Influence of Sequence-Dependent Cytosine Protonation and Methylation on DNA Triplex Stability†

Dietmar Leitner,‡,§ Werner Schröder,|| and Klaus Weisz*,‡

Institut für Chemie der Freien Universität Berlin, Takustrasse 3, und Institut für Biochemie der Freien Universität Berlin, Fabeckstrasse 36A, D-14195 Berlin, Germany

Received November 15, 1999; Revised Manuscript Received February 15, 2000

ABSTRACT: To investigate cytosine protonation and its influence on the sequence-dependent thermal stability of DNA triplexes in detail, we have employed homo- and heteronuclear NMR experiments on specifically ^{15}N-labeled oligodeoxynucleotides that were designed to fold into intramolecular triple helices of the pyrimidine motif under appropriate conditions. These experiments reveal that cytosines in central positions of the triplex are significantly protonated even at neutral pH. However, semiprotonation points for individual cytosine bases as determined from pH-dependent measurements show considerable differences depending on their position. Thus, protonation is disfavored for adjacent cytosines or for cytosines at the triplex termini, resulting in a smaller contribution to the overall free energy of the triple helical system. In contrast, protonation of the base upon substitution of 5-methylcytosine for cytosine in the triplex third strand is only affected to a minor extent, and triplex stabilization by the methyl substituent is shown to primarily arise from stacking energies and/or hydrophobic effects.

Since triple-helical complexes of nucleic acids were first reported in 1957 (*1*) interest in these novel structural variants of nucleic acids has been rapidly growing, especially during the past decade. On one hand, the discovery of H-form DNA, a triplex-containing structure, provoked speculations as to its existence and possible regulatory role in vivo (*2*, *3*). On the other hand, due to the highly specific interactions involved, oligonucleotide-directed triple helix formation was soon recognized as a potential tool to selectively target regions in double-stranded DNA or RNA, thus opening up a wealth of applications in medicine and biotechnology. Examples include the development of antigene oligonucleotides for the site-specific inhibition of transcription (*4*, *5*) or the use of triplex-forming oligonucleotides as sequence-specific artificial nucleases, e.g., for chromosome mapping (*6*).

The recognition of a double-stranded nucleic acid by oligonucleotides is based on the formation of specific hydrogen bonds between nucleobases of the single strand and the purine bases of the Watson−Crick duplex (*7*). The triple helices thus formed can be classified according to the composition and orientation of the third strand. In the pyrimidine motif a pyrimidine-rich oligonucleotide binds parallel to the purine strand of a homopurine−homopyrimidine duplex domain in the major groove of the double helix. Formation of Hoogsteen hydrogen bonds between third-strand pyrimidine and Watson−Crick purine bases results in C^+·GC and T·AT base triplets. In a second family of triple helices a purine-rich oligonucleotide binds antiparallel to the Watson−Crick purine strand of the duplex in its major groove, forming G·GC, A·AT, or T·AT base triplets via reverse Hoogsteen hydrogen bonds.

The contribution of cytosine protonation in providing a stable Hoogsteen base pair in pyrimidine triplexes is often the limiting factor for their stability under neutral solution conditions. To overcome the requirement of acidic pH, several base analogues of cytosine have been designed. Thus it was found that a 5-methyl-substituted cytosine has a stabilizing effect on pyrimidine triplexes even at higher pH, yet major interactions of the methyl group responsible for the increased stability are still a matter of debate. To gain more insight into triplex protonation and its influence on triplex stability, protonation equilibria at specific sites within the complexes have to be carefully monitored and related to the thermodynamics of triplex formation. There have been a few studies employing ethidium bromide fluorescence (*8*), UV absorption spectroscopy (*9*), and CD spectroscopy (*8*, *10*, *11*) yielding semiprotonation points for oligonucleotide triplexes with a third pyrimidine strand. However, the determination of such apparent pK_a values by the above methods is based on a two-state model with the cytosine protonation being the only pH-dependent process for the duplex−triplex transition. Moreover, analysis is restricted to the influence of pH on a global equilibrium process and protonation of individual bases within a sequence are not resolved. However, as has been shown by recent NMR studies on an intramolecular triple helix, protonation strongly depends on the cytosine position within a particular triple

† This work was supported by the Deutsche Forschungsgemeinschaft.

