Published in final edited form as:

Biochimie. 2018 March ; 146: 156–165. doi:10.1016/j.biochi.2017.12.008.

Effect of dC → d(m⁵C) Substitutions on the Folding of Intramolecular Triplexes with Mixed TAT and C+GC Base Triplets

Carolyn E. Carr, Rajkumar Ganugula¹, Ronald Shikiya², Ana Maria Soto³, and Luis A. Marky
Department of Pharmaceutical Sciences, University of Nebraska Medical Center, 986025
Nebraska Medical Center, Omaha, NE 68198-6025

Abstract

Oligonucleotide-directed triple helix formation has been recognized as a potential tool for targeting genes with high specificity. Cytosine methylation in the 5' position is both ubiquitous and a stable regulatory modification, which could potentially stabilize triple helix formation. In this work, we have used a combination of calorimetric and spectroscopic techniques to study the intramolecular unfolding of four triplexes and two duplexes. We used the following triplex control sequence, named *Control Tri*, d(AGAGAC₅TCTCTC₅TCTCT), where C₅ are loops of five cytosines. From this sequence, we studied three other sequences with dC → d(m⁵C) substitutions on the Hoogsteen strand (*2MeH*), Crick strand (*2MeC*) and both strands (*4MeHC*). Calorimetric studies determined that methylation does increase the thermal and enthalpic stability, leading to an overall favorable free energy, and that this increased stability is cumulative, i.e. methylation on both the Hoogsteen and Crick strands yields the largest favorable free energy. The differential uptake of protons, counterions and water was determined. It was found that methylation increases cytosine protonation by shifting the apparent pK_a value to a higher pH; this increase in proton uptake coincides with a release of counterions during folding of the triplex, likely due to repulsion from the increased positive charge from the protonated cytosines. The immobilization of water was not affected for triplexes with methylated cytosines on their Hoogsteen or Crick strands, but was seen for the triplex where both strands are methylated. This may be due to the alignment in the major groove of the methyl groups on the cytosines with the methyl groups on the thymines which causes an increase in structural water along the spine of the triplex.

Keywords

Triplex; Duplex; Methylation; Epigenetics; Thermodynamics

*To whom correspondence should be addressed. Tel: (402) 559-4628; Fax: (402) 559-9543; lmarky@unmc.edu.

¹Current Address: Aisin Cosmos R&D Co. Ltd, Hyderabad, Telangana, India.

²Current Address: Department of Med. Microbiology & Immunology, Creighton University, Omaha, Nebraska, USA.

³Current Address: Department of Chemistry, Towson University, Towson, Maryland, USA

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.

Conflict of Interest

The authors declare no conflict of interest.

Supporting Material

Eight figures and one table are available at <http://>

1. Introduction

Triple stranded DNA, or triplexes, are a regulatory structure found within chromosomal DNA.[1, 2] Because triplexes are inherently unstable, a feature which implicates the involvement of triplexes in cancer and neurodegenerative disorders,[3–9] they can be hard to detect and verify. Because of triplex instability, they have a high rate of generating repeat sequences such as in the case of Friedreich's ataxia, which become progressively worse in successive generations.[8, 10–12] Despite their perceived instability, triplexes have been associated with a variety of roles related to genomic regulation; this role is supported by the fact that most identified intramolecular DNA triplex forming sequences occur only once in the genome, which imparts specificity to triplexes as a regulatory binding element.[3, 13] Triplexes have been shown to be involved in genomic regulation by acting as pause sites during replication,[6, 14–18] posttranslational processing through mRNA splicing,[19, 20] and chromatin organization by triplex formation through distant sequences.[21–23] In addition, it has been shown that triplexes promote methyltransferase recruitment and methylation of downstream cytosines,[11, 24–26] which alters gene expression and is important for gene imprinting and cell differentiation.[27–33] Recent work has suggested that the solution conditions of the cell, not often replicated *in vitro*, significantly stabilize Hoogsteen hydrogen bonds, leading to stable formation of triplexes.[34] This research, coupled with the identification of over a thousand sites for triplex formation by long noncoding RNAs (lncRNAs),[35, 36] suggests that triplexes may be a more widespread regulatory mechanism than originally hypothesized. In addition, significant synthetic work has been done to generate a triplex forming third strand with greater stability and pH independence for use as an antigene therapy.[37–40]

Epigenetic affects due to cytosine methylation have long been known to be a critical factor in vertebrates,[28, 33, 41, 42] and are associated with key biological processes such as genomic imprinting,[32, 33, 43, 44] X-chromosome inactivation,[45–48] silencing of transposons[33, 49–51] and cell differentiation.[45, 52, 53] These methylation-dependent processes typically occur during early development and exclusively occur at CpG dinucleotides. However, current research has shown that dynamic methylation occurs during other times and in other cell types, with a preference for CpA dinucleotides.[54, 55] This has shown to be especially prevalent in repeat sequences and transposons of fungi, which display little CpG methylation but an abundance of CpA and CpT methylation.[56, 57] The methylation of the repeat sequences is thought to silence transposons and prevent the repeat sequences from being extended. CpA methylation is also seen in methylation studies of *Drosophila*, which was previously believed to not utilize methylation for gene regulation.[58] Specifically, significant CpH methylation has been found in mature human and mouse brains cells, but not in infant brain cells.[54, 59, 60] These studies revealed that non CpG methylation accumulates appreciably through human brain development.[54] Embryonic stem cells, pluripotent stem cells and oocytes have shown significant levels of non-CpG methylation which is then lost after cell differentiation, suggesting a dynamic role in gene regulation for non-CpG methylation rather than the static role of CpG methylation.[54, 55, 61–64] Because triplexes have been linked with chromatin organization and enhanced methylation of downstream cytosines,[4, 24, 65, 66] and studies have shown dynamic

methylation at non-CpG sites, we decided to investigate the thermodynamic stability of methylated pyrimidine triplexes, with specific focus on determining the differential binding of protons, counterions, and water. The immobilization of these molecules is important for determining DNA stability and for their ability to interfere with polymerase binding; they also have implications in antigene strategies that seek to generate a methylated triplex that could hinder polymerase binding, or disrupt a triplex by forming two duplexes.

Triplexes consist of three strands, with two of these strands forming canonical Watson-Crick base-pairs with each other, and a third strand binding to the Watson strand through Hoogsteen or reverse-Hoogsteen hydrogen bonds. There are two types of intramolecular DNA triplexes, a pyrimidine or parallel triplex, and a purine or anti-parallel triplex. The pyrimidine triplex has a pyrimidine-rich Hoogsteen strand, which runs parallel to the Watson strand and binds to it through Hoogsteen base-pairing. If this third strand contains cytosines, then only one hydrogen bond is possible without protonation at its N3 position, and thus these triplexes are stabilized by low pH to facilitate protonation.[38, 67] A pyrimidine triplex may also contain only thymines and these triplexes are stabilized by mono and divalent cations. If the third strand is rich in purines then it will run anti-parallel to the Watson strand and bind using reverse-Hoogsteen hydrogen bonds; purine rich triplexes are stabilized by mono and divalent cations.[39, 68] The triplexes studied below and shown in Figure 1 are examples of pyrimidine, or parallel, triplexes, with the third strand binding to the Watson strand with Hoogsteen hydrogen bonds.

Ions and water molecules are critical for the native function of DNA, regardless of structure, but are often overlooked. As a biopolymer, DNA is negatively charged and the repulsive forces from these charges play a role in dictating the structure of DNA. Structures such as intramolecular junctions, which have a high charge density around the junction, have a strong structural requirement for cations to neutralize the negative charge repulsion.[69, 70] The same is true for triplexes, which have a third strand inserted into the major groove and bound to the Watson strand.[71–73] While triplexes may have a higher counterion uptake, the cytosine protonation necessary for pyrimidine triplex formation would exclude counterions and may have an impact on the overall stability. In addition to influencing triplex stability, any ions bound to the surface of the DNA would need to be displaced, an entropically favourable process, in order to interact with a complementary strand, such as in an antigene therapy, or for a protein to bind.

