Published in final edited form as:

Biochem Biophys Res Commun. 2011 December 9; 416(1-2): 146–152. doi:10.1016/j.bbrc.2011.11.013.

Premutation *Huntingtin* Allele Adopts a Non-B Conformation and Contains a Hot Spot for DNA Damage

Daniel A. Jarem and Sarah Delaney*

Department of Chemistry, Brown University, Providence RI 02912

Abstract

The expansion of a CAG trinucleotide repeat (TNR) sequence has been linked to several neurological disorders, for example, Huntington's disease (HD). Healthy individuals have 5–35 CAG repeats. Those with 36–39 repeats have the premutation allele, which is known to be prone to expansion. In the disease state, greater than 40 repeats are present. Interestingly, the formation of non-B DNA conformations by the TNR sequence is proposed to contribute to the expansion. Here we provide the first structural and thermodynamic analysis of a premutation length TNR sequence. Using chemical probes of nucleobase accessibility, we found that similar to (CAG)₁₀, the premutation length sequence (CAG)₃₆ forms a stem-loop hairpin and contains a hot spot for DNA damage. Additionally, calorimetric analysis of a series of (CAG)_n sequences, that includes repeat tracts in both the healthy and premutation ranges, reveal that thermodynamic stability increases linearly with the number of repeats. Based on these data, we propose that while non-B conformations can be formed by TNR tracts found in both the healthy and premutation allele, only sequences containing at least 36 repeats have sufficient thermodynamic stability to contribute to expansion.

Keywords

Huntington's disease; trinucleotide repeat expansion; oxidative damage; non-B DNA conformation; chemical probes; differential scanning calorimetry

1. Introduction

Several neurodegenerative disorders have been associated with the expansion of trinucleotide repeat (TNR) regions of DNA [1,2,3]. The number of repeats falls into one of three categories: (1) shorter repeat lengths that are found in healthy individuals and are not prone to expand; (2) mid-range premutation lengths that are prone to expand to the disease state; and (3) longer repeat lengths found in individuals afflicted by the disease. For example, Huntington's disease (HD) is a disorder that results from the expansion of a CAG TNR region in exon 1 of the *huntingtin* gene [4]. Healthy individuals have 5–35 repeats, 36–39 repeats are considered the premutation range, and in the disease state individuals have 40 or greater repeats. Thus, 36 CAG repeats are of specific interest because this length defines a

© 2011 Elsevier Inc. All rights reserved.

*Corresponding Author: Telephone: (401) 863-3590. Fax: (401) 863-9368., sarah_delaney@brown.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

border between TNR tracts that are stable and are not prone to expand, and TNR tracts that are known to expand and result in HD.

In addition to B-form duplex conformations, structural analysis of CAG TNR sequences has shown, both *in vitro* [5,6] and *in vivo* [7], that these repetitive DNA sequences can adopt intramolecular stem-loop hairpins and other slipped-structure intermediates. It is this unique property of TNR sequences that provides the basis for proposed mechanisms of TNR DNA expansion. Non-B conformations are proposed to cause expansion via polymerase slippage and reiterative synthesis during replication of TNR DNA or during long-patch base excision repair (LP-BER) within a TNR tract of DNA [1,8]. Given the proposed role of non-B conformations in these expansion models, the stability of non-canonical structures is important to consider.

We have previously shown that the stem-loop hairpin structure formed by (CAG)₁₀ contains a hot spot for oxidative DNA damage. A guanine (G) in the loop of the hairpin has both increased susceptibility to a solution-borne oxidizing agent [9] and is repaired less efficiently [9,10] than damage in duplex DNA. Indeed, we [10] and others [11,12] have proposed a 'toxic oxidation cycle' in which TNR expansion occurs as part of a faulty LP-BER event.

Clinical observations have shown that across healthy, premutation, and disease state lengths of CAG TNR DNA, the propensity for expansion increases with increasing repeat number [4]. Similarly, *in vitro* experiments have shown that for repeat tracts within the healthy range, the thermodynamic stability of the non-B conformation is dependent on length [13,14]. However, to date, an *in vitro* characterization of the structures formed by TNR sequences in the premutation range has not been performed.