* To whom correspondence should be addressed: Fax +49-30-838-55310; e-mail weisz@chemie.fu-berlin.de.

‡ Institut für Chemie.

§ Present address: Forschungsinstitut für Molekulare Pharmakologie, Alfred-Kowalke-Str. 4, D-10315 Berlin, Germany.

|| Institut für Biochemie.

[1] Abbreviations: CD, circular dichroism; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single quantum coherence spectroscopy.

10.1021/bi992630n CCC: $19.00
Published on Web 04/21/2000

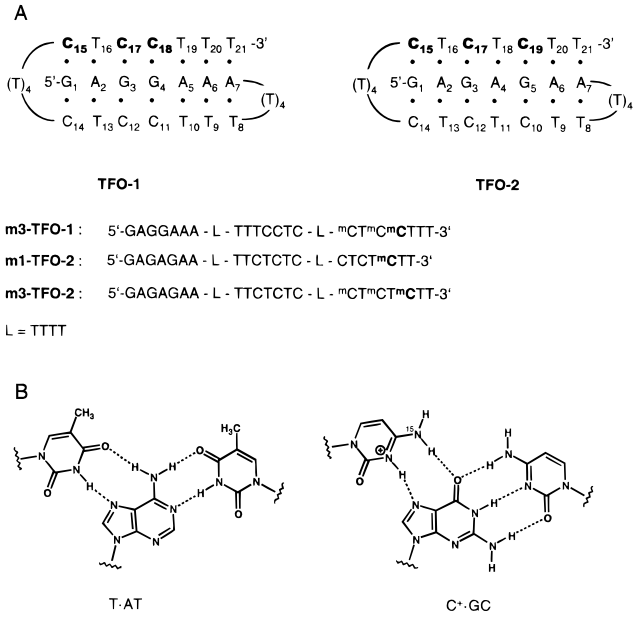

FIGURE 1: (A) Folding of oligodeoxynucleotides TFO-1 and TFO-2 into an intramolecular triplex structure and corresponding sequences with 5-methylcytosine (^{m}C) substitutions in the third strand. ^{15}N-labeled C and ^{m}C nucleotides are shown in boldface type. (B) Structure of T·AT and C^{+}·GC base triplet with Hoogsteen-bound cytosine specifically ^{15}N-labeled at the amino group.

helix and can be very different for various third-strand bases (*12*).

NMR studies on selectively ^{15}N isotope-labeled oligonucleotides have been performed in the past for the structural characterization of triple helices (*13*). Here we report on homo- and heteronuclear NMR experiments on specifically ^{15}N-labeled triple-helix-forming oligomers in an attempt to investigate the protonation of individual third-strand cytosines. All of the oligonucleotides were designed to allow the formation of an intramolecular triple helix with seven base triplets and two T_4 tetraloops by folding back twice on themselves under appropriate conditions. Such intramolecular folded constructs are particularly valuable in thermodynamic (*10*) and NMR structural studies of triple helices (*14*) because of their increased stability, which disfavors competing secondary structures.

MATERIALS AND METHODS

Synthesis and Purification of Oligonucleotides. Oligonucleotides for the UV and NMR studies were synthesized by the standard phosphoramidite method on a 0.2 μmol and 2 μmol scale by TIB MOLBIOL, Berlin, Germany. Following gel filtration, oligonucleotides for the UV melting experiments were further purified by FPLC. Isotope-labeled oligomers were synthesized by the standard protocol using protected β-cyanoethyl phosphoramidites of [4-^{15}N]-2′-deoxycytidine and [4-^{15}N]-5-methyl-2′-deoxycytidine. These were prepared according to published procedures (*15*).

UV Spectroscopy. UV thermal melting profiles were obtained with a Kontron Uvikon 922 spectrophotometer equipped with a Peltier element. Two measurements in 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM MES buffer, pH 5.0, and in 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM PIPES buffer, pH 7.0, were performed on each sample in quartz cuvettes of 1 cm optical path length. Prior to each melting experiment, the samples were annealed by heating to 90 °C for 5 min followed by slow cooling to room temperature and storage at 4 °C overnight. Readings of absorbance A at 260 nm versus temperature T were recorded with a rate of 0.5 °C/min controlled by home-written software. All melting profiles exhibited some hysteresis effects on successive heating and cooling cycles, and melting temperatures T_m were taken as the temperature for half-dissociation of the DNA triplex or duplex upon heating. These were determined from the first-derivative plot dA/dT vs temperature.

