are 370 nm for all of the spectra.

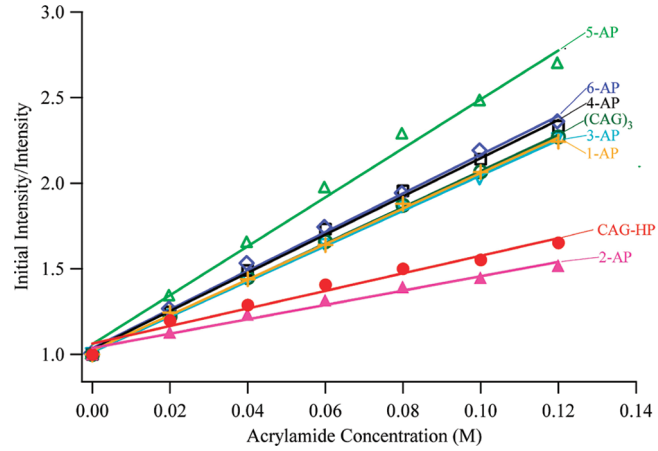

FIGURE 4: Stern−Volmer plot of the quenching of 2-aminopurine-modified $(CAG)_8$ with acrylamide (λ_{ex} = 307 nm, λ_{em} = 370 nm): 1-AP (yellow, crosses), 2-AP (pink, solid triangles), 3-AP (cyan, inverted open triangles), 4-AP (black, open squares), 5-AP (green, open triangles), 6-AP (blue, rhombuses), CAG-HP (red, solid circles), and $(CAG)_3$ (dark green, open circles).

fully sequestered base in CAG-HP, the substitution in the second repeat has a comparable intensity, suggesting that this A/A mismatch is incorporated into the stem as well as a canonical base pair. The relatively high fluorescence for 5-AP suggests that this region of the loop has a high degree of solvent exposure. However, the influence of the surrounding bases is evident, as the intensity for this substituted oligonucleotide is ≈20-fold lower than the free nucleobase (Supporting Information Figure 4S).

Acrylamide Quenching of Substituted (CAG)$_8$. The position dependence of solvent exposure was also assessed via acrylamide quenching of the 2-aminopurine fluorescence (Table 1 and Figure 4). The current studies focused on lower concentrations to understand the limiting behaviors of the substituted oligonucleotides. At higher concentrations, curvature was observed in the Stern−Volmer plots, and we are now investigating this effect in relation to the structure of the hairpin (*16*). The quenching constants are clustered into three groups and follow the structural patterns deduced from the fluorescence intensities. The smallest quenching constants are observed for the 2-aminopurine substitutions in 2-AP and in CAG-HP, suggesting that bases in the vicinity of the second repeat are most protected from the solvent, as expected for a stem motif. The substitutions in 1-AP and in $(CAG)_3$ have higher quenching constants, which is attributed to their higher degree of solvent exposure in their denatured states. For 3-AP/6-AP and 4-AP, the quenching constants are comparable to that of the $(CAG)_3$, within experimental error. As with the trends in the fluorescence intensities, this similarity suggests that the stem bases in the vicinity of the loop adopt a more open conformation. The quenching constant is largest for 5-AP, consistent with the highest degree of solvent exposure for this region of the oligonucleotide. This quenching efficiency is similar to that measured for 2-aminopurine substituted at the exposed bulge in an RNA sequence but is smaller than that of the free nucleobase (Table 1) (*15*). Thus, the substitution in the fifth repeat still involves interactions of the 2-aminopurine with its neighboring bases.

Emission Studies via Thermal Denaturation. Because the modified oligonucleotides have common denatured states with equivalent sequence and structural environments for the 2-aminopurine substitutions, the structure and thermodynamics of the low temperature form of $(CAG)_8$ were further interrogated using thermal denaturation. For the free nucleobase, elevated temperatures suppress the fluorescence due to enhanced nonradiative decay (Supporting Information Figure 3S) (*30*). To dissect structural from photophysical changes in the fluorescence of the incorporated 2-aminopurines, the fluorescence from the $(CAG)_8$ sequences was normalized relative to the fluorescence from $(CAG)_3$ (*15*). This latter oligonucleotide was chosen because both the free and incorporated forms of the nucleobase follow the same behavior (Figure 5 and Supporting Information Figure 3S).

