ELSEVIER

Biophysical Chemistry 126 (2007) 154–164

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, $d(C_3TA_2)_4$

Mahima Kaushik [a], Nathan Suehl [a], Luis A. Marky [a,b,c,*]

[a] *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE 68198-6025, United States*
[b] *Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE 68198-6025, United States*
[c] *Eppley Institute for Cancer Research, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE 68198-6025, United States*

Received 24 April 2006; received in revised form 26 May 2006; accepted 30 May 2006
Available online 6 June 2006

Abstract

A combination of spectroscopic and calorimetric techniques is used to determine the unfolding thermodynamics of the complexes formed by the complementary sequence of the human telomere, $d(C_3TA_2)_4$, in the pH range of 4.2 to 6. Calorimetric melting curves show biphasic transitions; both transitions are shifted to higher temperatures as the pH is decreased, indicative of cytosine protonation, which favors the formation of $C{\bullet}C^+$ base pairs. Furthermore, the transition temperature, T_M, of the lower transition depends on strand concentration, while the T_M of the higher transition is independent of strand concentration, indicating the following sequential melting: bimolecular complex(s)→ intramolecular complex→random coil. The thermodynamic profiles for the formation of each complex, bimolecular and i-motif reveals small favorable free energy terms resulting from favorable enthalpy–unfavorable entropy compensations, uptake of protons, marginal uptake of counterions (i-motif) and marginal release of water molecules (i-motif). Furthermore, an enthalpy of 3.2 kcal/mol (bimolecular complex) and 5.0 kcal/mol (i-motif) is estimated for a single $C{\bullet}C^+/C{\bullet}C^+$ base-pair stack.

Keywords: i-Motif DNA; DSC; Thermodynamics; Proton/counterion/water binding

1. Introduction

This article is in honor of Professor Julian Sturtevant, who was one of the leading scientists in the field of biothermodynamics. The author met Professor Sturtevant for the first time in 1982 at a Biopolymers Gordon Conference, where he gave a lecture on "The heat capacities of proteins". This was an excellent lecture. At this early time, he was already correlating the measured heat capacities of several proteins with vibrational degrees of freedom. Later on, in 1988, Microcal gave him a prototype titration calorimeter (Omega) to test the unit, and I was given all of his experimental data. He had numerous controls that convinced me that the Omega titration calorimeter was an excellent instrument; I ended up buying the second Omega unit, and it is still in good working condition. In this paper, I am reporting on the unfolding thermodynamics of the complexes that are formed by a cytosine-rich oligonucleotide. The main conclusions are very close to the experimental data, which was the style of Professor Sturtevant.

Cytosine-rich sequences are frequently found in eukaryotic genomes located near or within regions of functional and/or regulatory importance, and play important biological roles in replication, recombination, transcription and chromatin organization [1]. Forty years ago, it was proposed that poly dC could form parallel duplexes [2] stabilized by $C{\bullet}C^+$ base pairs at acidic pH (Fig. 1); this particular base pairing scheme has been confirmed with the crystal structure of acetyl cytosine [3], poly(dC) and poly(rC) [4,5]. Cytosine-rich oligonucleotides can potentially adopt complex pH-dependent conformations; for instance, the crystal structure of d-TCCCCC [6,7] indicated that two parallel stranded $C{\bullet}C^+$ duplexes with an antiparallel orientation form a four stranded structure. The

* Corresponding author. Department of Pharmaceutical Sciences, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE 68198-6025, United States. Tel.: +1 402 559 4628; fax: +1 402 559 9543.
E-mail address: lmarky@unmc.edu (L.A. Marky).

0301-4622/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2006.05.031

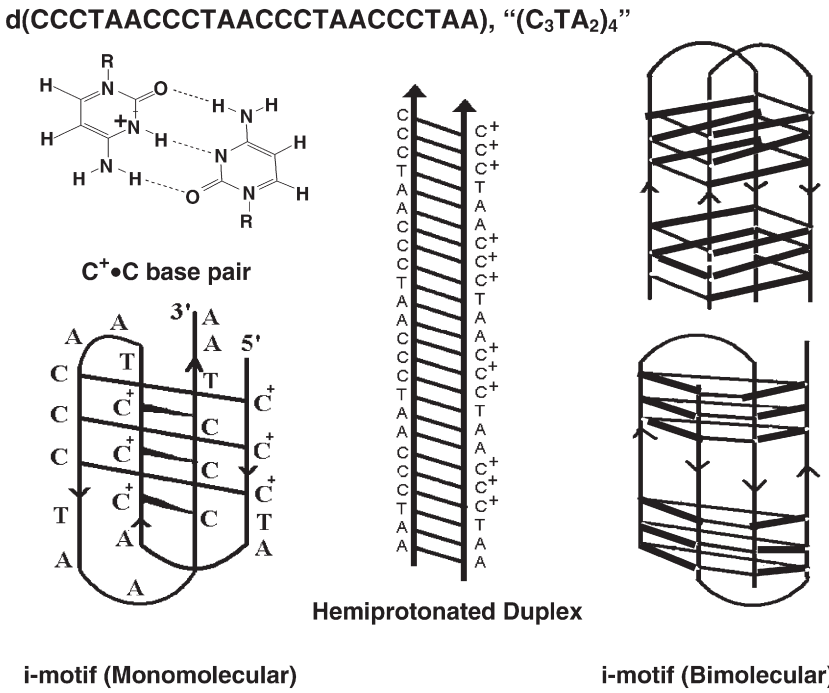

Fig. 1. Sequence, scheme of a C^+•C base-pair and cartoons of the possible conformations adopted by cytosine-rich oligonucleotides.

base pairs of one C•C^+ duplex intercalate into those of the other duplex; hence, the structure is called the "i-motif" or intercalated motif (Fig. 1); the base planes of the stacked C•C^+ base pairs are perpendicular to each other with all the bases in the anti-conformation. This type of alternating stacking of C•C^+ base pairs further stabilizes the i-motif [8]. The formation of i-motifs using different DNA oligonucleotides (ODNs) have been studied extensively by a variety of experimental techniques [6,7,9–18]; however, differential scanning calorimetric techniques have not been used to determinate model-independent unfolding thermodynamics. Furthermore, C-rich ODNs have been used to block gene expression via antisense or antigene strategies [19,20]; however, the formation of i-motif within these ODNs could well interfere with the success of these strategies because the oligonucleotide may get trapped in a stable folded conformation [21]. Therefore, it is necessary to investigate their conformational flexibility and the solution conditions that lead to their stable formation. In addition, since telomeric C-rich ODNs can form i-motif structures in vitro, it has been proposed that these sequences may play a biological role analogous to that of G-quadruplexes [6,22].

In this work, we used a combination of spectroscopic and calorimetric techniques to study the unfolding of the complementary sequence of the human telomere: $d(CCCTAA)_4$. At acidic pH, it yields a bimolecular complex/i-motif equilibrium at low temperatures and their temperature-unfolding takes place through sequential transitions: bimolecular complex→i-motif →random coil. Thus, acidic pH stabilizes both structures through the formations of C•C^+ base pairs. We report complete thermodynamic profiles for their unfolding, including the energetic contributions of the C•C^+/C•C^+ base-pair stacks in each structure and their uptake of protons.