In this work we sought to characterize the TNR sequence (CAG)₃₆ that defines the threshold length for expansion *in vivo*. Non-B conformations were investigated by chemical probing and, surprisingly, we determined that (CAG)₃₆ adopts a simple stem-loop hairpin as opposed to a more complex folded structure. Furthermore, similar to (CAG)₁₀, we found that the loop of (CAG)₃₆ contains a hotspot for oxidative damage. Despite the similarities in structure and susceptibility to damage observed for (CAG)₁₀ and (CAG)₃₆, a difference was observed when comparing the thermodynamic parameters of the sequences. We observed a linear correlation between repeat length and thermodynamic stability of the non-B conformation, specifically, the thermodynamic stability (CAG)₃₆ is ~5-times greater than (CAG)₁₀. Here, we analyzed a premutation length allele and define *in vitro* that it forms a non-B conformation similar to the shorter repeat lengths found in the healthy allele. Since these sequences form similar non-B structures *in vitro*, we suggest that it is thermodynamic stability of the non-B conformations that is important for distinguishing TNR tracts that are stable *in vivo* and do not expand, from TNR tracts that expand to cause disease.

2. Materials and Methods

2.1. DEPC Chemical Probe Analysis

Oligonucleotides were synthesized and purified by HPLC as described previously [15] and were 5'-³²P end-labeled using T4 polynucleotide kinase. To obtain the thermodynamically favored DNA conformation(s), oligonucleotides (1 μ M) in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5 were incubated for 5 min at 95 °C and cooled to room temperature over ~2.5 h. Oligonucleotides were then incubated with 0.5–2% diethyl pyrocarbonate (DEPC) (v/v) (20 μ L final sample volume) for 30 min at 37 °C. Following incubation with DEPC, the samples were dried *in vacuo*, treated with 10% piperidine (v/v) for 30 min at 90 °C, and again dried *in vacuo*. Samples were resuspended in denaturing

loading buffer (80% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue), incubated for 3 min at 95 °C, and analyzed using a 16%, for (CAG)₁₀, or 10%, for (CAG)₃₆, denaturing polyacrylamide gel. The products were visualized by phosphorimager.

2.2. Br₂ Chemical Probe Analysis

Oligonucleotides were prepared as described above (3 μM for (CAG)₁₀, 1 μM for (CAG)₃₆) in 50 mM sodium phosphate, 10 mM NaCl, pH 7.5. Oligonucleotides were incubated with 2–4.5 mM Br₂ (20 μL final sample volume) for 30 min at 25 °C. The reaction was quenched with 0.3 M sodium acetate and 20 μg tRNA. The samples were ethanol precipitated twice, treated with 10% piperidine (v/v) for 30 min at 95 °C, and dried *in vacuo*. Samples were resuspended in denaturing loading buffer and electrophoresed as described above.

2.3. Oxidation by Peroxynitrite

Oligonucleotides (1 μM) in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5 were prepared as described above. Since the chemistry of peroxynitrite is modulated by CO₂ [16,17], all experiments were conducted in the presence of 25 mM NaHCO₃ to maintain the concentration of CO₂ in the buffer at ~1 mM, which is comparable to physiological levels. Oligonucleotides were incubated with 0.625–2.5 mM ONOO[−] (in 0.3 M NaOH) for 30 min at 37 °C. Following incubation, the samples were dried *in vacuo*, treated with 10% piperidine (v/v) for 30 min at 90 °C, and again dried *in vacuo*. Samples were resuspended in denaturing loading buffer and electrophoresed as described above.

2.4. Optical Melting Analysis

Quantification of oligonucleotides was performed at 95 °C using the ε₂₆₀ values estimated for single-stranded DNA [18]. Oligonucleotides had a final concentration of 0.5–2.0 μM and optical melting analysis was performed in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5. Prior to analysis, thermodynamically favored DNA conformation(s) were obtained as described above. Samples were subjected to a heating/cooling cycle (20 to 95 °C at 1 °C/min, held at 95 °C for 5 min, 95 to 20 °C at 1 °C/min) while the absorbance was monitored at 260 nm. The maximum of the first derivative of the absorbance versus temperature data was used to determine melting temperature (T_m). Thermodynamic parameters were extracted from melting profiles using van't Hoff analysis [19]. Student's t-test was used to determine if the average values were statistically different.