Table 1: UV Melting Temperatures T_m at pH 5.0[a] and pH 7.0[b] and Apparent pK_a^{app} Values for Third-Strand C and ^{m}C[c] of the Triplex-Forming Oligodeoxynucleotides

oligomer	third-strand sequence (5′→3′)	T_m[d] (°C) pH 5.0[f]	T_m[d] (°C) pH 7.0[g]	pK_a^{app}[e] C15	pK_a^{app}[e] C17	pK_a^{app}[e] C18/C19
TFO-1	**C**T**CC**TTT	71.9	18.6/66.4	5.3	6.8	6.8
TFO-2	**C**T**C**T**C**TT	73.7	30.5/64.2	5.1	7.4	7.4
m3-TFO-1	^{m}CT^{m}C**^{m}C**TTT	76.0	31.8/65.4			7.0
m1-TFO-2	CTCT**^{m}C**TT	75.8	34.3/64.2			7.5
m3-TFO2	^{m}CT^{m}CT**^{m}C**TT	80.0	42.9/64.4			7.5

[a] 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM MES buffer. [b] 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM PIPES buffer. [c] Shown in boldface type. [d] Uncertainty ±1 °C. [e] Uncertainty ±0.1. [f] Triplex−coil transition. [g] Triplex−duplex and duplex−coil transition, respectively.

NMR Measurements. NMR experiments were performed on a Bruker AMX500 spectrometer equipped with 5 mm probes. For one-dimensional ^{1}H NMR spectra a 1-1 echo scheme (*16*) with gradient pulses during the delays to increase water suppression and a spectral width of 12.8 kHz was used. The carrier was placed on the solvent resonance and the delay between the first two 90° pulses was set to 60 μs for maximum excitation in the imino proton region of the spectrum. Phase-sensitive ^{1}H−^{15}N gradient-selected HSQC spectra (*17*) were recorded with WALTZ16 decoupling during acquisition. A total of 400 t_1 increments were collected in the States−TPPI mode with the sweep width along the ^{15}N dimension set to 70 ppm.

All data were processed on a SGI Indy R4600 workstation using the XWIN NMR software package. Prior to Fourier transformation, HSQC data sets were apodized with exponential and Gaussian window functions in ω_2 and ω_1, respectively. The final processed matrix size was 1K × 1K in real points. ^{1}H chemical shifts were referenced relative to HDO (δ_H = 4.96 at 10 °C) and ^{15}N chemical shifts relative to external $^{15}NH_4Cl$ in 10% HCl (δ_N = 0).

Lyophilized samples (130−260 OD) were dissolved in 90% H_2O/10% D_2O containing 100 mM NaCl and 10 mM $MgCl_2$. The pH was adjusted by addition of NaOH or HCl and measured before and after the experiment. Spectra obtained at increasing and decreasing pH were identical in each case.

RESULTS

The two oligonucleotides TFO-1 and TFO-2 were designed to allow the formation of an intramolecular triple helix with three C·GC and four T·AT base triplets by folding back twice on themselves via two T_4 tetraloops (Figure 1). As observed previously, UV melting experiments on the intramolecular triplexes at pH 5.0 only show a single cooperative absorbance change reflecting a direct triplex to coil transition. In contrast, at neutral pH the triplexes melt in two steps. The low-