2. Materials and methods

2.1. Materials

The $^{5'}$-d(CCCTAACCCTAACCCTAACCCTAA)-$^{3'}$ oligonucleotide, designated as "C_3TA_2", was synthesized by the Core Synthetic Facility of the Eppley Institute at UNMC, purified by reverse phase HPLC, and further desalted by gel-permeation chromatography, using a Sephadex G-10 column. The solution concentration of C_3TA_2 was determined optically at 260 nm and 80 °C using a molar absorptivity of 2.20×10^5 M^{-1} cm^{-1} (in strands). This value is obtained from the molar absorptivity at 25 °C, obtained from the tabulated values of the dimers and monomer bases [23], and extrapolated to high temperatures using the upper portions of the UV melting curves, following procedures described earlier [24]. All measurements were performed in buffer solutions consisting of 10 mM sodium cacodylate, adjusted to the appropriate pH, salt and osmolyte concentrations with HCl, NaCl and ethylene glycol, respectively. We also use a 10 mM sodium acetate buffer, 100 mM NaCl at pH 4.2 or 4.8. Stock solutions of C_3TA_2 were prepared by dissolving dry ODN in the appropriate buffer. All other reagent grade chemicals, from Sigma Chemicals, were used without further purification.

2.2. Temperature-dependent UV spectroscopy

The spectra of C_3TA_2 were obtained with a thermoelectrically controlled Aviv 14-DS spectrophotometer (Lakewood, NJ) from 330 nm to 200 nm at wavelength steps of 1 nm, as a function of temperature and pH. The reported spectra correspond to the average of three scans. Absorbance versus temperature profiles (UV melting curves) were measured

simultaneously at two wavelengths, 265 nm and 295 nm, and the temperature was scanned at a heating rate of ~0.6 °C/min. Shape analysis of the melting curves yielded transition temperatures, T_M, and model-dependent van't Hoff enthalpies, ΔH_{vH}, which are obtained with the following relationship [25]: $\Delta H_{vH}=(2n+2)RT_M^2(\partial\alpha/\partial T)_{T=T_M}$, where "$n$" is the transition molecularity, equal to 1 for monomolecular transitions and 2 for bimolecular transitions; "α" is the fraction of single strands in the helical state; and $(\partial\alpha/\partial T)_{T=T_M}$ is the slope of the α vs. T curve measured around the T_M, using an interval of 1 °C. The T_M's of biphasic melting curves were obtained from the peaks of the corresponding differential melting curves.

UV melting curves were carried out as a function of strand concentration to check for transition molecularities; intramolecular complexes formed when the T_M is independent of strand concentration, while higher molecularities are obtained when the T_M depends on strand concentration.

2.3. Circular dichroism spectroscopy (CD)

The conformation of C_3TA_2 in appropriate solution conditions was obtained by simple inspection of its CD spectra. The spectra were obtained on a thermoelectrically controlled AVIV Circular Dichroism Spectrometer, Model 202 SF (Lakewood, NJ), from 320 to 220 nm at wavelength steps of 1 nm, using free-strained quartz cuvettes with path length of 10 mm. The reported spectra correspond to the average of at least three scans. Ellipticity versus temperature profiles (CD melting curves) were obtained at 286 nm using a heating rate of ~0.6 °C/min.

2.4. Differential scanning calorimetry (DSC)

Heat capacity functions of the helix-coil transition of C_3TA_2 were measured with a VP-DSC microcalorimeter (Northampton, MA). Two cells, the sample cell containing 0.7 ml of oligomer solution and the reference cell filled with the same volume of buffer solution, were heated from 0 °C to 100 °C at a heating rate of 0.75 °C/min. The biphasic DSC scans were deconvoluted using a "non two-state zero ΔC_p model" or a "two-state zero ΔC_p model" of the Microcal software. Both approaches yielded similar results, and the resulting heat capacity profiles were then normalized by the strand concentration, according to the molecularity of the complex. Analysis of these thermograms yielded model-independent thermodynamic profiles (ΔH_{cal}, ΔS_{cal} and ΔG°_{cal}) and van't Hoff enthalpies. These parameters are measured with the following relationships: $\Delta H_{cal}=\int\Delta C_p(T)\,dT$ and $\Delta S_{cal}=\int(\Delta C_p(T)/T)dT$, where ΔC_p is the anomalous heat capacity of the oligonucleotide solution along the transition curve; the free energy, ΔG°_{cal}, is calculated at 10 °C using the Gibbs relationship: $\Delta G^\circ_{(10)}=\Delta H_{cal}-T\Delta S_{cal}$.

Shape analysis of the experimental DSC curve allows us to calculate ΔH_{vH}'s according to the relationship: $\Delta H_{vH}=A/[(1/T_1)-(1/T_2)]$, where A is a constant, equal to 10.14 for bimolecular transitions and 7.0 for monomolecular transitions [25]; T_1 and T_2 correspond to the lower and upper temperatures, respectively, at the half height width of the DSC curve. The $\Delta H_{vH}/\Delta H_{cal}$ ratio allows us to inspect if DNA unfolding takes place through two-state transitions or through the formation of intermediates. If the $\Delta H_{vH}/\Delta H_{cal}$ ratio is equal to 1, then the transition takes place in an all-or-none fashion [25], while a larger ratio indicates the formation of aggregates or a larger cooperative unit.

2.5. Differential binding of protons, counterions and water molecules

The helical and coil states of an oligonucleotide are associated with a different number of bound protons, ions and water molecules; therefore, their helix→coil transition is accompanied by a differential release (or uptake) of protons, counterions and water molecules. The differential release (or uptake) of each of these species can be measured experimentally using the following reaction:

$$\mathrm{Helix}(a\mathrm{H^+}, b\mathrm{Na^+}, c\mathrm{H_2O})\rightarrow\mathrm{Coil}(x\mathrm{H^+}, y\mathrm{Na^+}, z\mathrm{H_2O}) + \Delta n_{\mathrm{H^+}}\mathrm{H^+} + \Delta n_{\mathrm{Na^+}}\mathrm{Na^+} + \Delta n_w\mathrm{H_2O},$$

where $\Delta n_{H^+}=x-a$, $\Delta n_{Na^+}=y-b$ and $\Delta n_w=z-c$; each of these terms correspond to the differential binding of protons, counterions and water, respectively, and are written on the right hand side of the reaction to indicate "releases"; however, if there is an "uptake" of any of these species, they should be written on the left hand side.

The corresponding reaction constant can be written as:

$$K=\{(\mathrm{Coil})/(\mathrm{Helix})\}(\mathrm{H^+})^{\Delta n_{H^+}}(\mathrm{Na^+})^{\Delta n_{Na^+}}(\mathrm{H_2O})^{\Delta n_w} \quad (1)$$

Using a logarithmic function for simplicity: $\ln K=\ln K\{T, P, \ln(H^+), \ln(Na^+), \ln(H_2O)\}$; its total differential is: $d\ln K=(\partial\ln K/\partial T)dT+(\partial\ln K/\partial P)dP+(\partial\ln K/\partial\ln(H^+))d\ln(H^+)+(\partial\ln K/\partial\ln(Na^+))d\ln(Na^+)+(\partial\ln K/\partial\ln(H_2O))d\ln(H_2O)$; the last three partial differentials correspond to the desired quantities or linking numbers: $\Delta n_{H^+}=(\partial\ln K/\partial\ln(H^+))_{(H_2O),\ (Na^+)}$, $\Delta n_{Na^+}=(\partial\ln K/\partial\ln(Na^+))_{(H^+),\ (H_2O)}$ and $\Delta n_w=(\partial\ln K/\partial\ln(H_2O))_{(H^+),\ (Na^+)}$. All three linking numbers are measured experimentally with the assumption that proton, counterion or water binding to the helical and coil states of the oligonucleotide takes place with a similar type of binding. This is a good assumption for protons and counterions but it may not be correct for water, because the helical state of DNA immobilizes electrostricted water (around charges), while the random coil state immobilizes primarily structural water (around polar and non-polar groups) [26–28]. Applying the chain rule to each partial differential and converting ionic activities to concentrations and natural logarithms for decimal logarithms, we obtain:

$$\Delta n_{H^+} = (\partial\ln K/\partial T_M)(\partial T_M/\partial\ln(H^+)) = -0.434[\Delta H_{cal}/RT_M^2](\partial T_M/\partial\mathrm{pH}), \quad (2)$$

$$\Delta n_{Na^+} = (\partial\ln K/\partial T_M)(\partial T_M/\partial\ln(Na^+)) = 0.483[\Delta H_{cal}/RT_M^2](\partial T_M/\partial\log[Na^+]), \quad (3)$$

$$\Delta n_w = (\partial\ln K/\partial T_M)(\partial T_M/\ln(H_2O)) = 0.434[\Delta H_{cal}/RT_M^2](\partial T_M/\partial\log[H_2O]) \quad (4)$$

The first term in brackets of Eqs. (2)–(4), $[\Delta H_{cal}/RT_M^2]$, is a constant determined directly in differential scanning calorimetric experiments and R is the gas constant, while the second term in parenthesis is also determined experimentally from UV melting curves or DSC curves, from the T_M dependences on the concentration of protons, counterions and water, respectively.