2.5. Differential Scanning Calorimetry (DSC) Melting Analysis

Nanocalorimetry was performed using a TA instruments NanoDSC III. Oligonucleotides had a final concentration of 8–50 μM and analysis was performed in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5. All samples were degassed *in vacuo* for 12 min at 25 °C before analysis. Data were obtained by monitoring the excess power required to maintain the same temperature in the sample cell and a reference cell that was filled with degassed buffer. Each cycle of the experiment (10 cycles were performed for each sample: 10 to 105 °C at 1 °C/min; 105 to 10 °C at 1 °C/min; equilibration period of 10 min before each scan) resulted in thermograms displaying excess heat capacity as a function of temperature. Each thermogram was normalized for concentration, background corrected using a buffer-buffer scan, and fit to a sigmoidal baseline correction using NanoAnalyze software. The change in enthalpy (ΔH) was calculated by integrating the area under the thermogram. The change in entropy (ΔS) was determined by plotting C_p/T versus T and integrating under the resulting curve. The change in free energy (ΔG) was calculated at 37 °C using the values for ΔH and ΔS and the equation ΔG = ΔH – TΔS.

3. Results

3.1. Structural Analysis by Chemical Probing

In order to characterize the structure adopted by (CAG)₃₆, the oligonucleotide was subjected to chemical probing using DEPC and Br₂. DEPC and Br₂ modify purines (A>>G) [20,21] and cytosines [22], respectively, and provide a measure of the solution accessibility of these nucleobases. For example, adenines in loops, bulges, and overhangs are more solution accessible than those in duplex and will show greater reactivity toward DEPC [23,24]. With respect to Br₂, unpaired cytosines will be modified preferentially over cytosines in duplex [22]. Therefore, these two chemical probes can be used to identify DNA loops, bulges, and overhangs and thus provide a molecular-level picture of the structure adopted by (CAG)₃₆. For each chemical probe, only a single region of hyper-reactivity is observed. Specifically, incubation with DEPC reveals hyper-reactivity at adenines and guanines in repeats 18 and 19 for (CAG)₃₆ (Fig. 1A, Supplementary Data). Incubation with Br₂ reveals hyper-reactivity of the cytosine in repeat 19 for (CAG)₃₆ (Fig. 1B, Supplementary Data).

In order to be able to compare the structure adopted by (CAG)₃₆ to the structure adopted by a repeat length that falls in the healthy range, we also analyzed (CAG)₁₀ and determined its reactivity to DEPC and Br₂. Hyper-reactivity at the adenines and guanines in repeats 5 and 6 was observed when (CAG)₁₀ was exposed to DEPC (Fig. 1C, Supplementary Data), and hyper-reactivity of the cytosine in repeat 6 was observed with Br₂ (Fig. 1D, Supplementary Data).

3.2. Peroxynitrite-Induced DNA Damage Patterns

Peroxynitrite is a powerful oxidizing and nitrating agent that is produced as a part of an immune response. Peroxynitrite is known to modify several types of biomolecules, including proteins, lipids, and DNA [25]. With respect to peroxynitrite modifying DNA, G is the most readily modified nucleobase [26]. In fact, much like DEPC and Br₂, peroxynitrite can serve as a chemical probe of solution-accessible guanines, and has been used previously in our laboratory [9] to show that guanines in the loop region of stem-loop hairpins are modified to a greater extent than those in the stem region, which is more duplex-like in structure. In one example, we showed that the G in the loop of a (CAG)₁₀ stem-loop hairpin is more susceptible to oxidation than guanines in the stem region. Those results were confirmed here (Fig. 2A) and will be used as a control.

Upon treatment of (CAG)₃₆ with peroxynitrite, hyper-reactivity is observed at the G in repeat 18 (Fig. 2B). The other guanines in the sequence do show reactivity to peroxynitrite, but this reactivity is minimal when compared to that of the hyper-reactive G.