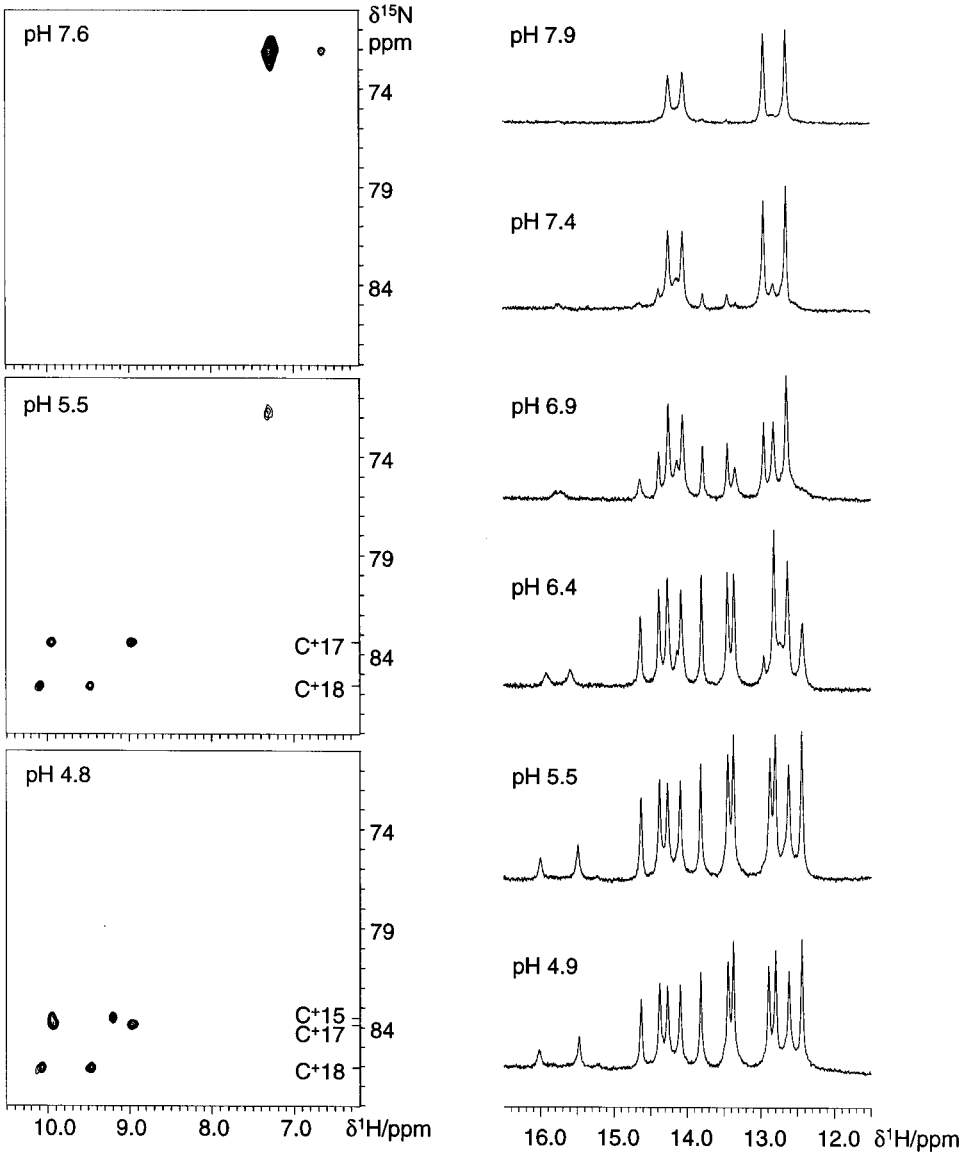

FIGURE 2: (Right) ^{1}H NMR spectra of TFO-1 at 10 °C as a function of pH showing the imino proton spectral region. (Left) ^{1}H−^{15}N HSQC spectra of TFO-1 at 10 °C at different pH values. ^{15}N assignments are indicated at the right of the spectra.

temperature transition can be assigned to the melting of the triplex into a hairpin, and the second transition, to the melting of the hairpin duplex into a single-stranded structure (*10*, *18*−*20*). The melting temperatures T_m given in Table 1 also reveal a less stable TFO-1 triplex as compared to TFO-2 in acidic as well as neutral solution conditions. This can be attributed to the presence of neighboring cytosine bases in the third strand of TFO-1 (*19*, *21*−*23*). Note, however, that destabilization is far more pronounced at neutral pH with a drop in T_m from TFO-2 to TFO-1 of 12 °C as compared to 2 °C at pH 5.0.

To follow protonation equilibria we have performed homo- and heteronuclear NMR studies on the two oligonucleotides after having specifically introduced ^{15}N isotopes at the 4-amino group of the three Hoogsteen-bound cytosine bases. Although not the actual protonation site, this label has been shown to serve as a convenient probe for cytosine protonation with an accompanying downfield shift of the ^{15}N amino resonance by more than 10 ppm (*12*).