In the determination of Δn_{H^+}, DSC melting experiments were carried out in the pH range of 4.2 to 6.0 and 100 mM NaCl, while for Δn_{Na^+}, UV melting curves were measured in the salt range of 10 mM to 1 M NaCl at pH 5.2. UV melts as a function of the concentration of ethylene glycol (0.6 m to 3.0 m) are carried out to determine Δn_w at pH 5.2 and 100 mM NaCl; this cosolute does not interact specifically with DNA [29]. The osmolality of the latter solutions is obtained with a UIC vapor pressure osmometer Model 830 (Utah), which was calibrated with standardized solutions of NaCl. These osmolalities were then converted into water activities (a_w) using the following equation [30]: $\ln a_w = -\mathrm{Osm}/M_w$, where Osm is the solution osmolality and M_w is the molality of pure water equal to 55.5 mol/kg of H_2O.

3. Results

3.1. UV melting curves of C_3TA_2

The unfolding of C_3TA_2 was investigated initially by UV melting techniques. Fig. 2a shows the UV melts at 265 nm at several pH values and Fig. 2b the corresponding melting curves at 295 nm; in these curves, their final and initial absorbances were normalized to 1, respectively. The shape of all curves is sigmoidal and show more or less single transitions, characterized by a hyperchromic effect at 265 nm and a hypochromic effect at 295 nm, which are consistent with previous reports [17,31]. The T_M increases from 19 °C to 45.6 °C as the pH is decreased from 6.4 to 5.2, which clearly indicates that complex formation is stabilized by acidic pH. The complexes that could be forming at low temperatures include the bimolecular complex, the i-motif or a combination of these complexes in equilibrium. To differentiate which of these complexes are actually forming at these temperatures, we obtained UV melts as a function of strand concentration at pH 4.8. The resulting melting curves at pH 4.8 are biphasic (Fig. 2c) and their T_M dependences on strand concentration are shown in Fig. 2d. Over a 10-fold concentration range, the T_M of the first transition increases with the increase in strand concentration, while the T_M of the second transition remains constant. This indicates the formation of bimolecular and monomolecular complexes, respectively. Therefore, the UV melting behavior of C_3TA_2 is through the unfolding of a bimolecular complex at lower temperatures, followed by an intramolecular complex that yields the final random coil state at high temperatures.

To further investigate complex formation at low temperatures, the UV spectra of C_3TA_2 as a function of pH at 10 °C is shown in Fig. 3. The overlay of the spectra shows three

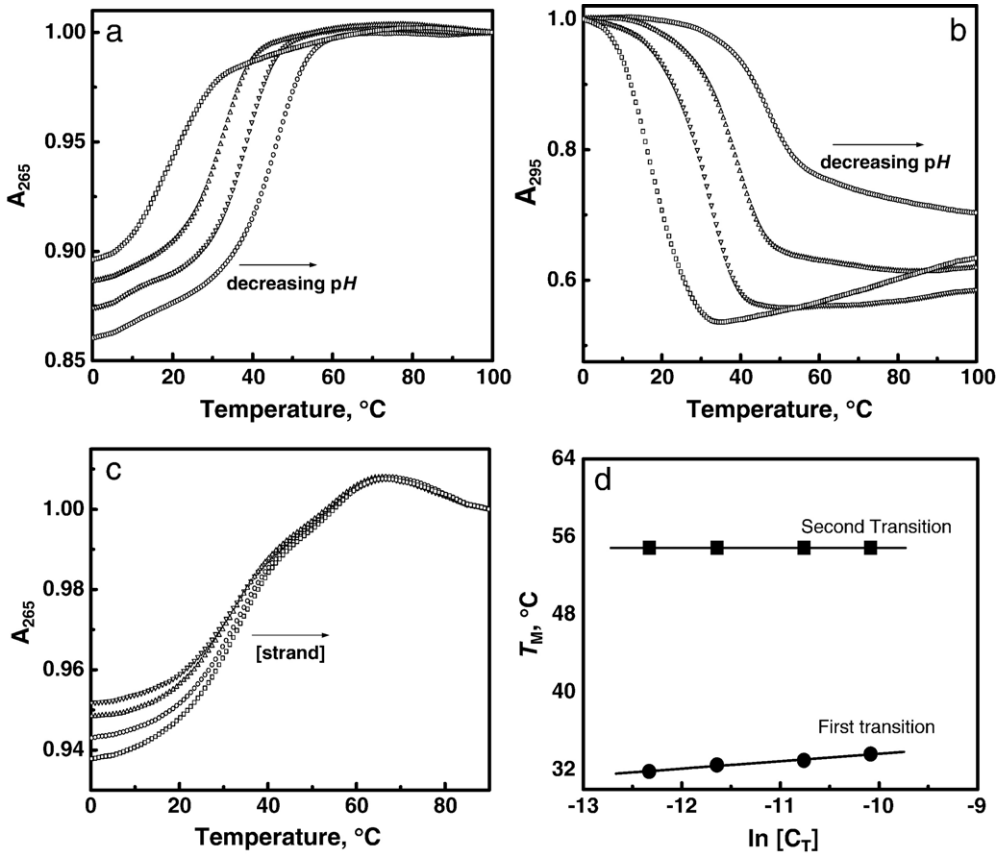

Fig. 2. UV melting curves of C_3TA_2 in 10 mM sodium cacodylate buffer as a function of pH (a) at 265 nm, (b) at 295 nm, (c) at 265 nm in 10 mM sodium acetate buffer at pH 4.8 as a function of strand concentration, and (d) T_M dependence on total strand concentration.

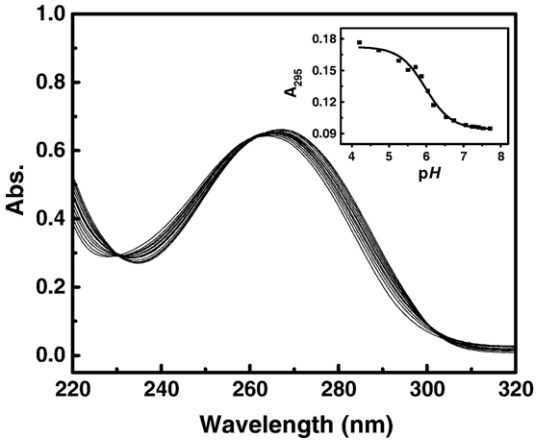

Fig. 3. UV absorption spectrum of C_3TA_2 in 10 mM sodium acetate buffer as a function of pH. The inset shows the A_{295} as a function of pH.

isosbestic points (230 nm, 262 nm and 302 nm), which indicates the presence of an equilibrium of at least two different species. The inset of this figure shows the pH dependence of the absorbance at 295 nm, which yielded a sigmoidal curve with a midpoint at a pH of 6.0 (±0.2), providing further evidence that at least two protonated species exists at low temperatures.