Water also plays a fundamental role in determining the secondary and tertiary structure of oligonucleotides,[74] as previous research has indicated that nucleic acids are heavily hydrated.[75, 76] The overall hydration of an oligonucleotide is dependent on its conformation, nucleic acid composition, and sequence.[77–80] The precise details and determinants of hydration have yet to be fully elucidated, especially when taking into account the different types of water bound to the surface of nucleic acids; hydrophobic or structural water, associated with polar and non-polar groups, and electrostricted water, which are immobilized by charged groups.[81, 82] These two types of water are difficult to differentiate and the results are further confused when considering the hydration sphere of the ions bound to oligonucleotides. In a similar manner to ions, water must be displaced in order for another molecule to bind to the surface of DNA. Cation and water release have a

major effect on the binding of proteins to nucleic acids due to the contribution to a favourable binding entropy.

Our lab is focused on understanding the unfolding thermodynamics of DNA structures, the physical factors that control which structures they form and the free energy of these forms. In addition, we seek to complete these thermodynamic profiles by measuring the uptake of water, protons and counterions, which are sequence and structure dependent. In this study, we used a combination of spectroscopic and calorimetric techniques to investigate the unfolding behaviour of a pyrimidine triplex (*Control Tri*) at pH 5.2 and 6.2, and compared its behaviour with that of a triplex containing 5-methylcytosines on the Hoogsteen strand (*2MeH*), the Crick strand (*2MeC*), and both strands (*4MeHC*). Our results suggest that methylation increases the thermal stability and free energy of formation of the duplex domain, but not the triplex. Methylation increases the pH range at which the cytosines become protonated; a higher amount of protonation causes a greater exclusion of counterions due to the positively charged cytosines. Methylation did not appear to increase water immobilization to any significant extent, with *2MeH* and *2MeC* showing the same amount of immobilization as *Control Tri*. However, *4MeHC* did have higher water immobilization, although the reason for this is unclear.

2. Materials and Methods

2.1 Materials

All oligomers were synthesized and HPLC purified by the Core Synthetic Facility of the Eppley Research Institute at the University of Nebraska Medical Center. The oligomers were further desalted by gel permeation chromatography using a G-10 Sephadex column. The concentration of each oligomer shown in Figure 1 was measured at 260 nm at 80°C using the molar extinction coefficients reported below. Oligomer sequences, their designations, and extinction coefficients (in $\text{mM}^{-1}\text{cm}^{-1}$ of strands): d(AGAGAC₅TCTCTC₅TCTCT), *Control Tri*, 210; d(AGAGAC₅TCTCTC₅T^mCT^mCT), *2MeH*, 210; d(AGAGAC₅T^mCT^mCTC₅TCTCT), *2MeC*, 213; d(AGAGAC₅T^mCT^mCTC₅T^mCT^mCT), *4MeHC*, 213; d(AGAGAC₅TCTCTCC), *Control Du*, 150; d(AGAGAC₅T^mCT^mCTCC), *2MeC-Du*, 150. All values were calculated by extrapolating the tabulated values of the monomers and dimers from 25 °C to high temperatures, using procedures reported earlier. [83, 84] The buffer solutions consisted of 10 mM sodium cacodylate or 10 mM sodium phosphate, adjusted to different pH, salt, and water activity with HCl, NaCl and ethylene glycol, respectively. All chemicals used in this study were reagent grade and used without further purification.

2.2 Temperature-Dependent UV Spectroscopy

Absorbance versus temperature profiles for all oligonucleotides were obtained with a thermoelectrically controlled Aviv 14-DS Spectrophotometer (Lakewood, NJ). The absorbance of the oligonucleotides was recorded continuously at 260 nm, while the temperature was increased from 0 to 100°C at a heating rate of 0.6 °C min⁻¹. UV melting curves allow measurement of the midpoint of the order-disorder transition, or T_M , and the shape of the melting curves allows determination of the model dependent van't Hoff

enthalpy, H_{vH} . The transition temperatures, T_{M} s, were determined by using the Lorentz model in Origin to fit the first derivative of the melting curves; the peak maxima obtained was reported as the mid-point temperature of the triplex-duplex and duplex-coil transitions. Melting curves were done as a function of total strand concentration in order to determine the molecularity of the transition(s) of each oligonucleotide. If the T_{M} is constant across all concentrations then the transition is intramolecular or monomolecular. A T_{M} dependence on strand concentration is indicative of higher order molecularities. Additional UV melting curves were obtained to measure the model-independent release/uptake of protons, n_{H^+} , counterions, n_{Na^+} , and water molecules, n_{W} .

2.3 Circular Dichroism (CD)

An Aviv stopped flow circular dichroism spectrometer model 202SF (Lakewood, NJ) was used to evaluate the overall conformation of the DNA triplexes and control duplexes. The spectra were recorded from 200 to 320 nm at 20 °C in 1 nm increments in a 1 cm pathlength quartz cell using solutions with an absorbance of 1.0. Each spectrum was recorded after 20 minutes of pre-equilibration time and the spectra shown correspond to an average of at least two scans.

2.4 Differential Scanning Calorimetry (DSC)

The calorimetric measurements were performed with a VP-DSC from Malvern MicroCal (Northampton, MA). This calorimeter consists of two cells; a sample cell, containing an oligomer solution (0.6 mL) and the reference cell filled with the same volume of buffer. Both cells were equilibrated at 1 °C for 10 minutes before being heated adiabatically from 1 to 100 °C at a heating rate of 0.75 °C min⁻¹. The heat capacity, C_p , was measured as a function of temperature during the unfolding process of each oligonucleotide. The corresponding buffer-buffer baselines were subtracted from the C_p versus temperature curves and the resultant melting curves were deconvoluted using a non two-state zero C_p model and analyzed to obtain the thermal midpoint, T_{M} , calorimetric enthalpy, H_{cal} , calorimetric entropy, S_{cal} , and van't Hoff enthalpy, H_{vH} . The H_{cal} and S_{cal} were determined by integrating the normalized area under a transition peak, ($H_{\text{cal}} = \int C_p dT$; $S_{\text{cal}} = \int (C_p/T) dT$), while the H_{vH} is determined by the shape of the transition peak.[85] The folding free energy at 5 °C, $G^\circ_{(5)}$, is calculated from the Gibbs equation, $G^\circ_{(5)} = H_{\text{cal}} - T S_{\text{cal}}$.

Each experiment is composed of at least three scans with a pre-equilibration time of 20 to 40 minutes for successive scans. In each experiment at pH 5.2, we observed that the scans of the modified triplexes with 5-methylcytosines were not superimposable due to oligonucleotide degradation from a combination of low pH and high temperature. However, the first scans of the experiments performed each time with purified oligomers are superimposable. The data reported at this pH are the average of at least three first scans of three individual experiments. At pH 6.2 or 7.0, all scans of each experiment are superimposable and the data reported are the average of at least three scans.