^{1}H NMR spectra of TFO-1 showing imino resonances of the oligonucleotide at 10 °C are plotted as a function of pH in Figure 2 (right). Below pH 5.5, formation of an intramolecular triplex is indicated by the signals of Watson−Crick and Hoogsteen hydrogen-bonded imino protons between 12 and 16.5 ppm. NOESY measurements at low pH confirm the formation of a triple-helical structure and also allow the unambiguous assignment of all imino as well as amino resonances (data not shown). With increasing pH, additional signals of a duplex emerge at about pH 6.9 that are of similar intensity compared to those of the triplex. The presence of triplex and partially unfolded hairpin duplex in slow exchange on the NMR time scale is common for triplexes (*12*, *14*, *24*). Above pH 7.4, resonances of the triple helix disappear and only a reduced set of imino signals expected for a duplex structure remains observable.

We have acquired a series of ^{1}H−^{15}N HSQC spectra over a pH range of 4.8−7.6. Representative spectra are also shown in Figure 2 (left). Cross-peaks connect the amino nitrogens of the three labeled cytosines with their attached amino protons. Downfield- and upfield-shifted ^{1}H resonances are associated with hydrogen-bonded and non-hydrogen-bonded protons, respectively. From the assignment of the amino

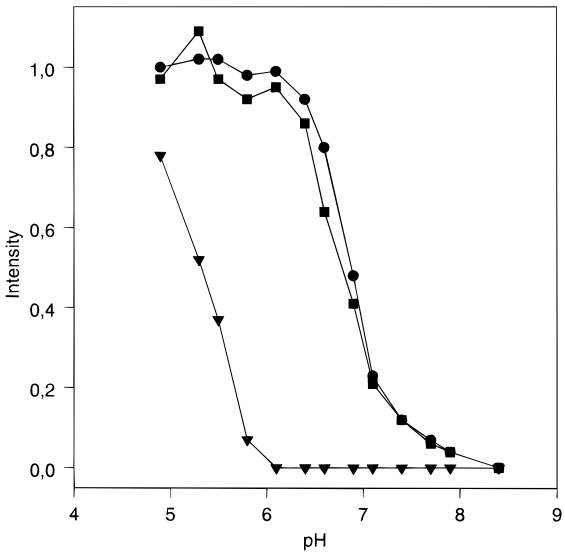

FIGURE 3: $^{1}H-^{15}N$ HSQC cross-peak intensities for protonated third strand cytosines C15 (▼), C17 (■), and C18 (●) of TFO-1 as a function of pH. The intensity of the C18 cross-peak at pH 4.9 was arbitrarily set to 1.0.

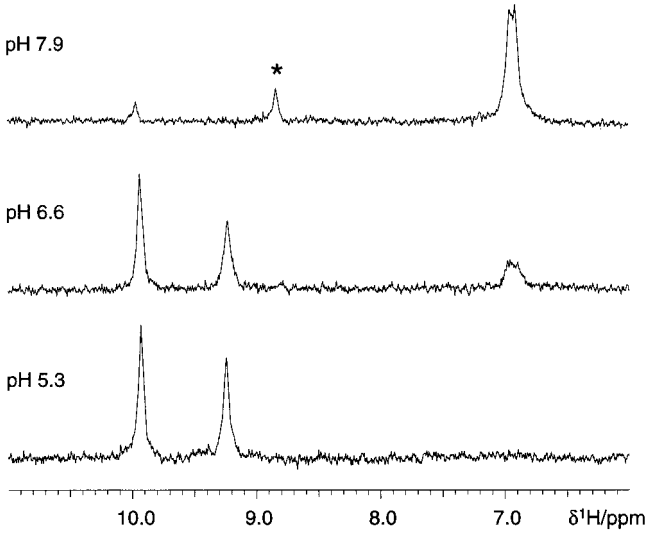

FIGURE 4: Representative one-dimensional $^{1}H-^{15}N$ HSQC spectra of m1-TFO-2. The spectra were acquired at 10 °C with different pH values and show resonances of the ^{m}C amino protons. On the basis of the corresponding ^{15}N chemical shift, the signal appearing at higher pH and marked by an asterisk can be assigned to an unprotonated ^{m}C of an unknown structure.

protons at low pH, the three nitrogen signals with chemical shifts of 83.5, 83.9, and 86.1 ppm at pH 4.80 can be unambiguously attributed to the terminal cytidine C15 and the central cytidines at positions 17 and 18, respectively. The significantly downfield-shifted ^{1}H and ^{15}N amino resonances indicate that at pH 4.8 all of the three labeled cytidines are essentially in their protonated state. As seen from Figure 2, cross-peaks due to protonated cytidine 15 at the triplex terminus have disappeared upon raising the pH from 4.8 to 5.5. In contrast, resonances of protonated cytidines 17 and 18 remain observable up to pH 7 (not shown) but have also disappeared at pH 7.6. At the same time, increasingly strong cross-peaks at higher field typical of nonprotonated cytosine bases appear with ^{1}H and ^{15}N chemical shifts clustered at 7.3, 6.6, and 72.1 ppm.