3.2. CD spectra and CD melting curves

Fig. 4a shows the CD spectra of C_3TA_2 at several pH values. At acidic pH, the CD spectra show a large positive band with a peak at 286 nm and a negative band centered at 254 nm. This type of spectrum has been attributed to the formation of C•C$^+$ base pairs [32], characteristic of the i-motif structure [33]. As the pH is gradually increased, a blue shift is observed followed by a sharp decrease in the magnitude of the ellipticity at 285 nm. Furthermore, the overlay of the CD spectra also shows an isodichroic point at 275 nm, suggesting the equilibrium of two species. Fig. 4b shows the CD spectra of C_3TA_2 at different temperatures. The overall comparison of the CD spectra as a function of pH (Fig. 2a) with the ones as a function of temperature (Fig. 2b) indicates that the observed spectral changes with the increase in temperature correspond to the destabilization of the complexes with the increase in pH. For instance, the spectrum of C_3TA_2 at pH 7.4 is similar to the one obtained at pH 5.2 and 80 °C, which is assigned to the spectrum of the random coil state. Fig. 4c shows the corresponding CD melting curves at several pH values. The T_M decreases (45.6 °C to 19 °C) with the increase in pH (5.2 to 6.4), in a similar way as the UV melts. These CD results also confirm the formation of structures that are stabilized by C•C$^+$ base pairs at acidic pH. Fig. 4d shows the dependence of the ellipticity at 286 nm on pH at 10 °C, the resulting curve is sigmoidal with a midpoint at pH=6.2 (±0.2), which is in good agreement with the one obtained from the UV spectra (inset of Fig. 3). The narrow pH range strongly indicates a cooperative transition between protonated species and their random coil states. The combined results are consistent with the presence of the equilibrium of protonated species with the random coil state. Furthermore, the

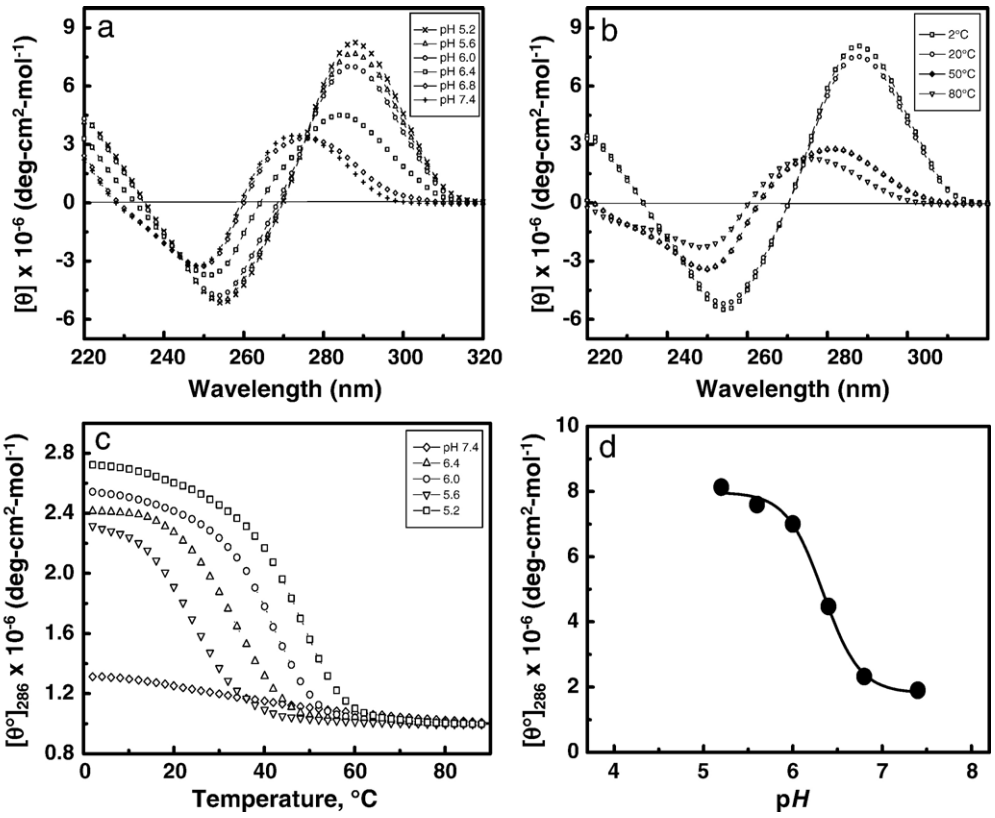

Fig. 4. CD spectra of C_3TA_2 (a) in 10 mM sodium cacodylate buffer at several pHs, (b) in 10 mM sodium cacodylate buffer at pH 5.2 as a function of temperature, (c) CD melting curves at several pHs, and (d) pH dependence of the molar ellipticity at 286 nm.

midpoints of the UV and CD observables as a function of pH (6.0 and 6.2) may well correspond to the average pK_a of the protonated species.

3.3. Differential scanning calorimetry

DSC melting curves were carried out to determine complete thermodynamic profiles for the unfolding of C_3TA_2. While obtaining the DSC scans, we observed that, at pH ≥ 5.2, all scans were super imposable after a short period of equilibration; however, at pHs below 5.2, the scans were not reproducible and smaller heats were obtained with each subsequent scan. Hence, to ensure reproducibility of the DSC curves at the lower pHs, we developed a protocol in which C_3TA_2 solutions were prepared in the DSC sample cell by first heating and keeping it at 90 °C for 5 min, slowly cooled to room temperature, then kept at 4 °C for few hours and a heating scan recorded. All DSC scans were reproducible among samples undergoing this temperature protocol, and only these scans were used for the determination of thermodynamic profiles. To confirm the molecularity of each complex, DSC melts of C_3TA_2 were carried out at pH 4.8 using two different total strand concentrations, 75 μM and 251 μM. The resulting DSC curves are shown in Fig. 5a, each curve yielded biphasic transitions, the T_M of the first transition increases with the increase of strand concentration, while the T_M of the second transition remains the same. This confirms that the first transition corresponds to the unfolding of a bimolecular complex while the second transition involves the unfolding of an intramolecular complex. The curves yielded total heats of 1.25 mcal and 3.96 mcal that when normalized by the total strand concentration yielded similar overall enthalpies, as seen in the similarity of the whole areas under these curves, Fig. 5a. This confirms that the observed transitions correspond to the sequential transition of a bimolecular complex(s)→intramolecular→random coil state. Following this observation, the associated enthalpies are calculated by using one half of the total concentration of strands for the bimolecular transition and the total concentration of strands for the monomolecular transition. This exercise yields average enthalpies of 31.5 kcal/mol and 14.5 kcal/mol for the unfolding of the bimolecular and intramolecular complexes, respectively. We speculate that at low temperatures single bimolecular species are present, which could be either the hemiprotonoted duplex or a bimolecular i-motif; alternatively, these two conformations may be present in equilibrium, while at higher temperatures, the intramolecular i-motif is most likely to be the only conformation present.