2.5 Differential Binding of Protons, Sodium Ions and Water Molecules

The following equations were used to measure the thermodynamic uptake of protons, n_{H^+} , counterions, n_{Na^+} , and water molecules, n_{W} , upon folding of each triplex and control duplexes: [86, 87]

$$\Delta n_{\text{H}^+} = 0.434 \left[\Delta H_{\text{cal}} / RT_{\text{M}}^2 \right] (\partial T_{\text{M}} / \partial \text{pH}) \quad (1)$$

$$\Delta n_{\text{Na}^+} = -1.11 \left[\Delta H_{\text{cal}} / RT_{\text{M}}^2 \right] (\partial T_{\text{M}} / \partial \ln [\text{Na}^+]) \quad (2)$$

$$\Delta n_{\text{W}} = - \left[\Delta H_{\text{cal}} / RT_{\text{M}}^2 \right] (\partial T_{\text{M}} / \partial \ln [a_{\text{W}}]) \quad (3)$$

where 0.434 and 1.11 are correction factors that correspond to conversion of decimal logarithms into natural logarithms and concentrations into ionic activities, respectively. The $[\Delta H_{\text{cal}} / RT_{\text{M}}^2]$ term is a constant that is determined from DSC experiments where the enthalpy is model independent and R is the gas constant. The values in parenthesis are determined from UV melting curves by measuring the T_{M} dependencies on the concentration of protons (by varying the pH), counterions (by varying the salt concentration), and water (by varying the osmolyte, ethylene glycol, concentration).

In determining n_{Na^+} , the UV experiments were carried out with a [NaCl] from 10 mM to 200 mM at pH 5.2. For n_{H^+} , the UV melting curves were obtained from a pH range of 5.2 to 7.0 with a concentration of 10 mM NaCl. For n_{W} the UV melting curves were carried out at pH 5.2 in 10 mM NaCl. The activity of water was varied by increasing the concentration of ethylene glycol, an osmolyte that does not interact with DNA, from 0.5 – 2.5 M.[88] The osmolality of the ethylene glycol solutions was measured with a Model 830 UIC vapor pressure osmometer which was calibrated with standardized NaCl solutions. The osmolalities were converted into water activity using the equation $\ln a_{\text{W}} = -\text{Osm}/M_{\text{W}}$, [89] where Osm is the measured solution osmolality and M_{W} is the molality of pure water. Measurement of the differential binding of water using the osmotic stress method yields only the immobilization of structural, not electrostricted, water.

3. Results and Discussion

3.1 Experimental Design

All oligonucleotides studied were designed to form an intramolecular triplex with a common triple helical domain consisting of three TAT and two C⁺GC base triplets (Figure 1). The cytosines in both the Crick and/or Hoogsteen strand were methylated, as cytosines are the most abundant methylated nucleic acid, with nearly all organism utilizing cytosine methylation as a regulatory mechanism. In naming these molecules, *Control Tri* refers to the triplex with no modifications, *2MeH* refers to the replacement of two cytosines with

3.4 Methylation increases the stability of the duplex

The DSC thermograms of the triplexes at pH 5.2 and 6.2 are shown in Figure 3 while the thermograms of the duplex controls are shown in Figure S4. Because the triplexes studied contain cytosines in the Hoogsteen strand they should be stabilized at low pH; this can be seen in Table 1, which list the thermodynamic parameters obtained at pH 5.2 and 6.2 for all oligonucleotides studied. The change in the thermodynamic parameters of the methylated triplexes compared to *Control Tri* is shown in Table 2. In all cases the triplexes are significantly thermally destabilized at pH 6.2 compared to pH 5.2.

Control Tri is monophasic at both pHs and gains significant thermal and enthalpic stability as the pH is decreased. At pH 6.2 *Control Tri* has a T_M of 38.5 °C and a H_{cal} of 87.6 kcal/mol. While this is greater than the predicted nearest-neighbor enthalpy of the duplex (40.4 kcal/mol), indicating that the triplex is forming, it is less than what would be expected for the addition of a third strand (100.5 kcal/mol),[72] indicating poor base-triplet stacking or lack of base-triplet stacks. At pH 5.2 *Control Tri* has a T_M of 56.7 °C and a H_{cal} of 93.7, indicating improved triplex formation although it is still below the expect value for complete formation of the triplex. This is an increase of 18.2 °C and –6.3 kcal/mol due to protonation of the cytosines in the Hoogsteen strand.

In *2MeH* the two cytosines on the Hoogsteen strand are methylated and, similar to *Control Tri*, it gains significant thermal stability when lowering the pH. At pH 6.2, *2MeH* has two transitions at 35.2 and 49.1 °C with H_{cal} s of 50.4 and 45.2 kcal/mol, respectively. The increased thermal stability suggests significant effects of methylation on the unfolding of triplexes, consistent with previous literature on methylated triplexes.[40, 92–94] However, in previous studies intermolecular triplexes with a methylated Hoogsteen strand were studied and only the triplex was stabilized. For intramolecular triplexes, methylation uncouples the simultaneous unfolding of the triplex and duplex seen in *Control Tri* with the duplex domain of *2MeH* unfolding at a significantly higher temperature (T_M of 10.6 °C) than that of *Control Tri*, and with a higher enthalpy, suggesting that methylation of the Hoogsteen strand affects the stability of the unmethylated duplex; this increase in stability is not seen in the unfolding of the triplex. The total enthalpy (95.6 kcal/mol) is higher than *Control Tri* at pH 6.2, indicating that methylation facilitates base-triplet stacking or that the methyl groups engage in enthalpically favorable interactions with the surrounding nucleotides, as previously described.[93, 95, 96] At pH 5.2 *2MeH* is monophasic, with a T_M of 61.6 °C and a H_{cal} of 91.4 kcal/mol. While the total enthalpy is not increased (H_{cal} –2.5 kcal/mol, well within experimental error), the thermal stability of the duplex and triplex is increased (T_M of 4.9 °C). This suggests that at physiological pH, without partial protonation of the cytosines, methylation of the Hoogsteen strand stabilizes duplex formation, but at low pH when the cytosines are protonated methylation stabilizes both duplex and triplex formation.

In *2MeC* the two cytosines in the Crick strand are methylated. At pH 6.2 *2MeC* looks similar to *2MeH*, with two transitions at 37.9 and 53.5 °C and H_{cal} s of 46.4 and 56.0 kcal/mol, respectively. The triplex and duplex are uncoupled, similar to *2MeH*, with the duplex unfolding at a much higher temperature than in *Control Tri* or even *2MeC-Dup*, indicating that methylation of cytosines causes stabilization of the duplex regardless of where the methylation occurs. The enthalpy of the duplex folding transition (53.5 kcal/mol) is higher

than expected based on *2MeC-Dup*, indicating a greater number of base-pair stacks occurring, or perhaps contributions from the methylated cytosines, consistent with previous reports of hydrophobic effects from the methyl groups that increase the enthalpy of formation.[95, 96] The total enthalpy is higher than what is seen for *Control Tri* at pH 6.2 (H_{cal} of 14.8 kcal/mol), indicating that there is complete formation of all possible base-triplet stacks and the triplex is fully formed. At pH 5.2 *2MeC* is biphasic, with the first transition corresponding to the unfolding of the triplex with a T_M of 57.1 °C, identical to the unfolding of the triplex in *Control Tri*. *2MeC* is methylated on the Crick strand and the duplex portion unfolds at 66.7 °C, 9.6 °C higher than the triplex. This is also higher than the formation of the duplex domain of *2MeH* (61.6 °C), indicating increased stability of the duplex domain when the Crick strand is methylated instead of the Hoogsteen strand.