3.3. Optical Melting Analysis

UV-visible optical spectrophotometry was used to obtain the optical melting profile for (CAG)₃₆. For comparison, (CAG)₁₀, (CAG)₂₀ and (CAG)₃₀ were also analyzed by optical melting (Fig. 3). Monitoring the absorbance at 260 nm as a function of temperature, a hyperchromic shift was observed for all sequences, consistent with the transition of DNA from a structured to an unstructured state. In addition, minimal hysteresis was observed between the melting and annealing profiles. Interestingly, (CAG)₁₀, the shortest sequence, displayed the highest melting temperature (T_m) and lowest ΔH of all four sequences, with a T_m ~2 °C higher and a ΔH ~8 kcal/mol lower, compared to the other sequences (Table 1). There was no significant change in the T_m between (CAG)₂₀, (CAG)₃₀, and (CAG)₃₆, despite a difference of 48 nucleotides in length between the shortest and longest sequences. These results indicate that across the range of 20 to 36 repeats, the thermal stability of the structure adopted by a CAG sequence is not influenced by the number of repeats.

Furthermore, van't Hoff analysis revealed no difference in the thermodynamic parameters (change in enthalpy, ΔH ; change in entropy, ΔS ; change in free energy, ΔG) associated with the transitions of these three sequences from a structured to an unstructured state (Table 1).

3.4. Calorimetric Melting Analysis

DSC was also used to characterize the thermal and thermodynamic properties for the transition of (CAG)₁₀, (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆ from structured to unstructured states. The resulting thermograms, in which excess heat capacity is plotted as a function of temperature, show a single transition for all sequences (Supplementary Data). As was observed by optical analysis, the melting temperatures for (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆ are similar (Table 2). This indicates that for the repeat range studied, the thermal stability of the structures adopted by CAG repeats is not measurably affected with increasing number of repeats. Also, consistent with results observed by optical analysis, (CAG)₁₀ has the highest melting temperature, with a $T_m \sim 1.5^\circ\text{C}$ higher, compared to the other three sequences.

Interestingly, and quite different from the results obtained by van't Hoff analysis of the optical melting data, calorimetric measurements show that there is a linear increase in ΔH with respect to repeat length (Table 2, Fig. 4).

4. Discussion

Prior to this work, *in vitro* structural and thermodynamic analyses of CAG repeat DNA had been performed on sequences containing a number of repeats that falls within the healthy range of HD [5,6,13,24,27,28]. Notably, repeat tracts of this length are not prone to expansion *in vivo*. Typically, researchers have used results obtained with shorter repeat sequences to draw conclusions about the behavior of longer repeat sequences that are prone to TNR expansion. These extrapolations were used to develop hypotheses for the mechanism of TNR expansion, including replication-dependent and BER-dependent TNR expansion. However, the technical challenges associated with synthesizing and purifying longer repeat oligonucleotides, and obtaining enough material to perform calorimetry has limited the experimental conformation of these predictions. Here, we have overcome the above-mentioned challenges and the result is a full *in vitro* analysis of (CAG)₃₆, which is the number of repeats that separates sequences that do not expand *in vivo* from those that expand to cause HD.

For both (CAG)₁₀ and (CAG)₃₆, the chemical probe reactivity pattern is consistent with the formation of a stem-loop hairpin with 4 nucleobases in the loop and a blunt-ended stem. Previous studies of (CAG)₁₀ have revealed that this sequence forms a stem-loop hairpin [9,29]. Surprisingly, (CAG)₃₆, the expansion-prone counterpart of (CAG)₁₀, also adopts a simple stem-loop structure. Under our experimental conditions there is no indication that (CAG)₃₆ contains multiple loops and/or bulges, contrary to what is predicted by modeling programs such as mfold [30].

Our conclusion that (CAG)₁₀ and (CAG)₃₆ each form a stem-loop hairpin is further supported by the reactions with peroxynitrite. The G that is modified by peroxynitrite, for both (CAG)₁₀ and (CAG)₃₆, is in the loop of the proposed stem-loop structure. Therefore, (CAG)₃₆ forms a stem-loop hairpin where the G in the loop represents a hot spot for oxidative DNA damage. The oxidation pattern for (CAG)₃₆ helps to support that the proposed toxic oxidation cycle [10,11,12] may be viable for this longer, expansion-prone sequence.