Because no noticeable changes in chemical shifts and signal line widths are observed with pH, protonated and nonprotonated bases slowly exchange on the NMR chemical shift time scale, setting an upper limit of 500 s^{-1} for the overall exchange rate. Thus, by examining corresponding signal intensities in the NMR spectra the extent of deprotonation of individual bases can be followed directly as a function of increasing pH. Due to the partial or complete overlap of amino proton signals in the 1D NMR spectra, the well-resolved HSQC cross-peak intensities were evaluated and plotted in Figure 3 for the cytosine bases of TFO-1. Note that only cross-peak intensities for the protonated bases are given. Cross-peak intensities for nonprotonated cytidines are not suited for a quantitative evaluation because of overlap and their chemical shift being close to the water resonance.

Apparently, deprotonation for the terminal and the two central cytidines takes place at very different pH values. Whereas C15 is already partially deprotonated at pH 5.0 and fully deprotonated at pH 6.0, deprotonation at C17 and C18 does not occur until pH 6.4. Semiprotonation points at half-intensity for the two central cytosines at about pH 6.8 are the same within experimental error. Corresponding NMR experiments on TFO-2 also indicate about equal apparent pK_as for the two internal cytosines C17 and C19 but with a shift to higher values (data not shown). The pH values for semiprotonation as determined for the individual cytosines in the two triplexes are summarized in Table 1.

Effect of Cytosine Methylation. To study the influence of methyl substituents on protonation and pH-dependent triplex stabilities, cytosine bases in the third strand of the triplex-forming oligonucleotides TFO-1 and TFO-2 were replaced by 5-methylcytosine (^{m}C) to give oligomers m3-TFO-1, m1-TFO-2, and m3-TFO-2 (see Figure 1).

UV melting temperatures for the three methylated oligomers m3-TFO-1, m1-TFO-2, and m3-TFO-2 at pH 5.0 and 7.0 are summarized in Table 1. As expected from previous studies, the substitution of ^{m}C for C increases the thermal stability of the intramolecular triplexes (*9*, *21*, *25*−*27*). Thus, with one ^{m}C substitution the triplex−coil transition of m1-TFO-2 at pH 5.0 increases by about 2 °C and the triplex is even more stabilized by about 4 °C at pH 7.0. Replacing all three cytosines in the third strand by ^{m}C results in a further triplex stabilization of about 6 and 12 °C at pH 5.0 and 7.0, respectively. A similar increase in melting temperatures is observed for m3-TFO-1 when compared to TFO-1.

To follow protonation equilibria on the substituted pyrimidines by NMR, a single [4-^{15}N]-5-methylcytosine was introduced at position C18/19 of the three substituted oligonucleotides (Figure 1). With only one methylcytosine ^{15}N-labeled at the amino nitrogen, a one-dimensional gradient-selected $^{1}H-^{15}N$ HSQC experiment suppresses resonances other than signals of the amino protons scalar coupled to the ^{15}N nucleus of the methylated base. As seen from Figure 4, the intensity of the two downfield-shifted hydrogen-bonded and non-hydrogen-bonded $^{m}C^{+}$ amino resonances between 9 and 10 ppm reduces on increasing the pH, while at the same time increasingly strong upfield-shifted amino resonances typical of nonprotonated ^{m}C nucleobases appear around 7 ppm. Again, because of slow exchange pH-dependent signal intensities of the downfield-shifted amino resonance can be used to construct a titration curve allowing

the determination of the pH value for semiprotonation. This apparent pK_a value for ^{m}C18/19 of m3-TFO-1, m1-TFO-2, and m3-TFO-2 is also presented in Table 1. Apparently, the methyl substituent slightly increases pK_a^{app} by 0.1−0.2 pH unit. From pH-dependent ^{15}N chemical shifts of free [4-^{15}N]-2′-deoxycytidine and [4-^{15}N]-5-methyl-2′-deoxycytidine, pK_a values for the two nucleosides of 4.2 and 4.4 were determined, respectively (data not shown). These values are in good agreement with pK_a values of 4.3 and 4.4 reported for the unsubstituted and 5-methyl-substituted cytidine (*28*). Thus, pK_a changes observed upon methylation in the oligonucleotides essentially correspond to differences between the free cytosine and 5-methylcytosine base.