4MeHC has four methylated cytosines, two on the Hoogsteen strand and two on the Crick strand. As methylation increased the stability of both *2MeH* and *2MeC*, it is expected that *4MeHC* will have even more stability. This is seen to be true at pH 6.2, where *4MeHC* has two transitions, at 35.6 and 57.9 °C. Higher thermal stability translates to increased enthalpy, as *4MeHC* has a slightly higher total enthalpy than *2MeC* (−111.0 kcal/mol). At this pH, the effects of methylation on the thermodynamics of triplex unfolding are additive, in that the difference between *2MeH/2MeC* and *Control Tri* add together to give the thermodynamic difference between *4MeHC* and *Control Tri*. At pH 5.2 the total enthalpy remains the same and the unfolding of the triplex at 55.3 °C has gained no stability compared to the same transition in *Control Tri*, but the duplex domain has the highest T_M of all the oligonucleotides in this study, including *2MeC*, which is also methylated on the Crick strand, indicating that stability from methylation is additive, consistent with the stability being from hydrophobic interactions of the methyl groups.[95, 96]

In summary, substitution of cytosines with methylated cytosines increases the thermal stability and improves base-triplet stacking, leading to a more favorable folding reaction. The order of thermal and enthalpic stability follow the same trend and are as follows: *4MeCH* > *2MeC* > *2MeH* > *Control Tri*.

Standard thermodynamic profiles, including entropy and free energy, for the formation of each molecule are summarized in Table 1 and comparison of the methylated triplexes with *Control Tri* are in Table 2. The $G^\circ_{(5)}$ and $T^\circ S$ terms are estimated at 5 °C, where all the molecules are in the helical state. Inspection of Table 1 indicates that the folding of each molecule is accompanied by a favorable free energy term, which results from the characteristic compensation of favorable enthalpy and unfavorable entropy contributions. [97] The favorable enthalpy term is caused by the formation of base-pairs and base-pair stacks, while the unfavorable entropy term is due to the order of the strands, which is minimized for intramolecular complexes, and the immobilization of water, protons and counterions. In the case of the methylated triplexes, the total free energy and entropy of folding follow the same trend as the enthalpy: *4MeCH* > *2MeC* > *2MeH* > *Control Tri*. This is primarily due to the increased thermal stability, related to the entropy, with slight contributions from the higher enthalpy term, although the overall difference in enthalpy of *4MeHC*, which has the highest $G^\circ_{(5)}$, compared to *Control Tri* is modest.

A comparison between the thermodynamic profiles of the methylated triplexes with *Control Tri* allowed us to determine the thermodynamic contribution of a single methylated ATA/CGC⁺ base-triplet stack (Table 3). Methylation on the Hoogsteen strand decreased the enthalpy and entropy but increases the free energy due to the increase in thermal stability. Methylation on the Crick and both Hoogsteen and Crick strands increases all parameters.

3.5 Methylation increases the uptake of protons

The differential binding of protons was calculated using Equation 1. The dependence of T_M on pH was determined from UV melting curves and are shown in Figure S5. The H_{cal}/RT_M^2 term was determined from DSC experiments at several pHs and the n_{H^+} values are shown in Table 4 and displayed in Figure 4. The folding of all triplexes is accompanied by an uptake of protons. *Control Tri* has the smallest uptake of protons among the studied triplexes, which suggests that methylation enhances protonation of the cytosines, whether it be in the stem or the loops. The increased uptake of protons for the methylated triplexes is attributed to enhanced protonation caused by a shift of the pK_a of the cytosines toward a more neutral pH; how methylation is achieving this is unclear. However, previous studies have shown that methylation alters the pK_a by up to at least one pH unit, allowing for protonation at a more physiological pH.[93, 94] The proton uptake of the control molecules does highlight the fact that triplex formation itself alters the amount of protonation and likely the pK_a of the cytosines, as unless the electronic properties of the cytosines are altered no additional protonation would occur compared to the duplexes. Methylation increases proton uptake although the location of the methylation does not appear to matter, since the uptake of protons for all methylated triplexes is within experimental error of each other. In addition, the proton uptake for the first transition (white bars) and second transition (black bars) are all similar, again indicating that position is unimportant regarding protonation. The large increase in protonation of the methylated triplexes compared to *Control Tri* is possibly due to a shift in the apparent pK_a of both the Hoogsteen and Crick strand cytosines toward a more basic pH, which would lead to ~100% protonation of the cytosines on both strands, as well as the surrounding cytosines in the loop, leading to an overall more protonated triplex. This may also explain the overall exclusion of counterions by methylated triplexes, specifically in the duplex domain which would normally not be protonated (see below).

3.6 Methylation leads to an overall exclusion of sodium ions upon formation of the triplex

The differential binding of sodium ions was calculated using Equation 2. The T_M dependence on salt concentration obtained from UV melting curves is shown in Fig S6. The H_{cal}/RT_M^2 term was obtained from DSC thermograms at several salt concentrations at pH 5.2. The resultant n_{Na^+} values are summarized in Table 4 and Figure 5. Typically, the unfolding of an oligonucleotide is associated with an uptake of counterions; this can be seen for the control molecules at pH 7.0 in Table 3, where both *Control Dup* and *2MeC-Dup* have an associated uptake of counterions during unfolding. At pH 5.2, *Control Tri* also has an overall uptake of counterions during folding. However, the uptake is very small compared to duplex DNA[87, 98] and triplexes consisting of TAT base-triplets[72, 73] but is consistent with the values from triplexes containing C⁺GC base-triplets.[73, 99] This is attributed to the presence of protonated cytosines, which excludes counterions from the phosphate backbone. For all methylated triplexes at pH 5.2 as the salt concentration increases (Table S1) there are

two transitions, the first corresponding to the folding of the triplex and the second the folding of the duplex. *2MeH* has only one peak at pH 5.2 in the absence of salt; in the presence of salt the two transitions are stabilized to different amounts, leading to separation of the folding events (Table S1). In all cases the first transition, corresponding to the triplex, has an overall repulsion of counterions; i.e. as the duplex folds into a triplex, counterions are expelled from the structure. This may be explained by the increase in protonation caused by methylation, as a decreased counterion uptake would be caused by repulsion of the positive counterions by the positive charge of the cytosines. The second transition, corresponding to the duplex, has an overall uptake of counterions as expected, although the values are small compared to typical double stranded DNA. This may be due to the relatively close or simultaneous folding of the triplex soon after the duplex at pH 5.2 for all oligonucleotides. Another possibility is that, as mentioned above, methylation shifts the pK_a of both the Hoogsteen cytosines as well as those in the Crick strand and loop regions to a more physiological pH, leading to the greater protonation observed and which would cause greater counterion exclusion. Taken together, at low pH with full triplex formation, methylated triplexes repel, rather than uptake, counterions.

3.7 Methylation does not affect immobilization of water molecules

The differential binding of water was calculated using Equation 3. The dependence of T_M on pH was determined from UV melting curves and are shown in Figure S7 for the duplexes and in Figure S8 for the triplexes. The H_{cal}/RT_M^2 term was determined from DSC experiments at pH 5.2 and the tabulated values are in Table 4 and displayed in Figure 6. The folding of all triplexes is accompanied by an immobilization of water molecules. The grey bars in Figure 6 represent total water immobilization; *Control Tri*, *2MeH* and *2MeC* have identical total water immobilization, in contrast to the results from proton and counterion uptake. It can be seen from the n_W values of *Control Dup* and *2MeC-Dup* that methylation does increase the amount of water immobilized. However, the values for the control duplexes are much smaller than *Control Tri*, indicating that DNA triplexes have a much higher amount of water immobilization than DNA duplexes, in contrast to what was observed for counterion uptake. Surprisingly, despite the results for the other methylated triplexes, *4MeHC* has increased water immobilization. It can be seen in Figure 6 that the second transition (black bars) of *4MeHC* has increased water immobilization compared to the second transition of *2MeC*; this transition corresponds to the duplex and thus if methylation of the Crick strand, which is the methylated strand of the duplex, was responsible for the increased water uptake then *2MeC* should also have a higher n_W value. Since *4MeHC* is the only triplex with increased water immobilization, it can be inferred that methylation of both the Crick and Hoogsteen strand act in conjunction to increase the amount of water immobilized. This may be due to a double chain of methyl groups from the cytosines and thymines in the major groove from both the Crick and Hoogsteen strands which increases the amount of hydration around the triplex. Previous research suggested that the increased hydration from this chain can be disrupted by a single base substitution, and so it can only be seen in *4MeHC*. [100]