It is of note that for DEPC and Br₂, the difference in reactivity between the nucleobases in the loop and those in the stem is more pronounced for (CAG)₁₀, as compared to (CAG)₃₆.

One explanation for this observation is that a structure besides the stem-loop hairpin shown in Fig. 2 forms, and that in this structure nucleobases that are identified as being in the stem in Fig. 2 are actually in a loop or bulge region. However, this difference in modification between the loop and stem is negligible for reactions with peroxynitrite; the nucleobases in the loop are easily distinguished from those in the stem for both (CAG)₁₀ and (CAG)₃₆. Therefore, it is less likely that (CAG)₃₆ adopts multiple structures, but rather that chemical probing on longer sequences produces more background cleavage.

Analysis of the optical melting data reveals how the number of CAG repeats affects both thermal and thermodynamic stability. Interestingly, (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆ all have comparable values for ΔH and ΔS , and these values are larger than (CAG)₁₀. However, the ΔG for all the sequences, aside from (CAG)₃₆, are the same within error. This is due to the increase in the entropic contribution to free energy for the longer sequences, with respect to (CAG)₁₀, which is a result of both the decrease in T_m and the increased length of the sequences. Importantly, the difference in ΔH for (CAG)₁₀ and (CAG)₃₆ is consistent with the structural data, as there are more nucleobases involved in hydrogen bonding and base stacking in (CAG)₃₆ than (CAG)₁₀. Although the effect of increased ΔH is somewhat mitigated by the increase in ΔS for (CAG)₃₆, the $\Delta\Delta G$ between (CAG)₁₀ and (CAG)₃₆ is positive, indicating that (CAG)₃₆ is thermodynamically more stable than (CAG)₁₀.

Interestingly, there is an inconsistency in the values of ΔH , ΔS , and ΔG obtained by optical melting analysis and DSC. For example, both methods are in good agreement for the ΔH of (CAG)₁₀, ~36 kcal/mol, however, the ΔH values for (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆ obtained by DSC are significantly greater than those obtained from optical melting (Fig. 4). This difference is due to assumptions made during van't Hoff analysis of the optical melting data. Van't Hoff analysis assumes a two-state model in which no intermediates populate the transition from the structured stem-loop hairpin to the unstructured state [19]. When the transition from structured to unstructured is not a simple two-state transition, and intermediates populate this transition, van't Hoff analysis yields an underestimate of the total enthalpy of the transition. In contrast, DSC does not rely on a two-state assumption and instead directly measures the heat associated with the transition; thus, any intermediates involved with the melting transition are included in ΔH . In fact, the difference in ΔH for (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆ obtained by optical melting and DSC indicates that while the melting transition for (CAG)₁₀ is two-state, the melting transition for these longer sequences proceeds through thermodynamically-significant intermediates.

Another important observation from the optical melting and DSC data is a lack of correlation between repeat length and T_m . One might expect that (CAG)₃₆ would have a higher T_m than (CAG)₁₀ due to the greater number of base pairs and base stacking. However, we observe a small decrease in T_m from 10 to 20 CAG repeats followed by a plateau of the T_m . This lack of dependence of T_m on repeat length has been reported previously for TNR stem-loop hairpins, specifically for CAG, CTG and CCG repeats ranging from 4 to 30 repeats [13,14,28,31].