DISCUSSION

The evaluation of individual equilibria for cytosine protonation is expected to give a more detailed understanding of pH and sequence-dependent triplex stabilities as determined by UV denaturation experiments. Following NMR signal intensities of specifically ^{15}N-labeled cytosines over a wide pH range allows the extraction of semiprotonation points for all three individual cytosine bases in TFO-1 and TFO-2. Apparent pK_a values for intermolecular triplexes containing C$^+$·GC triplets are typically in the range between 5.0 and 6.0 (*8, 9*), whereas intramolecular triplexes display apparent pK_a values closer to 7.0 and are thus significantly shifted to higher pH as compared to a pK_a of about 4.2 for free cytosine (*10, 11*). Correspondingly, semiprotonation points for the central but not for the terminal cytosines as obtained from the NMR titrations described above range from 6.8 to 7.4 (see Table 1). In recent studies on intramolecular triple helices having a single internal C$^+$·GC base triplet, protonated cytosines were observed in the NMR spectra even at pH > 8 with an apparent pK_a estimated to be >9 (*24*). Such high pK_a values may result from the stabilization at high pH expected for a triplex with a single C$^+$·GC triplet when compared to sequences with more cytosine bases in the third strand.

It is generally accepted that deprotonation in base pairs requires the disruption of corresponding hydrogen bonds (*29*−*31*). Ignoring triplex−duplex equilibria with complete dissociation of the Hoogsteen-bound strand, cytosine deprotonation in a triple helix can be represented by a two-state process:

$$\underset{\text{closed state}}{[\text{CH}\cdot\text{GC}]^+_{\text{T}}} \overset{K_{\text{diss}}}{\rightleftharpoons} \underset{\text{open state}}{[\text{CH} + \text{GC}]^+_{\text{T}}} \overset{K_a}{\rightleftharpoons} [\text{C} + \text{GC}]_{\text{T}} + \text{H}^+ \quad (1)$$

Local opening of the base pair with a dissociation constant K_{diss} is followed by proton abstraction from the open state described by the acid dissociation constant K_a. Consequently, the ratio of unprotonated and protonated species is given by

$$\frac{c([\text{C} + \text{GC}]_{\text{T}})c(\text{H}^+)}{c([\text{CH}\cdot\text{GC}]^+_{\text{T}})} = K_{\text{diss}}K_a \quad (2)$$

Base pair dissociation constants in DNA duplexes have been estimated to be about 10^{-5} (*30*). Thus, the protonated open state is not expected to be significantly populated and the pH of semiprotonation, pK_a^{app}, can be written as

$$\text{p}K_a^{\text{app}} = -\log K_{\text{diss}} + \text{p}K_a \quad (3)$$

Assuming that the cytosine residues in an open state are not strongly perturbed by adjacent nucleotides, Hoogsteen base pair dissociation constants K_{diss} can be estimated by use of pK_a = 4.2 for monomeric cytosine (vide supra). From the apparent pK_a values listed in Table 1, dissociation constants of about 10^{-3} to 10^{-1} are derived for C$^+$·G Hoogsteen base pairs at central and terminal positions of the intramolecular triple helix, respectively. Clearly, fraying at the triplex termini will result in increased base pair dissociation constants, and hence lower semiprotonation points are expected for terminal cytosines. Although the estimates for K_{diss} are necessarily imprecise given the various simplifications, the results indicate larger dissociation constants for Hoogsteen base pairs as compared to the reported values for Watson−Crick base pairs. Correspondingly, lifetimes of Hoogsteen base pairs have been found to be shorter than that of the Watson−Crick pairs contained in the same triplet (*32*).