4. Conclusions

We report here a complete thermodynamic description of the unfolding of an intramolecular triplex containing mixed TAT/C⁺GC base-triplets and the differential binding of protons, counterions and water. Substitution of the cytosines on the Hoogsteen strand with methylated cytosines (*2MeH*) increased the thermal stability of both the triplex and duplex domain. Methylation of the cytosines on the Crick strand (*2MeC*) and on both the Hoogsteen and Crick strands (*4MeHC*) increase the stability of the duplex only. Thus, methylation of the cytosines only on the Hoogsteen strand would lead to stabilization of the triple, but methylation on any other strand or a combination of strands leads to stabilization of the duplex only. This indicates that methylation must be precise in order to prevent genetic mutation caused by strand slippage due to instability of triplexes and their repeat sequences. This is in contrast with previous reports of intermolecular triplexes, where methylation of the Hoogsteen strand caused an increase in stability of triplex formation but did not affect duplex formation.[93–95] The stabilization of the duplex by methylation of cytosines may be a feature specific to intramolecular pyrimidine triplexes.

This study also investigates the binding of protons, counterions and water; DNA is highly hydrated, and binding of counterions is necessary to counteract the negative phosphate backbone, which allows folding of DNA into various secondary and tertiary structures. In addition, pyrimidine triplexes require protonation of the cytosines in order to stabilize the triplex structure. Our studies show that methylation increases the uptake of protons, regardless of methylation position. However, this uptake is dependent solely on the presence of methylation, and not on the position or cumulative amount, i.e. the uptake of protons is same regardless of whether the Hoogsteen, Crick, or both strands are methylated. In essence, methylation increases the apparent pK_a of cytosine to a more physiological pH, leading to higher protonation amounts and may be partially responsible for the increased stability of the duplexes, although it is not solely responsible since *4MeHC* is the most stable but does not have significantly higher proton uptake. The increased stability is likely attributed to the hydrophobic effect of the methyl groups that was seen in studies of other methylated triplexes.[92, 94, 95] In conjunction with a higher amount of protonation, methylated triplexes release counterions when folding, a result which is in opposition to duplex DNA which uptakes counterions due to the higher charge density. Although triplexes typically have a much lower uptake of counterions than duplex DNA, release of counterions was only seen for triplexes of mixed bases caused by counterion exclusion from the backbone due to protonation, a feature which is enhanced when the cytosines are methylated. Thus, methylation increases the pK_a of cytosines, increasing the protonation of the cytosines in the third strand at low pH which leads to increased repulsion of the counterions. Interestingly, counterions typically stabilize DNA; in the case of methylated triplexes, the release of counterions should destabilize the triplex but this is counteracted by the increased stability caused by protonation and methylation. Finally, methylation does increase water uptake, but only when there the cytosines on both the Hoogsteen and Crick strands are methylated. The reason for this is unclear, but may possibly due to the increase in structural water that forms when there is a continuous line of methyl groups within the major groove formed by the thymines and methylated cytosines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by Grant MCB-1122029 and Grant GM42223 (LAM) from the National Institutes of Health.

References

1. Diviacco S, Rapozzi V, Xodo L, Hélène C, Quadrifoglio F, Giovannangeli C. Site-directed inhibition of DNA replication by triple helix formation. *FASEB J*. 2001; 15:2660–2668. [PubMed: 11726542]
2. Duval-Valentin G, Thuong NT, Hélène C. Specific inhibition of transcription by triple helix-forming oligonucleotides. *Proc Natl Acad Sci U S A*. 1992; 89:504–508. [PubMed: 1731320]
3. Bacolla A, Collins JR, Gold B, Chuzhanova N, Yi M, Stephens RM, Stefanov S, Olsh A, Jakupciak JP, Dean M, Lempicki RA, Cooper DN, Wells RD. Long homopurine*homopyrimidine sequences are characteristic of genes expressed in brain and the pseudoautosomal region. *Nucleic Acids Res*. 2006; 34:2663–2675. [PubMed: 16714445]
4. Jain A, Wang G, Vasquez KM. DNA triple helices: biological consequences and therapeutic potential. *Biochimie*. 2008; 90:1117–1130. [PubMed: 18331847]
5. Javadekar SM, Raghavan SC. Snaps and mends: DNA breaks and chromosomal translocations. *FEBS J*. 2015; 282:2627–2645. [PubMed: 25913527]
6. Liu G, Myers S, Chen X, Bissler JJ, Sinden RR, Leffak M. Replication fork stalling and checkpoint activation by a PKD1 locus mirror repeat polypurine-polypyrimidine (Pu-Py) tract. *J Biol Chem*. 2012; 287:33412–33423. [PubMed: 22872635]
7. Raghavan SC, Chastain P, Lee JS, Hegde BG, Houston S, Langen R, Hsieh CL, Haworth IS, Lieber MR. Evidence for a triplex DNA conformation at the bcl-2 major breakpoint region of the t(14;18) translocation. *J Biol Chem*. 2005; 280:22749–22760. [PubMed: 15840562]
8. Rajeswari MR. DNA triplex structures in neurodegenerative disorder, Friedreich's ataxia. *J Biosci*. 2012; 37:519–532. [PubMed: 22750988]
9. Wells, RD., Warren, ST., Sarmiento, M. *Genetic Instabilities and Hereditary Neurological Diseases*. Academic Press; San Diego: 1998.
10. Bacolla A, Wells RD. Non-B DNA conformations as determinants of mutagenesis and human disease. *Mol Carcinog*. 2009; 48:273–285. [PubMed: 19306308]
11. Ruan H, Wang YH. Friedreich's ataxia GAA.TTC duplex and GAA.GAA.TTC triplex structures exclude nucleosome assembly. *J Mol Biol*. 2008; 383:292–300. [PubMed: 18775729]
12. Gerhardt J, Bhalla AD, Butler JS, Puckett JW, Dervan PB, Rosenwaks Z, Napierala M. Stalled DNA Replication Forks at the Endogenous GAA Repeats Drive Repeat Expansion in Friedreich's Ataxia Cells. *Cell Rep*. 2016; 16:1218–1227. [PubMed: 27425605]
13. Wu Q, Gaddis SS, MacLeod MC, Walborg EF, Thames HD, DiGiovanni J, Vasquez KM. High-affinity triplex-forming oligonucleotide target sequences in mammalian genomes. *Mol Carcinog*. 2007; 46:15–23. [PubMed: 17013831]
14. Hoynes PR, Maher LJ. Functional studies of potential intrastrand triplex elements in the Escherichia coli genome. *J Mol Biol*. 2002; 318:373–386. [PubMed: 12051844]
15. Hile SE, Eckert KA. Positive correlation between DNA polymerase α -Primase pausing and mutagenesis within polypyrimidine/polypurine microsatellite sequences. *J Mol Biol*. 2004; 335:745–759. [PubMed: 14687571]
16. Krasilnikova MM, Mirkin SM. Replication stalling at Friedreich's Ataxia (GAA)_n repeats in vivo. *Mol Cell Biol*. 2004; 24:2286–2295. [PubMed: 14993268]
17. Patel HP, Lu L, Blaszkak RT, Bissler JJ. PKD1 intron 21: triplex DNA formation and effect on replication. *Nucleic Acids Res*. 2004; 32:1460–1468. [PubMed: 14990751]