Our data reveal that the structure adopted by the longer (CAG)₃₆ sequence mimics that of the (CAG)₁₀ control, and that both sequences form a stem-loop hairpin. However, our calorimetric analysis revealed an important distinction. While the stem-loop hairpins have similar thermal stability, the ΔG value for (CAG)₃₆ is ~5-fold larger than (CAG)₁₀ and indicates the greater thermodynamic stability of the stem-loop formed by the premutation TNR sequence. Furthermore, the data obtained for (CAG)₁₀, (CAG)₂₀, (CAG)₃₀, and (CAG)₃₆ indicate that this increase in thermodynamic stability is linear with respect to repeat length; longer TNR tracts form more stable stem-loop structures. It has previously been established that (CAG)₃₆ is the length at which CAG repeats are prone to expansion *in*

vivo. By providing the thermodynamic data that bridges the gap between healthy and expansion-prone repeat lengths we propose that the non-B conformation formed by (CAG)₃₆ has sufficient thermodynamic stability to form and persist *in vivo* and therefore can contribute to the expansion mechanisms; repeat tracts shorter than (CAG)₃₆ do not meet this thermodynamic requirement for expansion. Thus, thermodynamic stability serves as a metric to evaluate the propensity of a TNR sequence to expand and reconciles the observation that while TNR sequences of varying lengths can form non-B conformations, only those that reach the threshold length are prone to expansion.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful for support from the National Institute of Environmental Health Sciences (ES019296).

Abbreviations

BER	base excision repair
DEPC	diethyl pyrocarbonate
DSC	differential scanning calorimetry
EDTA	ethylenediaminetetraacetic acid
HD	Huntington's disease
LP-BER	long-patch base excision repair
T_m	melting temperature
TNR	trinucleotide repeat

References

1. McMurray C. Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet.* 2010; 11:786–799. [PubMed: 20953213]
2. López Castel A, Cleary JD, Pearson CE. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol.* 2010; 11:165–170. [PubMed: 20177394]
3. Kozłowski P, De Mezer M, Krzyżosiak W. Trinucleotide repeats in human genome and exome. *Nucleic Acids Res.* 2010; 38:4027–4039. [PubMed: 20215431]
4. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell.* 1993; 72:971–983. [PubMed: 8458085]
5. Gacy A, Goellner G, Juranić N, Macura S, McMurray C. Trinucleotide repeats that expand in human disease form hairpin structures *in vitro*. *Cell.* 1995; 81:533–540. [PubMed: 7758107]
6. Pearson C, Tam M, Wang Y, Montgomery S, Dar A, Cleary J, Nichol K. Slipped-strand DNAs formed by long (CAG)_n(CTG) repeats: Slipped-out repeats and slip-out junctions. *Nucleic Acids Res.* 2002; 30:4534–4537. [PubMed: 12384601]
7. Liu G, Chen X, Bissler J, Sinden R, Leffak M. Replication-dependent instability at (CTG)_n(CAG) repeat hairpins in human cells. *Nat Chem Biol.* 2010; 6:652–659. [PubMed: 20676085]
8. Mirkin S. Expandable DNA repeats and human disease. *Nature.* 2007; 447:932–940. [PubMed: 17581576]
9. Jarem DA, Wilson NR, Delaney S. Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry.* 2009; 48:6655–6663. [PubMed: 19527055]