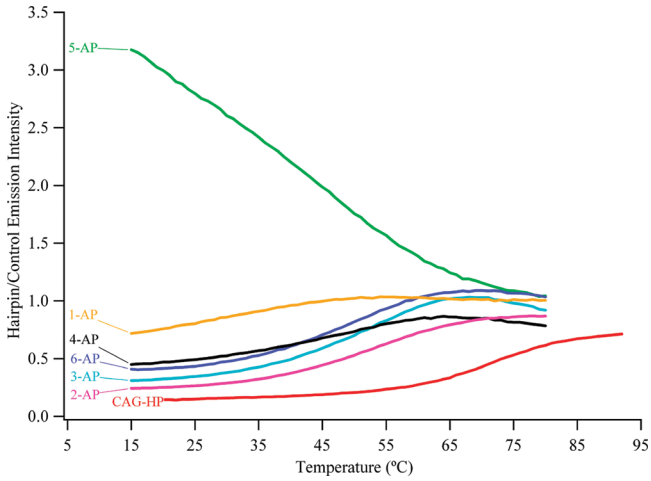

FIGURE 5: Emission intensities of 2-aminopurine-modified $(CAG)_8$ as a function of temperature normalized using the emission from $(CAG)_3$ ($\lambda_{ex} = 307$ nm, $\lambda_{em} = 370$ nm): 1-AP (yellow), 2-AP (pink), 3-AP (cyan), 4-AP (black), 5-AP (green), 6-AP (blue), and CAG-HP (red).

The broad profile for 1-AP has a low melting temperature, further supporting end-fraying. The substitutions in 2-AP and 3-AP/6-AP exhibit increases in their fluorescence intensities with temperature that are comparable to CAG-HP. Thus, these 2-aminopurine substitutions in $(CAG)_8$ are sequestered to a comparable degree as in a canonical duplex, and the bases are more solvent exposed in the denatured state. Relative to modifications in the central part of the stem, 2-aminopurine placement in the fourth repeat has a smaller relative change in the fluorescence, which is ascribed to a greater degree of solvent exposure due to greater loop character. The full development of the loop is evident in 5-AP, in which the fluorescence decreases with the temperature. This observation suggests that unfolding relaxes a constraint imposed by the loop, thus resulting in a greater degree of base stacking in the denatured form. Similar behavior has also been observed for 2-aminopurine substitutions in bulged and looped regions of nucleic acids (*15, 31*).

The thermodynamic properties of the stem and loop motifs were studied using the thermal denaturation profiles of the modified oligonucleotides (Table 1). The results show a distinction between the global unfolding as assessed via absorbance and the local denaturation measured via fluorescence. No difference is observed for the melting temperatures for CAG-HP that has a canonical 2-aminopurine/thymine base pair, which shows that the primary sequence context does not influence the melting temperature derived from the fluorescence measurements (*32*). However, for the $(CAG)_8$ oligonucleotides, the melting temperatures derived from fluorescence are lower than from absorbance. The contrasting behavior for the hairpins with complementary and mismatched base pairs indicates that the mismatch promotes enhanced solvent exposure of the 2-aminopurine prior to the global denaturation (*26, 33*). Relative to the other oligonucleotides, 4-AP and 5-AP have larger differences between the global and local melting temperatures, which suggests that the loop relaxes prior to the overall unfolding of the hairpin. Furthermore, the enthalpy changes in the loop region are smaller than for the global denaturation, suggesting that weaker interactions stabilize the bases in this region of the hairpin. In the stem, the similarity of the enthalpy changes for global and local denaturation suggests that the interactions of the mismatched bases are dominated by the surrounding bases (*26, 33*).