18. Hacia JG, Dervan PB, Wold BJ. Inhibition of Klenow fragment DNA polymerase on double-helical templates by oligonucleotide-directed triple-helix formation. *Biochemistry*. 1994; 33:6192–6200. [PubMed: 8193133]
19. Ohno M, Fukagawa T, Lee JS, Ikemura T. Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma*. 2002; 111:201–213. [PubMed: 12355210]
20. Vuyisich M, Beal PA. Regulation of the RNA-dependent protein kinase by triple helix formation. *Nucleic Acids Res*. 2000; 28:2369–2374. [PubMed: 10871369]
21. Hampel KJ, Burkholder GD, Lee JS. Plasmid dimerization mediated by triplex formation between polypyrimidine-polypurine repeats. *Biochemistry*. 1993; 32:1072–1077. [PubMed: 8424937]
22. Westin L, Blomquist P, Milligan JF, Wrangé Ö. Triple helix DNA alters nucleosomal histone-DNA interactions and acts as a nucleosome barrier. *Nucleic Acids Res*. 1995; 23:2184–2191. [PubMed: 7610046]
23. Goobes R, Cohen O, Minsky A. Unique condensation patterns of triplex DNA: physical aspects and physiological implications. *Nucleic Acids Res*. 2002; 30:2154–2161. [PubMed: 12000835]
24. Schmitz KM, Mayer C, Postepska A, Grummt I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev*. 2010; 24:2264–2269. [PubMed: 20952535]
25. Han YJ, de Lanerolle P. Naturally extended CT . AG repeats increase H-DNA structures and promoter activity in the smooth muscle myosin light chain kinase gene. *Mol Cell Biol*. 2008; 28:863–872. [PubMed: 17991897]
26. Soragni E, Herman D, Dent SY, Gottesfeld JM, Wells RD, Napierala M. Long intronic GAA*TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia. *Nucleic Acids Res*. 2008; 36:6056–6065. [PubMed: 18820300]
27. Vaillant I, Paszkowski J. Role of histone and DNA methylation in gene regulation. *Curr Opin Plant Biol*. 2007; 10:528–533. [PubMed: 17692561]
28. Ambrosi C, Manzo M, Baubec T. Dynamics and Context-Dependent Roles of DNA Methylation. *J Mol Biol*. 2017; 429:1459–1475. [PubMed: 28214512]
29. Edwards JR, Yarychivska O, Boulard M, Bestor TH. DNA methylation and DNA methyltransferases. *Epigenetics Chromatin*. 2017; 10:23. [PubMed: 28503201]
30. Kennedy AJ, Sweatt JD. Drugging the methylome: DNA methylation and memory. *Crit Rev Biochem Mol Biol*. 2016; 51:185–194. [PubMed: 26915423]
31. Koukoura O, Sifakis S, Spandidos DA. DNA methylation in endometriosis. *Mol Med Rep*. 2016; 13:2939–2948. [PubMed: 26934855]
32. Leenen FA, Muller CP, Turner JD. DNA methylation: conducting the orchestra from exposure to phenotype? *Clin Epigenetics*. 2016; 8:92. [PubMed: 27602172]
33. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*. 2007; 447:425–432. [PubMed: 17522676]
34. Tateishi-Karimata H, Nakano M, Sugimoto N. Comparable stability of Hoogsteen and Watson-Crick base pairs in ionic liquid choline dihydrogen phosphate. *Sci Rep*. 2014; 4:3593. [PubMed: 24399194]
35. Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, Mitra S, Mohammed A, James AR, Hoberg E, Moustakas A, Gyllenstein U, Jones SJ, Gustafsson CM, Sims AH, Westerlund F, Gorab E, Kanduri C. MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. *Nat Commun*. 2015; 6:7743. [PubMed: 26205790]
36. O'Leary VB, Smida J, Buske FA, Carrascosa LG, Azimzadeh O, Maugg D, Hain S, Tapio S, Heidenreich W, Kerr J, Trau M, Ovsepian SV, Atkinson MJ. PARTICLE triplexes cluster in the tumor suppressor WWOX and may extend throughout the human genome. *Sci Rep*. 2017; 7:7163. [PubMed: 28769061]
37. Doluca O, Withers JM, Filichev VV. Molecular engineering of guanine-rich sequences: Z-DNA, DNA triplexes, and G-quadruplexes. *Chem Rev*. 2013; 113:3044–3083. [PubMed: 23391174]
38. Fox KR, Brown T. Formation of stable DNA triplexes. *Biochem Soc Trans*. 2011; 39:629–634. [PubMed: 21428952]

39. Li Y, Syed J, Sugiyama H. RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem Biol.* 2016; 23:1325–1333. [PubMed: 27773629]
40. Plum GE, Park YW, Singleton SF, Dervan PB, Breslauer KJ. Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: A spectroscopic and calorimetric study. *Proc Natl Acad Sci U S A.* 1990; 87:9436–9440. [PubMed: 2251285]
41. Schubeler D. Function and information content of DNA methylation. *Nature.* 2015; 517:321–326. [PubMed: 25592537]
42. Tillo D, Mukherjee S, Vinson C. Inheritance of Cytosine Methylation. *J Cell Physiol.* 2016; 231:2346–2352. [PubMed: 26910768]
43. Elhamamsy AR. Role of DNA methylation in imprinting disorders: an updated review. *J Assist Reprod Genet.* 2017; 34:549–562. [PubMed: 28281142]
44. Satyaki PR, Gehring M. DNA methylation and imprinting in plants: machinery and mechanisms. *Crit Rev Biochem Mol Biol.* 2017; 52:163–175. [PubMed: 28118754]
45. Adhikari S, Curtis PD. DNA methyltransferases and epigenetic regulation in bacteria. *FEMS Microbiol Rev.* 2016; 40:575–591. [PubMed: 27476077]
46. Chaligne R, Heard E. X-chromosome inactivation in development and cancer. *FEBS Lett.* 2014; 588:2514–2522. [PubMed: 24937141]
47. Hoffman RM. Is DNA methylation the new guardian of the genome? *Mol Cytogenet.* 2017; 10:11. [PubMed: 28396696]
48. Peeters SB, Cotton AM, Brown CJ. Variable escape from X-chromosome inactivation: identifying factors that tip the scales towards expression. *Bioessays.* 2014; 36:746–756. [PubMed: 24913292]
49. Lauria M, Rossi V. Epigenetic control of gene regulation in plants. *Biochim Biophys Acta.* 2011; 1809:369–378. [PubMed: 21414429]
50. Saze H, Kakutani T. Differentiation of epigenetic modifications between transposons and genes. *Curr Opin Plant Biol.* 2011; 14:81–87. [PubMed: 20869294]
51. Saze H, Tsugane K, Kanno T, Nishimura T. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant Cell Physiol.* 2012; 53:766–784. [PubMed: 22302712]
52. An J, Rao A, Ko M. TET family dioxygenases and DNA demethylation in stem cells and cancers. *Exp Mol Med.* 2017; 49:e323. [PubMed: 28450733]
53. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science.* 2017; 357
54. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghighi FG, Sejnowski TJ, Behrens MM, Ecker JR. Global epigenomic reconfiguration during mammalian brain development. *Science.* 2013; 341:1237905. [PubMed: 23828890]
55. Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development.* 2011; 138:811–820. [PubMed: 21247965]
56. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of Eukaryotic DNA methylation. *Science.* 2010; 328:916–919. [PubMed: 20395474]
57. Goyon C, Rossignol JL, Faugeron G. Native DNA repeats and methylation in *Ascomobolus*. *Nucleic Acids Res.* 1996; 24:3348–3356. [PubMed: 8811089]
58. Takayama S, Dhahbi J, Roberts A, Mao G, Heo SJ, Pachter L, Martin DI, Boffelli D. Genome methylation in *D. melanogaster* is found at specific short motifs and is independent of DNMT2 activity. *Genome Res.* 2014; 24:821–830. [PubMed: 24558263]
59. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, Cross MK, Williams BA, Stamatoyannopoulos JA, Crawford GE, Absher DM, Wold BJ, Myers RM. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* 2013; 23:555–567. [PubMed: 23325432]
60. Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell.* 2012; 148:816–831. [PubMed: 22341451]

61. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462:315–322. [PubMed: 19829295]
62. Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H. Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. *PLoS Genet*. 2013; 9:e1003439. [PubMed: 23637617]
63. Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A*. 2000; 97:5237–5242. [PubMed: 10805783]
64. Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, Jenuwein T, Xu G, Leonhardt H, Wolf V, Walter J. In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet*. 2012; 8:e1002750. [PubMed: 22761581]
65. Bacolla A, Wang G, Vasquez KM. New Perspectives on DNA and RNA Triplexes As Effectors of Biological Activity. *PLoS Genet*. 2015; 11:e1005696. [PubMed: 26700634]
66. Buske FA, Mattick JS, Bailey TL. Potential in vivo roles of nucleic acid triple-helices. *RNA Biology*. 2014; 8:427–439.
67. Frank-Kamenetskii MD, Mirkin SM. Triplex DNA structures. *Annu Rev Biochem*. 1995; 64:65–95. [PubMed: 7574496]
68. Duca M, Vekhoff P, Oussedik K, Halby L, Arimondo PB. The triple helix: 50 years later, the outcome. *Nucleic Acids Res*. 2008; 36:5123–5138. [PubMed: 18676453]
69. Furlong JC, Sullivan KM, Murchie AIH, Gough GW, Lilley DMJ. Localized Chemical Hyperreactivity in Supercoiled DNA: Evidence for Base Unpairing in Sequences That Induce Low-Salt Cruciform Extrusion. *Biochemistry*. 1989; 28:2009–2017. [PubMed: 2541769]
70. Sullivan KM, Lilley DMJ. Influence of Cation Size and Charge on the Extrusion of a Salt-dependent Cruciform. *J Mol Biol*. 1987; 193:397–404. [PubMed: 3599077]
71. Lee HT, Khutsishvili I, Marky LA. DNA Complexes Containing Joined Triplex and Duplex Motifs: Melting Behavior of Intramolecular and Bimolecular Complexes with Similar Sequences. *J Phys Chem B*. 2010; 114:541–549. [PubMed: 19928823]
72. Soto AM, Loo J, Marky LA. Energetic contributions for the formation of TAT/TAT, TAT/CGC⁺, and CGC⁺/CGC⁺ base triplet stacks. *J Am Chem Soc*. 2002; 124:14355–14363. [PubMed: 12452709]
73. Soto AM, Marky LA. Thermodynamic contributions for the incorporation of GTA triplets within canonical TAT/TAT and C⁺GC/C⁺GC base-triplet stacks of DNA triplexes. *Biochemistry*. 2002; 41:12475–12482. [PubMed: 12369838]
74. Saenger, W. *Principles of Nucleic Acid Structure*. 1. Springer-Verlag; New York: 1984.
75. Westhof E. Water: an integral part of nucleic acid structure. *Annu Rev Biophys Chem*. 1988; 17:125–144. [PubMed: 2456073]
76. Berman HM. Hydration of DNA. *Curr Opin Struct Biol*. 1991; 1:423–427.
77. Rentzeperis D, Kupke DW, Marky LA. Volume changes correlate with entropies and enthalpies in the formation of nucleic acid homoduplexes: Differential hydration of A and B conformations. *Biopolymers*. 1993; 33:117–125. [PubMed: 8427928]
78. Buckin VA, Kankiya BI, Bulichov NV, Lebedev AV, Gukovsky IY, Chuprina VP, Sarvazyan AP, Williams AR. Measurement of anomalously high hydration of (dA)_n•(dT)_n double helices in dilute solution. *Nature*. 1989; 340:321. [PubMed: 2747797]
79. Buckin VA, Kankiya BI, Sarvazyan AP, Uedaira H. Acoustical investigation of poly(dA)•poly(dT), poly[d(A-T)]•poly[d(A-T)], poly(A)•poly(Y) and DNA hydration in dilute aqueous solutions. *Nucleic Acids Res*. 1989; 17:4189–4203.
80. Macgregor RB, Chen MY. ΔV° of the Na⁺-induced B-Z transition of poly[d(G-C)] is positive. *Biopolymers*. 1990; 29:1069–1075. [PubMed: 2369616]
81. Frank HS, Evans MW. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J Chem Phys*. 1945; 13:507.

82. Kauzmann W. Some factors in the interpretation of protein denaturation. *Adv Protein Chem.* 1959; 14:1–63. [PubMed: 14404936]
83. Cantor CR, Warshaw MM, Shapiro H. Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers.* 1970; 9:1059–1077. [PubMed: 5449435]
84. Borer, PN. Optical properties of nucleic acids, absorptions and circular dichroism spectra. In: Fasman, GD., editor. *Handbook of Biochemistry and Molecular Biology.* CRC Press; Cleveland: 1975. p. 589-595.
85. Marky LA, Breslauer KJ. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers.* 1987; 26:1601–1620. [PubMed: 3663875]
86. Kaushik M, Suehl N, Marky LA. Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, d(C(3)TA(2))(4). *Biophys Chem.* 2007; 126:154–164. [PubMed: 16822606]
87. Olsen CM, Shikiya R, Ganugula R, Reiling-Steffensmeier C, Khutsishvili I, Johnson SE, Marky LA. Application of differential scanning calorimetry to measure the differential binding of ions, water and protons in the unfolding of DNA molecules. *Biochim Biophys Acta.* 2016; 1860:990–998. [PubMed: 26450631]
88. Spink CH, Chaires JB. Effects of Hydration, Ion Release, and Excluded Volume on the Melting of Triplex and Duplex DNA. *Biochemistry.* 1999; 38:496–508. [PubMed: 9890933]
89. Colombo MF, Bonilla-Rodriguez GO. The Water Effect on Allosteric Regulation of Hemoglobin Probed in Water/Glucose and Water/Glycine Solutions. *J Biol Chem.* 1996; 271:4895–4899. [PubMed: 8617761]
90. Kypr J, Kejnovská I, Reniuk D, Vorlíková M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* 2009; 37:1713–1725. [PubMed: 19190094]
91. Escude C, François J-C, Sun J-s, Ott G, Sprinzl M, Garestier T, Hélène C. Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies. *Nucleic Acids Res.* 1993; 21:5547–5553. [PubMed: 7506827]
92. Hardin CC, Corregan M, Brown BA II, Frederick LN. Cytosine-cytosine⁺ base pairing stabilizes DNA quadruplexes and cytosine methylation greatly enhances the effect. *Biochemistry.* 1993; 32:5870–5880. [PubMed: 8504107]
93. Povsic TJ, Dervan PB. Triple helix formation by oligonucleotides on DNA extended to the physiological pH range. *J Am Chem Soc.* 1989; 111:3059–3061.
94. Xodo LE, Manzini G, Quadrifoglio F, van der Marel GA, van Boom JH. Effect of 5-methylcytosine on the stability of triple-stranded DNA - a thermodynamic study. *Nucleic Acids Res.* 1991; 19:5625–5631. [PubMed: 1945840]
95. Plum GE, Park YW, Singleton SF, Dervan PB, Breslauer KJ. Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: A spectroscopic and calorimetric study. *Proc Natl Acad Sci U S A.* 1990; 87:943609440.
96. Roberts RW, Crothers DM. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science.* 1992; 258:1463–1466. [PubMed: 1279808]
97. Marky LA, Kupke DW. Enthalpy-entropy compensations in nucleic acids: Contribution of electrostriction and structural hydration. *Methods Enzymol.* 2000; 323:419–441. [PubMed: 10944762]
98. Carr CE, Marky LA. Investigation of the Melting Behavior of DNA Three-Way Junctions in the Closed and Open States. *Biophys J.* 2017; 113:1–11. [PubMed: 28700907]
99. Lee HT, Arciniegas S, Marky LA. Unfolding thermodynamics of DNA pyrimidine triplexes with different molecularities. *J Phys Chem B.* 2008; 112:4833–4840. [PubMed: 18358029]
100. Soto AM, Rentzeperis D, Shikiya R, Alonso M, Marky LA. DNA intramolecular triplexes containing dT → dU substitutions: Unfolding energetics and ligand binding. *Biochemistry.* 2006; 45:3051–3059. [PubMed: 16503660]