10. Jarem DA, Wilson NR, Schermerhorn KM, Delaney S. Incidence and persistence of 8-oxo-7, 8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. *DNA Repair*. 2011; 10:887–896. [PubMed: 21727036]
11. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*. 2007; 447:447–452. [PubMed: 17450122]
12. Liu Y, Prasad R, Beard W, Hou E, Horton J, McMurray C, Wilson S. Coordination between Polymerase β and FEN1 can modulate CAG repeat expansion. *J Biol Chem*. 2009; 284:28352–28366. [PubMed: 19674974]
13. Paiva A, Sheardy R. Influence of sequence context and length on the structure and stability of triplet repeat DNA oligomers. *Biochemistry*. 2004; 43:14218–14227. [PubMed: 15518572]
14. Amrane S, Sacca B, Mills M, Chauhan M, Klump H, Mergny J. Length-dependent energetics of (CTG) $_n$ and (CAG) $_n$ trinucleotide repeats. *Nucleic Acids Res*. 2005; 33:4065–4077. [PubMed: 16040598]
15. Jarem DA, Huckaby LV, Delaney S. AGG interruptions in (CGG) $_n$ DNA repeat tracts modulate the structure and thermodynamics of non-B conformations in vitro. *Biochemistry*. 2010; 49:6826–6837. [PubMed: 20695523]
16. Lymar SV, Hurst JK. Rapid reaction between peroxynitrite ion and carbon dioxide: implications for biological activity. *J Am Chem Soc*. 1995; 117:8867–8868.
17. Denicola A, Freeman B, Trujillo M, Radi R. Peroxynitrite reaction with carbon dioxide/bicarbonate: Kinetics and influence on peroxynitrite-mediated oxidations. *Arch Biochem Biophys*. 1996; 333:49–58. [PubMed: 8806753]
18. Warshaw MM, Tinoco I Jr. Optical properties of sixteen dinucleoside phosphates. *J Mol Biol*. 1966; 20:29–38. [PubMed: 5970660]
19. Marky LA, Breslauer KJ. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*. 1987; 26:1601–1620. [PubMed: 3663875]
20. Leonard N, McDonald J, Henderson R, Reichmann M. Reaction of diethyl pyrocarbonate with nucleic acid components. *Adenosine Biochemistry*. 1971; 10:3335–3342.
21. Vincze A, Henderson R, McDonald J, Leonard N. Reaction of diethyl pyrocarbonate with nucleic acid components. Bases and nucleosides derived from guanine, cytosine, and uracil. *J Am Chem Soc*. 1973; 95:2677–2682. [PubMed: 4694530]
22. Ross S, Burrows C. Cytosine-specific chemical probing of DNA using bromide and monoperoxysulfate. *Nucleic Acids Res*. 1996; 24:5062–5063. [PubMed: 9016685]
23. Huertas D, Bellolell L, Casasnovas J, Coll M, Azorín F. Alternating d(GA) $_n$ DNA sequences form antiparallel stranded homoduplexes stabilized by the formation of G•A base pairs. *EMBO J*. 1993; 12:4029–4038. [PubMed: 8404869]
24. Mitas M. Trinucleotide repeats associated with human disease. *Nucleic Acids Res*. 1997; 25:2245–2253. [PubMed: 9171073]
25. Pacher P, Beckman JS, Liaudet L. Nitric oxide and peroxynitrite in health and disease. *Physiol Rev*. 2007; 87:315–424. [PubMed: 17237348]
26. Niles J, Wishnok J, Tannenbaum S. Peroxynitrite-induced oxidation and nitration products of guanine and 8-oxoguanine: Structures and mechanisms of product formation. *Nitric Oxide*. 2006; 14:109–121. [PubMed: 16352449]
27. Degtyareva N, Reddish M, Sengupta B, Petty J. Structural studies of a trinucleotide repeat sequence using 2-aminopurine. *Biochemistry*. 2009; 48:2340–2346. [PubMed: 19170594]
28. Petruska J, Arnheim N, Goodman MF. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res*. 1996; 24:1992–1998. [PubMed: 8668527]
29. Ávila Figueroa A, Delaney S. Mechanistic studies of hairpin to duplex conversion for trinucleotide repeat sequences. *J Biol Chem*. 2010; 285:14648–14657. [PubMed: 20228068]
30. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003; 31:3406–3415. [PubMed: 12824337]
31. Amrane S, Mergny J. Length and pH-dependent energetics of (CCG) $_n$ and (CGG) $_n$ trinucleotide repeats. *Biochimie*. 2006; 88:1125–1134. [PubMed: 16690198]

Highlights

1. First structural and thermodynamic analysis of premutation allele of HD
2. Premutation allele of HD adopts a stem-loop non-B conformation
3. Healthy and premutation length stem-loops are hyper-susceptible to oxidative damage
4. Stability of stem-loop structures increases linearly with repeat length
5. Thermodynamic stability, not the ability to adopt non-B conformation, distinguishes DNA prone to expansion from stable DNA

Figure 1. Chemical probes show a single region of reactivity for (CAG)₁₀ and (CAG)₃₆. Histograms for the reaction of (A,B) (CAG)₃₆ and (C,D) (CAG)₁₀ with the chemical probes DEPC and Br₂, respectively.

Figure 2.