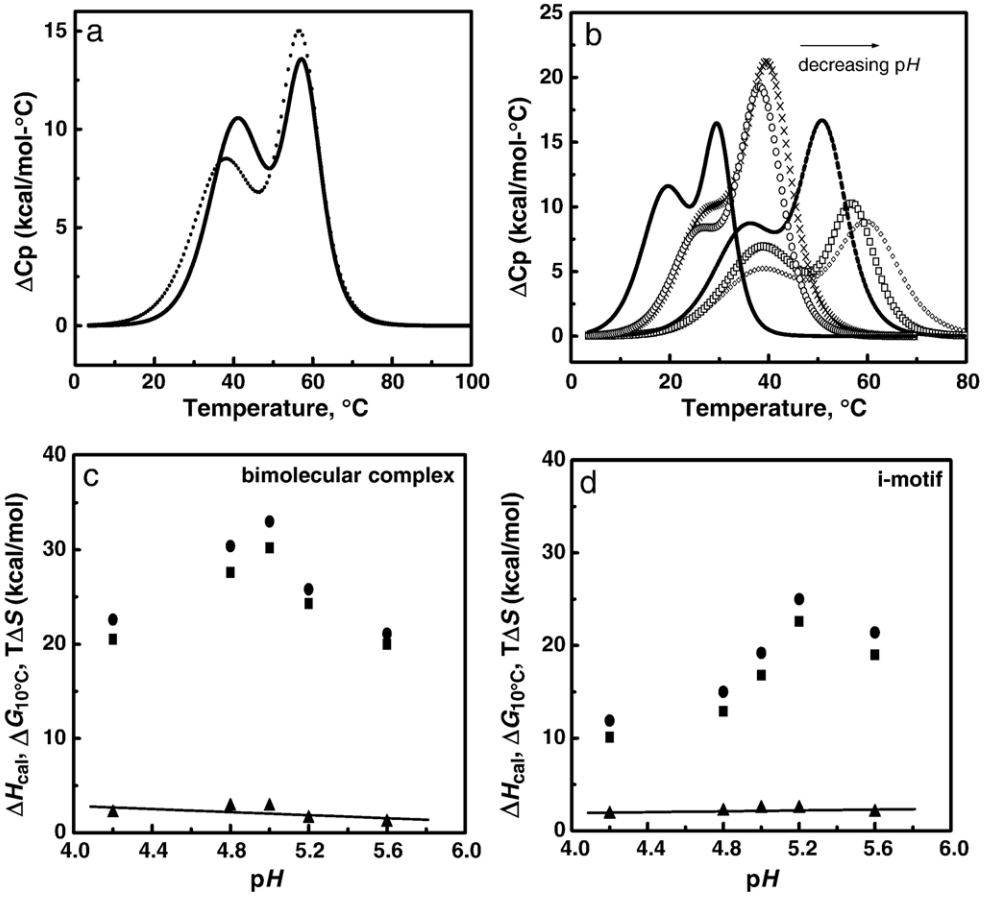

Fig. 5. DSC melting profiles of C_3TA_2 (a) in 10 mM sodium acetate buffer, 100 mM NaCl at pH 4.8 and strand concentration of 251 μM (solid line) and 75 μM (dotted line); (b) in 10 mM sodium cacodylate buffer at pHs of 5.0, 5.2, 5.6 and 6.0, or in 10 mM sodium acetate buffer, 100 mM NaCl at pHs of 4.2 and 4.8. Thermodynamic profiles for the unfolding of each complex, (c) bimolecular complex and (d) i-motif, as a function of pH; symbols as follows: enthalpy (circles), entropy (squares) and free energies (triangles).

DSC melting curves as a function of pH are shown in Fig. 5b. We obtained clear biphasic transitions at each pH, which show the formation of at least two protonated complexes. The T_M of each transition increases as the pH is decreased. DSC thermodynamic profiles for the formation of each complex (bimolecular and i-motif complexes) are shown in Fig. 5c and d. Inspection of these figures shows that the folding of each complex at 10 °C is accompanied by a favorable and small free energy term resulting from the relative large compensation of a favorable enthalpy and unfavorable entropy contributions. The favorable enthalpy terms correspond primarily to the formation of base-pair stacks, while the unfavorable entropy terms arise from contributions of the unfavorable ordering of strands and the putative uptake of counterions, protons and water molecules, as will be discussed in the following sections. A closer inspection of Fig. 5c and d shows for each complex that the enthalpy and entropy terms follow an initial increase at low pHs, go through a maximum at pH 5–5.2 and decrease with the further raise in pH, while the free energy terms remains insensitive to pH. These trends perhaps reflect an optimum pH range for the stacking of these type of C•C$^+$ base pairs, similar to the preferred pH range of 6–8, for the stacking of the natural W–C base pairs.

The associated calorimetric van't Hoff enthalpies are similar to the ones obtained from shape analysis of the UV melting curves, and range 47–74 kcal/mol as the pH increases from 4.2 to 6.0 for the unfolding of the bimolecular complex and remains constant, ~66 kcal/mol, for the unfolding of the i-motif. On the average, we obtained a $\Delta H_{vH}/\Delta H_{cal}$ ratio of 2.4 (bimolecular complex) and 4.1 (i-motif), suggesting the formation of aggregated states. However, the proximity of both transitions and/or the presence of the equilibria of protonated species make it difficult to assess the actual van't Hoff values. Furthermore, the analysis of the DSC melting curves allow us to measure the $\Delta H_{cal}/RT_M^2$ term that is used to obtain experimental measures of the release of protons, counterions and water molecules, without the need of using van't Hoff enthalpies and polyelectrolyte theory.

3.4. Differential binding of protons

The T_M dependence on pH is shown in Fig. 6a; the T_Ms were extracted from the DSC curves of Fig. 5b. Fig. 6a shows two sigmoidal curves with midpoints at pH 5.1 (bimolecular complex, lower line) and 5.0 (monomolecular i-motif, upper line), and negative slopes around the pH of these midpoints. This means, according to Eq. (2), that the unfolding of each complex is accompanied by a release of protons. The slopes determined around the midpoints, 40.7 °C (bimolecular) and 34.5 °C (i-motif), together with the $\Delta H_{cal}/RT_M^2$ terms, measured by averaging the values of the DSC melts of each complex at pH 5 and 5.2, yielded average proton releases of 1.9 mol of protons/mol of DNA for each of these complexes. This release of protons indicates that complex formation is accompanied by the protonation of cytosines, which should be proportional to the difference in the pK_a's between folded and unfolded cytosines.

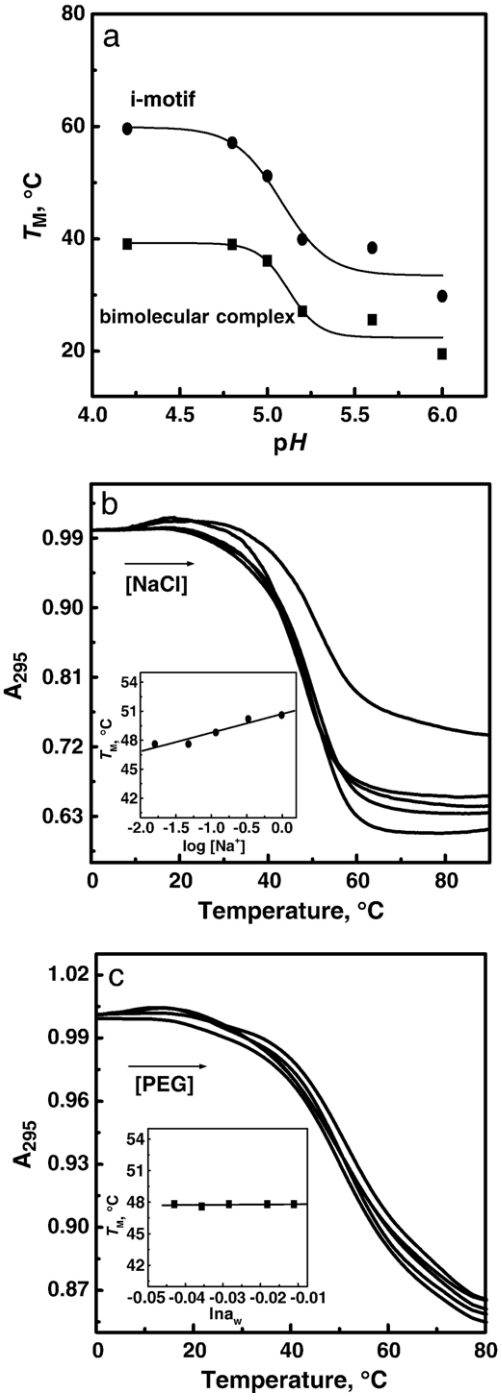

Fig. 6. (a) T_M dependence as a function of pH for each complex, all points obtained from the DSC experiments; (b) UV melting curves of C_3TA_2 in 10 mM sodium cacodylate buffer as a function of salt concentration at pH 5.2; inset shows the resulting plot of the dependence of T_M on salt concentration; (c) UV melting curves as a function of the concentration of ethylene glycol concentration at pH 5.2 and 100 mM NaCl; inset shows the resulting plot of the dependence of T_M on water activity.