DISCUSSION

These studies utilized the fluorescence properties of 2-aminopurine to understand the structural and thermodynamic characteristics of DNA comprised of trinucleotide repeats. Sequence redundancy suggests that multiple secondary structures are allowed, and the $(CAG)_8$ sequence was chosen because it favors a single hairpin structure with a duplex stem connected by a single-stranded loop. The solvent exposure at specific sites in $(CAG)_8$ was probed using 2-aminopurine modifications, and the results are consistent with the hairpin structure in Figure 1A. Stem-like properties are associated with suppressed fluorescence that is enhanced with temperature and with protection from quenching by acrylamide. Loop-like behavior is associated with enhanced fluorescence that is suppressed with temperature and with a greater efficiency of quenching by acrylamide. An important finding is that intermediate behavior is discerned by using 2-aminopurine. For example, relative to $(CAG)_3$, substitutions in the third and sixth positions exhibit both base stacking as evidenced by their low fluorescence intensity and solvent exposure as indicated by their similar quenching constants with acrylamide. The clearest illustration of intermediate stem and loop behavior is in 4-AP. The substitution has a fluorescence intensity between the fully sequestered base in CAG-HP and the solvent-exposed base in $(CAG)_3$. In this fourth repeat, the fluorescence increases in the denatured state, but the enhancement is small relative to other positions. Finally, the acrylamide quenching constant is comparable to the single-stranded $(CAG)_3$.

Further elaboration on the hairpin structure adopted by $(CAG)_8$ is obtained from the 2-aminopurine studies. Within the polymorphic structures that are feasible for repeated sequences, one distinction is hairpins with blunt and slipped termini (*34*). Substitutions in the third and sixth repeats argue in favor of the blunt-ended arrangement for $(CAG)_8$. For these two oligonucleotides, the fluorescence intensities, the fluorescence enhancements with temperature, and the acrylamide quenching constants are similar, thus suggesting that the 2-aminopurines have comparable chemical environments within the duplex stem. However, this symmetry is not reflected in the loop region. For (CAG) repeats, stacking interactions in the stem are maximized by the tetraloop AGCA, thus suggesting that the substitutions in the fourth and fifth positions should be equivalent (Figure 1A). (*34*) However, the 2-aminopurine in the fifth repeat has limited interactions with its neighboring bases, as indicated by the relatively high degree of solvent exposure and the low enthalpy change associated with thermal denaturation. For the fourth repeat, the 2-aminopurine exhibits an intermediate degree solvent exposure, suggesting that the transition from stem to loop occurs in this region of the hairpin. One interpretation of these differences is that the hairpin structure is not symmetric. Another possibility is that the conformations of the bases in the loop vary with the sequence position. In support of these changes in the loop environment, the NMR chemical shifts in the loop bases of (CTG) repeats vary with position. (*35*) Such behavior is expected because

variations in the solvent exposure of the bases are determined by interactions both within the loop and with the stem. (*36*)

A concept that is guiding our current studies is the importance of the long-range effect of the loop on the structure and thermodynamics of repeat sequence hairpins. The mismatches in the stem become increasingly solvent exposed in the vicinity of the stem, as demonstrated by the fluorescence properties of the 2-aminopurines. The range of behavior that can be discerned is illustrated by substitutions in the second and fifth repeats. For 2-AP, the degree of sequestration of the 2-aminopurine is comparable to that of a canonical base pair with thymine. The accommodation of bulky purine−purine pairs within the helix has been demonstrated by the cross-strand hydrogen bonds between adenine−adenine mismatches that are observed in NMR spectra (*2*). Thus, base pairing and stacking stabilize the mismatch, and the models for the hydrogen bonding suggest that 2-aminopurine and adenine can interact similarly (*7*). Substitution in the fifth repeat illustrates the opposite end of the spectral properties of 2-aminopurine in repeated sequences. The duplex stem constrains the loop, resulting in enhanced exposure relative to an unstructured, single-stranded DNA. Such behavior has also been observed for bulges and loops in other nucleic acid structures (*15, 31*). Thus, the fluorescence from 2-aminopurine allows not only the identification of structural motifs but also the characterization of the transition between these motifs. These studies complement NMR studies that can be restricted by higher concentrations that favor the duplex relative to the hairpin and by resonance overlap in longer sequences (*2*). The fluorescence studies provide local structural and thermodynamic information that also complements chemical and biochemical probing that is sensitive to the chemical environment of the bases (*7*). Our studies are now focusing on longer repeats, as these are most relevant to the development of diseases associated with the repeated sequences. The thermodynamic properties of the stem follow the expectation that localized base interactions are the dominant contribution to the stability of the hairpin (*37*). However, hairpins comprised of repeated sequences are destabilized by mismatched bases, and the present results indicate that this instability relaxes the integrity of the stem. We are particularly interested in how this effect influences the structure and thermodynamics of structures that deviate from the simple stem−loop hairpin (*9*).