Highlights

- Cytosine methylation on the Hoogsteen strand increases the thermodynamic stability of the triplex.
- Methylation on the Crick or both strands increase the thermodynamic stability of the duplex.
- Methylation increases the amount of protonation that occurs upon triplex formation.
- Cytosine methylation causes a release of sodium ions upon folding of the triplex.
- Immobilization of water is not increased upon methylation until a certain threshold is reached.

Figure 1.
Hypothesized structures of triplexes and duplexes. • indicates Watson-Crick base-pairing, * indicates Hoogsteen base-pairing, and mC⁺ represents methylated cytosines.

Figure 2. CD spectra at 20 °C in 10 mM sodium cacodylate at pH 5.2 of *Control Tri* (solid line), *2MeH* (dashed line), *2MeC* (dotted line) and *4MeHC* (short dashed line).

Figure 3. DSC thermograms of *Control Tri* (solid line), *2MeH* (dashed line), *2MeC* (dotted line) and *4MeHC* (short dashed line) in (Top) 10 mM sodium cacodylate at pH 5.2 and (Bottom) 10 mM sodium cacodylate at pH 6.2.

Figure 4.
Uptake of protons for all triplexes in 10 mM sodium cacodylate.

Figure 5.
Uptake of ions for all triplexes in 10 mM sodium cacodylate at pH 5.2.

Figure 6.
Uptake of water for all triplexes in 10 mM sodium cacodylate at pH 5.2.

Table 1

Thermodynamic Profiles for the Unfolding of Triplexes and Duplexes.

pH	Transition	T_M (°C)	H_{Cal} (kcal/mol)	$G^\circ_{(s)}$ (kcal/mol)	$T S_{cal}$ (kcal/mol)
<i>Control Tri</i>					
5.2		56.7	93.9	14.7	79.2
6.2		38.5	87.6	9.4	78.2
<i>2MeH</i>					
5.2		61.6	91.4	15.5	75.9
6.2	1 st	35.2	50.4	4.9	45.5
	2 nd	49.1	45.2	6.2	39.0
Total			95.6	11.1	84.5
<i>2MeC</i>					
5.2	1 st	57.1	57.8	9.1	48.7
	2 nd	66.7	42.2	7.7	34.5
Total			100.0	16.8	83.2
6.2	1 st	37.9	46.4	4.9	41.5
	2 nd	53.5	56.0	8.3	47.7
Total			102.4	13.2	89.2
<i>4MeHC</i>					
5.2	1 st	55.3	51.5	7.9	43.6
	2 nd	70.2	55.7	10.6	45.1
Total			107.2	18.5	88.7
6.2	1 st	35.6	46.4	4.6	41.8
	2 nd	57.9	64.6	10.3	54.3
Total			111.0	14.9	96.1
<i>Control Dup</i>					
7.0		31.9	29.2	2.6	26.6
<i>2MeC-Dup</i>					
7.0	1 st	13.7	4.3	0.1	4.2

pH	Transition	T_M (°C)	H_{Cal} (kcal/mol)	$G^{(S)}$ (kcal/mol)	$T S_{cal}$ (kcal/mol)
	2 nd	36.9	42.3	4.4	37.9
Total			46.6	4.5	42.1

All experiments are done in 10 mM sodium cacodylate buffer. Experimental errors are as follows: T_M (± 0.5 °C), H_{Cal} (± 5 %), $T S$ (± 5 %), G^S (± 7 %).

Table 2

Thermodynamic Contributions of the Methylated Triplexes.

pH	Transition	δT_M (°C)	H_{cal} (kcal/mol)	$G^\circ_{(S)}$ (kcal/mol)	$(T S_{cal})$ (kcal/mol)
<i>2MeH – Control Tri</i>					
5.2		+4.9	+2.5	-0.8	+3.3
6.2	1 st	-3.3			
	2 nd	+10.6			
Total			-8.0	-1.7	-6.3
<i>2MeC – Control Tri</i>					
5.2	1 st	-0.4			
	2 nd	+10.0			
Total			-6.1	-2.1	-4.0
6.2	1 st	-0.6			
	2 nd	+15.0			
Total			-14.8	-3.8	-11
<i>4MeHC – Control Tri</i>					
5.2	1 st	-1.4			
	2 nd	+13.5			
Total			-13.3	-3.8	-9.5
6.2	1 st	-2.9			
	2 nd	+19.4			
Total			-23.4	-5.5	-17.9
<i>2MeC-Dup – Control Dup</i>					
7.0		+5.0	-13.1	-1.8	-11.3

Values calculated relative to the control triplex and duplex.

Table 3

Thermodynamic Profiles for the Formation of Methylated Base-Triplet Stacks

Stack	H_{cal} (kcal/mol)	$T S_{\text{cal}}$ (kcal/mol)	$G^{\circ}_{(S)}$ (kcal/mol)
TAT/CG ^{5Me} C ⁺	−22.9	−18.3	−4.6
TAT/ ^{5Me} CGC ⁺	−23.4	−18.7	−4.7
TAT/ ^{5Me} CG ^{5Me} C ⁺	−23.8	−19.1	−4.8

Values calculated using thermodynamic parameters of base-triplet formation as described in ref [63].

Table 4

Differential Binding of Protons, Counterions and Water for the Unfolding of the Triplexes.

Transition	T_M/pH	n_{H^+} (mol H^+ /mol)	$T_M/\ln[\text{Na}^+]$	n_{Na^+} (mol Na^+ /mol)	$T_M/\ln(a_w)$	n_w (mol H_2O /mol)
<i>Control Tri</i>						
	-14.1	-2.9	0.197	-0.10	171.7	-75
<i>2MeH</i>						
1 st	-21.4	-2.7	-0.884	0.24	185.4	-76
2 nd	-16.3	-1.7	0.564	-0.15		
<i>2MeC</i>						
1 st	-19.3	-2.5	-0.662	0.20	107.9	-29
2 nd	-14.9	-1.3	0.417	-0.09	253.7	-47
<i>4MeHC</i>						
1 st	-19.7	-2.3	-0.475	0.13	142.8	-26
2 nd	-14.8	-1.7	0.091	-0.02	258.2	-61
<i>Control Dup</i>						
	-4.4	-0.3	1.300	-0.22	95.7	-15
<i>2MeC-Dup</i>						
	-5.2	-0.6	1.414	-0.35	252.2	-23

Experimental errors are as follows: n_{H^+} ($\pm 12\%$), n_{Na^+} ($\pm 12\%$), n_w ($\pm 12\%$).