3.5. Differential binding of counterions

UV melting curves as a function of salt concentration at pH 5.2 are shown in Fig. 6b, and the resulting T_M dependence on salt is shown in the inset of this figure. All UV melting curves show monophasic transitions with T_M's ranging from 47.6 °C to

50.6 °C, which after correlating with the T_M's of the DSC melts correspond to the unfolding of the intramolecular i-motif. The T_M dependence on log [Na^+] (inset of Fig. 6b) yielded a positive slope of 1.9 °C, and indicates a release of counterions upon unfolding. Using Eq. (3), we determine an average Δn_{Na^+} value of 0.1 mol Na^+/mol DNA over this range of salt concentration at pH 5.2. The main observation is that i-motif formation is accompanied by a marginal uptake of counterions, consistent with the protonation of cytosine that effectively excludes counterions.

3.6. Differential binding of water

UV melting curves of C_3TA_2, as a function of the concentration of ethylene glycol at pH 5.2 and 100 mM NaCl, are shown in Fig. 6c. All curves show monophasic transitions with T_M's (~47.8–48.9 °C) that correspond to melting of the i-motif. The T_M dependence on the activity of water is shown in the inset of Fig. 6c; the resulting line yields a small positive slope of 2.4 °C. Using Eq. (4), we calculate a marginal release of water molecules, $\Delta n_w = 0.3$ mol H_2O/mol strands, for the unfolding of the i-motif of C_3TA_2.

4. Discussion

We use a combination of UV and DSC melting protocols to investigate the unfolding of the complementary sequence of the human telomere. The UV and CD spectra of this oligonucleotide in solutions at acidic pH indicated the presence of multiple equilibria of helical complexes. All complexes contain C^+ bases, stabilized by the formation of C•C^+ base pairs and C•C^+/C•C^+ base-pair stacks. We were able to identify the species involved in this equilibria and determined complete thermodynamic profiles for their unfolding, including their differential binding of protons, counterions and water molecules. In the following sections, a discussion of the main findings is presented.

4.1. C_3TA_2 Forms several complexes at acidic pH

The unfolding of C_3TA_2 was investigated initially by UV and CD melting techniques as a function of pH. The melting behavior changed from monophasic to biphasic as the pH is decreased, and the T_M of each transition increases with the decreased in pH. This shows the presence of at least two DNA helical complexes, each stabilized by acidic pH. We concentrated on the biphasic melting behavior at acidic pH, UV and DSC melts at pH 4.8 show that the T_M of the first transition increases while the T_M of the second transition remains the same when the strand concentration is increased. This shows C_3TA_2 unfolds through a sequential melting of a bimolecular complex(s)→intramolecular complex→random coil. In addition, the CD spectra at acidic pH show the characteristic bands for the formation of C•C^+ base pairs with both complexes, in agreement with previous reports on the formation of double helical structures with cytosine-rich oligomer and polymer sequences [34–36]. Furthermore, the overlay of the UV and CD spectra of C_3TA_2 yielded several isosbestic points, confirming the equilibrium of at least two complexes. This observation is consistent with previous findings that cytosine-rich DNA sequences can adopt complex pH-dependent conformations. For instance, NMR solution studies have shown that d (CCCTAA) can form multiple i-motif four stranded complexes with different topologies [37]. Parallel stranded structures with heteropyrimidine oligonucleotides at pH 4–5.5 were detected using IR, UV melting and fluorescence techniques [38]. The d ($[C_3TA_2]_3CCCT$) and d($[C_4A_2]_3CCCC$) ODNs form in solution intramolecular i-motif structure at acidic pH [9,17]. All of these novel structures are stabilized by C•C^+ base pairs.

4.2. Thermodynamic profiles for the unfolding of C_3TA_2

Complete thermodynamic profiles for the unfolding of C_3TA_2 were obtained from DSC melting experiments. All DSC curves show biphasic transitions at each pH, and indicate a sequential transition of a bimolecular complex(s)→intramolecular→random coil state, which further confirmed the formation of at least two protonated complexes. This sequential melting behavior is similar to the one observed in the unfolding of the Dickerson-Drew dodecamer, which shows two transitions: duplex→hairpin→random coil at neutral pH [24]. This sequential melting can be explained in terms of the entropy penalty of complex formation, and consistent with conventional wisdom, that the bimolecular complex should form at lower temperatures because of its higher unfavorable entropy contribution. The formation of tetramolecular complexes is ruled out for a similar reason, if these were to form then their unfolding will show a T_M of ~15 °C at ~10 mM strand concentrations [39]. We speculate that at low temperatures only one type of bimolecular specie is present, the hemiprotonated duplex or a bimolecular i-motif (Fig. 1), but it is quite possible that both bimolecular complexes are present in equilibrium.

DSC thermodynamic profiles for the unfolding of each complex, bimolecular and i-motif, are shown in Fig. 5c and d. The favorable formation of each complex is enthalpy driven in the pH range of 4.2 to 5.6 and the low free energy values results from relatively large enthalpy–entropy compensation. To explain the enthalpy (and entropy) dependence on pH, we invoke differences in stacking contributions as the pH is increased. At pH<3.5, formation of C•C^+ base pairs and C•C^+/C•C^+ base-pair stacks is not possible because all cytosines will be protonated and their enthalpic contribution is negligible. As the pH is increased, pH>3.5, some of the cytosines get deprotonated, reaching eventually the required 50% deprotonation, which correspond to the optimum number of C•C^+ base pairs to form each of these complexes. The enthalpy values parallel this behavior and their increasing values result from the formation of a higher number of base-pair stacks. Further pH increases (pH>5.2) keep deprotonating the cytosines, which will reduce both the number of C•C^+ base pairs and C•C^+/C•C^+ base-pair stacks, yielding a net reduction of enthalpy contributions. The net enthalpy values in this pH range may include stacking contributions from the single strands. The overall enthalpy (and entropy) dependence on pH shows

formation of the optimum number of base-pair stacks at pH 5.2, where 50% of the cytosines are protonated, suggesting that the pK of the N3 of cytosine is 5.2. Complete thermodynamic profiles for the unfolding of each complex at this optimum pH are shown in Table 1.

4.3. Complex formation is accompanied by a net uptake of protons

We have obtained experimentally each of the two parameters needed of Eqs. (2)–(4) to measure the differential binding of protons, counterions and water molecules between the folded and unfolded states of each complex. Specifically, we determined that each complex (bimolecular and i-motif) binds 1.9 mol protons/mol DNA in the pH range of 5–5.2. In addition, marginal binding of counterions and water molecules are obtained by the intramolecular i-motif complex. The net binding of protons and marginal uptake of counterions in the formation of these complexes is in good agreement with what has been observed in earlier investigations of DNA triplexes containing C^+GC base triplets [39,40], the positively charged cytosine bases excludes counterions due to the local and fixed charged density parameter [41]. The marginal uptake of water molecules is explained in terms of a compensating effect between the two types of water molecules, electrostricted (around charges) and structural (around polar and non-polar groups), that are involved in the unfolding of a DNA complex. The helical state due to its higher charged density parameter immobilizes electrostricted water, while the random coil state with a higher base exposure to the solvent immobilizes structural water; these two contributions oppose to each other, and if their magnitudes are similar, would yield a marginal net effect. In addition, this parameter was measured in the presence of 100 mM NaCl, which will tend to screen the interactions of water in both DNA states yielding much lower effects.