CONCLUSIONS

The structure of the $(CAG)_8$ hairpin has been studied using the fluorescence properties of the adenine isomer 2-aminopurine. The results are consistent with a hairpin structure, but solvent exposure of the bases increases in the vicinity of the loop. This behavior is attributed to the destabilizing factors in the hairpin comprised of repeated sequences with mismatched bases. These concepts are being applied to understand the structures adopted by longer repeated sequences that are the basis of debilitating neurological diseases.

ACKNOWLEDGMENT

The authors appreciate discussions with L. Marky. The valuable comments of the reviewers were constructive and are appreciated.

SUPPORTING INFORMATION AVAILABLE

Supplemental figures including a gel image of $(CAG)_8$, a plot of the molar absorptivity of 5-AP vs temperature, fluorescence intensities of 2-aminopurine and $(CAG)_3$ vs temperature, and emission spectra of 2-aminopurine, 5-AP, and $(CAG)_3$. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wells, R. D., Dere, R., Hebert, M. L., Napierala, M., and Son, L. S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res. 33*, 3785–3798.
2. Mariappan, S. V., Silks, L. A., Chen, X., Springer, P. A., Wu, R., Moyzis, R. K., Bradbury, E. M., Garcia, A. E., and Gupta, G. (1998) Solution structures of the Huntington's disease DNA triplets, $(CAG)_n$. *J. Biomol. Struct. Dyn. 15*, 723–744.
3. McMurray, C. T. (1999) DNA secondary structure: A common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. U.S.A. 96*, 1823–1825.
4. Sinden, R. R., Potaman, V. N., Oussatcheva, E. A., Pearson, C. E., Lyubchenko, Y. L., and Shlyakhtenko, L. S. (2002) Triplet repeat DNA structures and human genetic disease: Dynamic mutations from dynamic DNA. *J. Biosci. 27*, 53–65.
5. Lenzmeier, B. A., and Freudenreich, C. H. (2003) Trinucleotide repeat instability: A hairpin curve at the crossroads of replication, recombination, and repair. *Cytogenet. Genome Res. 100*, 7–24.
6. Wang, G., and Vasquez, K. M. (2006) Non-B DNA structure-induced genetic instability. *Mutat. Res. 598*, 103–119.
7. Mitas, M. (1997) Trinucleotide repeats associated with human disease. *Nucleic Acids Res. 25*, 2245–2254.
8. Paiva, A. M., and Sheardy, R. D. (2004) Influence of sequence context and length on the structure and stability of triplet repeat DNA oligomers. *Biochemistry 43*, 14218–14227.
9. Amrane, S., Sacca, B., Mills, M., Chauhan, M., Klump, H. H., and Mergny, J.-L. (2005) Length-dependent energetics of $(CTG)_n$ and $(CAG)_n$ trinucleotide repeats. *Nucleic Acids Res. 33*, 4065–4077.
10. Zheng, 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.
11. Petruska, J., Arnheim, N., and Goodman, M. (1996) Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res. 24*, 1992–1998.
12. Rachofsky, E. L., Osman, R., and Ross, J. B. (2001) Probing structure and dynamics of DNA with 2-aminopurine: Effects of local environment on fluorescence. *Biochemistry 40*, 946–956.
13. Xu, D., Evans, K. O., and Nordlund, T. M. (1994) Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry 33*, 9592–9599.
14. Lee, B. J., Barch, M., Castner, E. W., Volker, J., and Breslauer, K. J. (2007) Structure and dynamics in DNA looped domains: CAG triplet repeat sequence dynamics probed by 2-aminopurine fluorescence. *Biochemistry 46*, 10756–10766.
15. Ballin, J. D., Bharill, S., Fialcowitz-White, E. J., Gryczynski, I., Gryczynski, Z., and Wilson, G. M. (2007) Site-specific variations in RNA folding thermodynamics visualized by 2-aminopurine fluorescence. *Biochemistry 46*, 13948–13960.
16. Ballin, J. D., Prevas, J. P., Bharill, S., Gryczynski, I., Gryczynski, Z., and Wilson, G. M. (2008) Local RNA conformational dynamics revealed by 2-aminopurine solvent accessibility. *Biochemistry 47*, 7043–7052.
17. Volker, J., Makube, N., Plum, G. E., Klump, H. H., and Breslauer, K. J. (2002) Conformational energetics of stable and metastable states formed by DNA triplet repeat oligonucleotides: Implications for triplet expansion diseases. *Proc. Natl. Acad. Sci. U.S.A. 99*, 14700–14705.
18. Fox, J. J., Wempen, I., Hampton, A., and Doerr, I. L. (1958) Thiation of nucleosides. I. Synthesis of 2-amino-6-mercapto-9-/8-D-ribofuranosylpurine ("thioguanosine") and related purine nucleosides. *J. Am. Chem. Soc. 80*, 1669–1675.
19. Mergny, J.-L., and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides 13*, 515–537.

20. Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
21. Ren, J., and Chaires, J. B. (1999) Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry 38*, 16067–16075.
22. Spink, C. H., Garbett, N., and Chaires, J. B. (2007) Enthalpies of DNA melting in the presence of osmolytes. *Biophys. Chem. 126*, 176–185.
23. Barone, G., Del Vecchio, P., Esposito, D., Fessas, D., and Graziano, G. (1996) Effect of osmoregulatory solutes on the thermal stability of calf-thymus DNA. *J. Chem. Soc., Faraday Trans. 92*, 1361–1367.
24. Markham, N. R., and Zuker, M. (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res. 33*, W577–W581.
25. Dimitrov, R. A., and Zuker, M. (2004) Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys. J. 87*, 215–226.
26. Ansari, A., Kuznetsov, S. V., and Shen, Y. Q. (2001) Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. U.S.A. 98*, 7771–7776.
27. Bailly, C., and Waring, M. J. (2001) Use of DNA molecules substituted with unnatural nucleotides to probe specific drug-DNA interactions. *Methods Enzymol. 340*, 485–502.
28. Puglisi, J. D., Wyatt, J. R., and Tinoco, I. (1990) Solution conformation of a RNA hairpin loop. *Biochemistry 29*, 4215–4226.
29. Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1986) Base pairing and mutagenesis: observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR. *Proc. Natl. Acad. Sci. U.S.A. 83*, 5434–5438.
30. Ward, D. C., Reich, E., and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem. 244*, 1228–1237.
31. Jiao, Y., Stringfellow, S., and Yu, H. (2002) Distinguishing "looped-out" and "stacked-in" DNA bulge conformation using fluorescent 2-aminopurine replacing a purine base. *J. Biomol. Struct. Dyn. 19*, 929–934.
32. Goodman, M. F., and Ratliff, R. L. (1983) Evidence of 2-aminopurine-cytosine base mispairs involving two hydrogen bond. *J. Biol. Chem. 258*, 12842–12846.
33. Xu, D., Evans, K. O., and Nordlund, T. M. (1994) Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry 33*, 9592–9599.
34. Hartenstine, M. J., Goodman, M. F., and Petruska, J. (2000) Base stacking and even/odd behavior of hairpin loops in DNA triplet repeat slippage and expansion with DNA polymerase. *J. Biol. Chem. 275*, 18382–18390.
35. Mariappan, S. V., Garcoa, A. E., and Gupta, G. (1996) Structure and dynamics of the DNA hairpins formed by tandemly repeated CTG triplets associated with myotonic dystrophy. *Nucleic Acids Res. 24*, 775–783.
36. (a) Varani, G. (1995) Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct. 24*, 379–404. (b) Senior, M. M., Jones, R. A., and Breslauer, K. J. (1988) Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc. Natl. Acad. Sci. U.S.A. 85*, 6242–6246.
37. (a) SantaLucia, J., Jr., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct. 33*, 415–440. (b) Tinoco, I., and Bustamante, C. (1999) How RNA folds. *J. Mol. Biol. 293*, 271–281.

BI802225Y