Modification of (CAG)₁₀ and (CAG)₃₆ by peroxynitrite. Autoradiograms revealing strand cleavage of (A) (CAG)₁₀ and (B) (CAG)₃₆ following exposure to peroxynitrite and a graphical representation of the structures adopted by the TNR sequences. Lanes 1, 2, and 3 are controls containing DNA alone, piperidine-treated DNA and piperidine/NaOH-treated DNA, respectively. Lanes 4, 5, and 6 contain DNA incubated for 30 min at 37 °C in the presence of peroxynitrite (0.625, 1.25, 2.5 mM, respectively), followed by piperidine treatment. A/G and C/T are Maxam-Gilbert sequencing reactions. The arrow indicates hyper-reactivity of the G in the loop. The asterisk represents the location of the ³²P-radiolabel.

Figure 3. Optical analysis of CAG repeat sequences. Melting/annealing curves for (A) (CAG)₁₀, (B) (CAG)₂₀, (C) (CAG)₃₀, and (D) (CAG)₃₆. The solid line represents data obtained while heating the sample and the dotted line represents data obtained while cooling.

Figure 4. Graphical representation of enthalpy data. Data obtained for (CAG)₁₀, (CAG)₂₀, (CAG)₃₀, and (CAG)₃₆ by UV-visible (closed circles) and DSC (open circles) are provided.

Table 1

UV-Visible-Derived Thermodynamic Parameters for (CAG)₁₀, (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆

Substrate	T _m ^a (°C)	ΔH ^{a,b} (kcal/mol)	ΔG ^{a,c} (kcal/mol)	ΔS ^{a,b} (cal/(mol•K))	ΔT _m ^a (°C)	ΔΔH (kcal/mol)	ΔΔG (kcal/mol)	ΔΔS (cal/(mol•K))
(CAG) ₁₀	61.6 ± 0.7	34.9 ± 1.4	2.6 ± 0.2	104 ± 4	-	-	-	-
(CAG) ₂₀	59.2 ± 0.5	42.5 ± 1.8	2.8 ± 0.1	128 ± 6	-2.4	+7.6	N/A ^d	+24
(CAG) ₃₀	59.3 ± 0.8	43.3 ± 2.2	2.9 ± 0.2	130 ± 7	-2.3	+8.4	N/A ^d	+26
(CAG) ₃₆	59.5 ± 0.8	44.1 ± 1.0	3.0 ± 0.1	133 ± 3	-2.1	+9.2	+0.4	+29

^aDNA in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5 (Na⁺ = 125 mM). Error represents standard deviation from a minimum of three experiments.

^bThermodynamic values derived from van't Hoff analysis as described by Marky and Breslauer.

^cValues reported at 37 °C.

^dN/A indicates values are not statistically different from (CAG)₁₀ as determined by Student's t-test.

Table 2

DSC-Derived Thermodynamic Parameters for (CAG)₁₀, (CAG)₂₀, (CAG)₃₀ and (CAG)₃₆

Substrate	T _m ^a (°C)	ΔH ^{a,b} (kcal/mol)	ΔG ^{a,c} (kcal/mol)	ΔS ^{a,b} (cal/(mol•K))	ΔT _m ^a (°C)	ΔΔH (kcal/mol)	ΔΔG (kcal/mol)	ΔΔS (cal/(mol•K))
(CAG) ₁₀	63.4 ± 0.3	37.1 ± 0.8	2.9 ± 0.1	110 ± 2	-	-	-	-
(CAG) ₂₀	61.8 ± 0.1	99.0 ± 1.5	7.3 ± 0.1	296 ± 4	-1.6	+61.9	+4.4	+186
(CAG) ₃₀	61.6 ± 0.3	150.7 ± 2.9	11.1 ± 0.3	450 ± 8	-1.8	+113.6	+8.2	+340
(CAG) ₃₆	62.3 ± 0.1	181.2 ± 4.6	13.7 ± 0.4	541 ± 14	-1.1	+144.1	+10.8	+431

^aDNA in 50 mM sodium phosphate, 10 mM NaCl, 25 mM NaHCO₃, pH 7.5 (Na⁺ = 125 mM). Error represents standard deviation from a minimum of three experiments.

^bThermodynamic values derived directly by integration of the excess heat capacity curve.

^cValues reported at 37 °C.