At the pH range of 5 to 5.2, where optimum base-pair stacking interactions takes place, we measured a similar proton uptake of 1.9 mol protons/mol DNA for each type of complex (Table 1). If we use the Henderson–Hasselbach relationship and a pK_a of 4.6 for the N3 protonation site of cytosine in the random coil state [42], we estimate a 20% degree of cytosine protonation takes place at pH 5.2. In a similar exercise, but including the experimental values for the uptake of protons by each helical complex, we obtain 28% of cytosine protonation for each complex. This translates to a protonation of 3.3 cytosines out of the 6 needed for an optimum i-motif complex and 6.6 protonated cytosines out of 12 needed in the bimolecular complex. Alternatively, these proton uptakes allow us to estimate pK_a's of 4.8 for these two complexes. These values are close to the pH of 5.2 where optimum base-pair stacking interactions were observed within each complex, and smaller than the midpoints of the T_M, absorbance and ellipticity dependences on pH, 5.0, 6.0 and 6.2, respectively, which the last two values are considered as the average pK_a of the cytosine group in the intramolecular i-motif. This discrepancy may well be in the strand concentration used in the calorimetric (~130 μM) and optical experiments (~7 μM); lower concentrations will tend to dissociate the acid form to a larger extent.

Table 1
Thermodynamic profiles for the unfolding of C_3TA_2 at pH 5.2 [a]

T_M (°C)	ΔH_{cal} (kcal/mol)	ΔH_{vH} (kcal/mol)	$T\Delta S_{cal}$ (kcal/mol)	$\Delta G^o_{(10)}$ (kcal/mol)	Δn_{H^+} (mol H^+/mol)
		Bimolecular complex			
27.1	25.8	69.1	24.3	1.5	1.7
		i-Motif			
39.9	25.0	62.5	22.6	2.4	2.3

[a] All parameters measured in a 10 mM sodium cacodylate buffer at pH 5.2. The Δn_{H^+} parameter was measured in the same buffer containing 100 mM NaCl. Experimental errors are shown in parentheses: T_M (±0.5 °C), ΔH_{cal} (±5%), $T\Delta S_{cal}$ (±5%), $\Delta G^o_{(10)}$ (±7%) and Δn_{H^+} (±5%).

Table 2
Thermodynamics of $C{\cdot}C^+/C{\cdot}C^+$ base-pair stacks [a]

Complex	$\Delta G°$ (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
Hemiprotonated	−0.2	−3.2	−3.0
Bimolecular i-motif	−0.2	−2.6	−2.4
i-Motif	−0.5	−5.0	−4.5

[a] All parameters measured in a 10 mM sodium cacodylate buffer at pH 5.2. Experimental errors as follows: ΔH (±5%), $T\Delta S$ (±5%) and $\Delta G°$ (±7%).

4.4. Thermodynamic contributions for the stacking of $C{\bullet}C^+$ base pairs

We use the thermodynamic profiles shown in Table 1 to determine the thermodynamic contributions for a single $C{\bullet}C^+/C{\bullet}C^+$ base-pair stack at pH 5.2, we simply divide the magnitude of each parameter by the possible number of $C{\bullet}C^+/C{\bullet}C^+$ base-pair stacks of each structure, and assuming that the other bases are not contributing at all. Since C_3TA_2 contains 12 cytosines, 6 should be protonated in the intramolecular i-motif, yielding 5 $C{\bullet}C^+/C{\bullet}C^+$ base-pair stacks; the hemiprotonated duplex and bimolecular i-motif both contain 8 and 10 $C{\bullet}C^+/C{\bullet}C^+$ base-pair stacks, respectively. The results are presented in Table 2; for instance, the free energy contribution for the formation of a single $C{\bullet}C^+/C{\bullet}C^+$ base-pair stack at pH 5.2 is estimated to be equal to −0.2 kcal/mol (bimolecular complex) and −0.5 kcal/mol (i-motif), while the enthalpy is −3.2 kcal/mol (hemiprotonated duplex) and −5.0 kcal/mol (i-motif). As expected, the overall profiles are much smaller than the average thermodynamic profiles for the formation of a canonical W–C base-pair stack and consistent with the non-vertically stacking of the cytosine aromatic rings. For instance, the enthalpies of −2.6 to −5.0 kcal/mol are lower than the −8 kcal/mol of a DNA base-pair stack, which corresponds to the extend of overlapping of three aromatic rings over three aromatic rings such as pyrimidine–purine/purine–pyrimidine.

5. Conclusion

We have used a combination of spectroscopic and calorimetric techniques to investigate the unfolding thermodynamics

of the structures formed by the cytosine-rich oligonucleotide $d(CCCTAA)_4$, which is a repeat of the complementary sequence of the human telomere. UV and DSC melting profiles show the following sequential melting: bimolecular complex(s)→intramolecular complex→random coil. As the pH is increased, the melting curves are shifted to lower temperatures, indicating that acidic pH stabilizes the structures formed at low temperatures, by protonating the cytosines that in turn form $C\bullet C^+$ base pairs. A marginal binding of counterions and water molecules to these protonated structures is obtained, which are coupled to the local protonation of cytosines. We also determined standard thermodynamic profiles for the formation of $C\bullet C^+/C\bullet C^+$ base-pair stacks in each of these complexes at the optimum pH of 5.2.

The present work should help in studying the different structures formed by cytosine-rich sequences and associated equilibriums. The biological significance of studying these sequences is that cytosine-rich strands are not only present naturally as the complement to the guanine-rich strand of telomeric DNA but they are also used as the third strand for triplex formation in the antigene strategy. Furthermore, this study provides information about protonation of the cytosine bases, which may help future studies with cytosine-rich sequences at acidic pH.

In addition, it has been suggested that the DNA i-motif structure may have some biological relevance, due to the discovery of nuclear proteins that bind to cytosine-rich sequences containing repeated stretches of at least three cytosines [43–45].

Acknowledgements

This work was supported by Grant MCB-0315746 from the National Science Foundation.

References

[1] R.D. Wells, D.A. Collier, J.C. Hanvey, M. Shimizu, F. Wohlrab, The chemistry and biology of unusual DNA structures adopted by oligopurine·oligopyrimidine sequences, FASEB J. 2 (1988) 2939–2949.
[2] R. Langridge, A. Rich, Molecular structure of helical polycytidylic acid, Nature 198 (1963) 725–728.
[3] R.E. Marsh, R. Bierstedt, L. Eichhorn, The crystal structure of cytosine-5-acetic acid, Acta Crystallogr. 15 (1962) 15,310–15,316.
[4] E.O. Akinrimisi, C. Sander, O. Ts'o, Properties of helical polycytidylic acid, Biochemistry 2 (1963) 340–344.
[5] R.B. Inman, Multistranded DNA homopolymer interactions, J. Mol. Biol. 10 (1964) 137–146.
[6] K. Gehring, J.-L. Leroy, Fueron, A tetrameric, DNA structure with protonated cytosine·cytosine base pairs, Nature 363 (1993) 561–565.
[7] J.-L. Leroy, K. Gehring, A. Kettani, Gueron, Acid multimers of oligodeoxycytidine strands: stoichiometry, base-pair characterization, and proton exchange properties, Biochemistry 32 (1993) 6019–6031.
[8] M. Gueron, -L. Leroy, The i-motif in nucleic acids, Curr. Opin. Struct. Biol. 10 (2000) 326–331.
[9] J.-L. Leroy, M. Gueron, J.-L. Mergny, Helene, intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif, Nucleic Acids Res. 22 (1994) 1600–1606.
[10] J.-L. Leroy, M. Gueron, Solution structures of the i-motif tetramers of d(TCC), d(5methylCCT) and d(T5methylCC): novel NOE connections between amino protons and sugar protons, Structure 3 (1995) 101–120.
[11] C.H. Kang, I. Berger, C. Lockshrin, R. Ratliff, R. Moyzis, A. Rich, Crystal structure of intercalated four-stranded $d(C_3T)$ at 1.4 Å resolution, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 11636–11640.
[12] C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, Stable loop in the crystal structure of the intercalated four-stranded cytosine-rich metazoan telomere, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 3874–3878.
[13] I. Berger, C. Kang, A. Fredian, R. Ratliff, R. Moyzis, A. Rich, Extension of the four-stranded intercalated cytosine motif by adenine·adenine base pairing in the crystal structure of d(CCCAAT), Nat. Struct. Biol. 2 (1995) 416–425.
[14] L. Chen, L. Cai, X. Zhang, A. Rich, Crystal structure of a four-stranded intercalated DNA: d(C4), Biochemistry 22 (1994) 13540–13546.
[15] S. Nonin, J.-L. Leroy, Structure and conversion kinetics of a bi-stable DNA i-motif: broken symmetry in the $[d(5mCCTCC)]_4$ tetramer, J. Mol. Biol. 261 (1996) 399–414.
[16] J.-L. Mergny, Following G-quartet formation by UV-spectroscopy, FEBS Lett. 435 (1998) 74–78.
[17] G. Manzini, N. Yathindra, L.F. Xodo, Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences, Nucleic Acids Res. 22 (1994) 4634–4640.
[18] J.M. Benevides, A.H.-J. Wang, A. Rich, Y. Kyogoku, G.A. Vander Marel, J.H. Van Boom, G.J. Thomas Jr., Raman spectra of single crystals of r(GCG)d(CGC) and d(CCCCGGGG) as models for A DNA, their structure transitions in aqueous solution, and comparison with double-helical poly(dG).poly(dC), Biochemistry 25 (1986) 41–50.
[19] C. Giovannangeli, M. Rougee, T. Garestier, N.T. Thuong, C. Helene, Triple-helix formation by oligonucleotides containing the three bases thymine, cytosine, and guanine, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 8631–8635.
[20] J.-L. Mergny, J.S. Sun, M. Rougee, T. Montenay-Garestier, F. Barcelo, J. Chomilier, C. Helene, Sequence specificity in triple-helix formation: experimental and theoretical studies of the effect of mismatches on triplex stability, Biochemistry 30 (1991) 9791–9798.
[21] L. Lacroix, J.-L. Mergny, J.-L. Leroy, C. Helene, Inability of RNA to form the i-motif: implications for triplex formation, Biochemistry 35 (1996) 8715–8722.
[22] S. Ahmed, A. Kintanar, E. Henderson, Human telomeric C-strand tetraplexes, Nat. Struct. Biol. 1 (1994) 83–88.
[23] C.R. Cantor, M.M. Warshow, H. Shapiro, Oligonucleotide interactions: III. Circular dichroism studies of the conformation of deoxyoligonucleotides, Bioploymers 9 (1970) 1059–1077.
[24] L.A. Marky, K.S. Blumenfeld, S. Kozlowski, K.J. Breslauer, Salt-dependent conformational transitions in the self-complementary deoxydodecanucleotide d(CGCGAATTCGCG): evidence for hairpin formation, Biopolymers 22 (1983) 1247–1257.
[25] L.A. Marky, K.J. Breslauer, Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, Biopolymers 26 (1987) 1601–1620.
[26] E.J. Kings, Volume changes for ionization of formic, acetic and n-butyric acids on the glycimiun ion in aqueous solutions at 25 °C, J. Phys. Chem. 73 (1969) 1220–1232.
[27] B.E. Conway, Ionic Hydration in Chemistry and Biophysics, Elsevier, New York, 1981.
[28] K. Zieba, T.M. Chu, D.W. Kupke, L.A. Marky, Differential hydration of dA·dT base pairing and dA and dT bulges in deoxyoligonucleotides, Biochemistry 30 (1991) 8018–8026.
[29] C.H. Spink, J.B. Chaires, Effects of hydration, ion release, and excluded volume on the melting of triplex and duplex DNA, Biochemistry 38 (1999) 496–508.
[30] M.F. Colombo, G.O. Bonilla-Rodriguez, The water effect on allosteric regulation of hemoglobin probed in water/glucose and water/glycine solutions, J. Biol. Chem. 271 (1996) 4895–4899.
[31] J.-L. Mergny, L. Lacroix, X. Han, J.-L. Leroy, C. Helene, Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA Motif, J. Am. Chem. Soc. 117 (1995) 8887–8898.
[32] E.L. Edwards, M.H. Patrick, R.L. Ratliff, D.M. Gray, $A \cdot T$ and $C \cdot C^+$ base pairs can form simultaneously in a novel multistranded DNA complex, Biochemistry 29 (1990) 828–836.

[33] H. Kanehara, M. Mizuguchi, K. Tajima, K. Kanaori, K. Makino, Spectroscopic evidence for the formation of four-stranded solution structure of oligodeoxycytidine phosphorothioate, Biochemistry 36 (1997) 1790–1797.
[34] D.M. Gray, F.J. Bollum, A circular dichroism study of poly dG, poly dC, and poly dG:dC, Biopolymers 13 (1974) 2087–2102.
[35] C. Marck, D. Thiele, C. Schneider, W. Guschlbauer, Protonated polynucleotides structures: 22. CD study of the acid–base titration of poly(dG).poly(dC), Nucleic Acids Res. 5 (1978) 1979–1996.
[36] H. Robinson, G.A. van der Marel, J.H. Van Boom, A.H. Wang, Unusual DNA conformation at low pH revealed by NMR: parallel-stranded DNA duplex with homo base pairs, Biochemistry 31 (1992) 10510–10517.
[37] K. Kananori, N. Shibayama, K. Gohda, K. Tajima, K. Makino, Multiple four-stranded conformations of human telomere sequence d(CCCTAA) in solution, Nucleic Acids Res. 29 (2001) 831–840.
[38] F. Geinguenaud, J. Liquier, M.G. Brevnov, O.V. Petrauskene, Y.I. Alexeev, E.S. Gromova, E. Taillandier, Parallel self-associated structures formed by T,C-rich sequences at acidic pH, Biochemistry 39 (2000) 12650–12658.
[39] A.M. Soto, J. Loo, L.A. Marky, Energetic contributions for the formation of TAT/TAT, TAT/CGC$^+$, and CGC$^+$/CGC$^+$ base triplet stacks, J. Am. Chem. Soc. 124 (2002) 14355–14363.
[40] A.M. Soto, L.A. Marky, Thermodynamic contributions for the incorporation of GTA triplets within canonical TAT/TAT and C$^+$GC/C$^+$GC base-triplet stacks of DNA triplexes, Biochemistry 41 (2002) 12475–12482.
[41] G.S. Manning, Electrostatic free energy of the DNA double helix in counterion condensation theory, Biophys. Chemist. 101–102 (2002) 461–473.
[42] Ch. Zimmer, H. Venner, Protonations in cytosine in DNA, Biopolymers 4 (1966) 1073–1079.
[43] E. Marsich, A. Piccini, L.E. Xodo, G. Manzini, Evidence for a HeLa nuclear protein that binds specifically to the single-stranded d(CCCTAA)n telomeric motif, Nucleic Acids Res. 24 (1996) 4029–4033.
[44] E. Marsich, L.E. Xodo, G. Manzini, Widespread presence in mammals and high binding specificity of a nuclear protein that recognizes the single-stranded telomeric motif $(CCCTAA)_n$, Eur. J. Biochem. 258 (1998) 93–99.
[45] L. Lacroix, H. Lienard, E. Labourier, M. Djavaheri-Mergny, L. Lacoste, H. Leffcos, J. Tazi, C. Helene, J.-L. Mergny, Identification of two human nuclear proteins that recognize the cytosine-rich strand of human telomeres in vitro, Nucleic Acids Res. 28 (2000) 1564–1575.