It is well-known that at low pH a triple helix with a sequence rich in C·GC triplets is thermally more stable than a triple helix of the same length containing solely T·AT triplets. Under acidic conditions third-strand cytosines are protonated, forming two Hoogsteen hydrogen bonds with a Watson−Crick GC base pair (*14, 33*−*35*). The increased stability of a C$^+$·GC base triplet, relative to a T·AT triplet can be attributed to the formation of stronger hydrogen bonds and to favorable electrostatic interactions between the positive charge on a protonated cytosine and the negatively charged sugar−phosphate backbone (*36*). In contrast, cytosine deprotonation at higher pH leads to an unprotonated C·GC base triplet, which is less stable than a T·AT triplet. Consequently, the degree of cytosine protonation largely determines the stability of a C·GC triplet. The titration curves for TFO-1 in Figure 3 indicate that at pH 7.0 the terminal cytosine C15 is fully deprotonated and the population of protonated cytosines C17 and C18 has diminished to about 30% at 10 °C. Hence, C·GC triplets exert a destabilizing effect on the triplex at neutral pH for all positions. In contrast, both internal cytosines are fully protonated at pH 5.0, resulting in a stabilization relative to T·AT base triplets. The less stabilizing effect of a terminal cytosine residue as observed in triple helices at acidic pH has been ascribed to the poorer stacking interaction with only one nearest neighbor at the 3′- or 5′-end (*24*). However, our data indicate that partial deprotonation even at acidic pH for the cytosine at the terminus is most likely reponsible for the observed end effects.

Electrostatic repulsion of adjacent protonated cytosines has been attributed to the observed decrease in T_m for a triplex with contiguous cytosine bases (*19, 22*). In addition, theoretical studies indicate much smaller base stacking energies in C_n sequences as compared to alternate $(CT)_n$ sequences (*36*). Unfavorable electrostatic interactions only occur between cytosines carrying a positive charge and hence are expected to decrease with progressive deprotonation at higher pH values. However, destabilization by adjacent cytosine bases is found to increase in neutral solution. Obviously, effects other than direct electrostatic interactions are responsible for the enhanced destabilization of adjacent cytosines at pH 7.0. As seen from the titration curves in Figure 5, the semiprotonation point for the central cytosine C17 in TFO-1 is

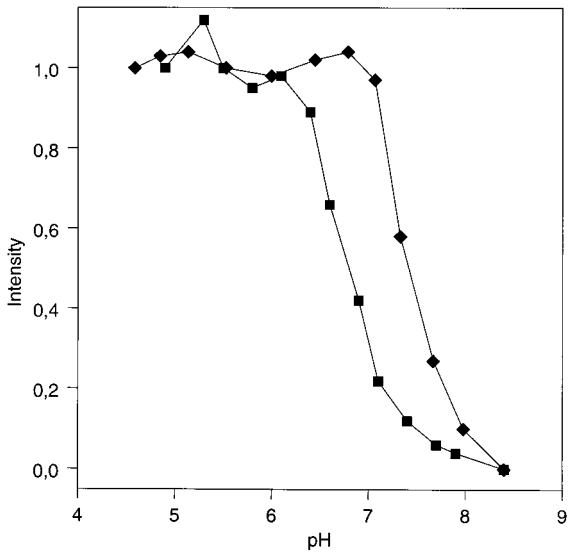

FIGURE 5: $^{1}H-^{15}N$ HSQC cross-peak intensity for the protonated third-strand cytosine C17 of TFO-1 (■) and TFO-2 (◆) as a function of pH. The intensity of the cross-peak at lowest pH was set to 1.0 for each of the oligomers.

considerably lowered by 0.6 pH unit as compared to the isolated cytosine C17 in TFO-2, i.e., adjacent cytosines disfavor protonation. Whereas central cytosines of both sequences are fully protonated at pH 5.0, there are large differences in the extent of protonation at pH 7.0 due to the steep slope of the titration curves between pH 6.4 and 7.9. Apparently, destabilization by adjacent cytosines at higher pH is mostly due to the bases being protonated to a lesser degree.

Effect of Cytosine Methylation on Triplex Stability. It has been shown that incorporation of 5-methylcytosine into the third strand of polynucleotide and oligonucleotide triplexes increases both the thermal stability and the upper pH limit for triple-helix stability (*9*, *21*, *25*−*27*). However, major sources of the stabilizing influence of a 5-methyl substituent is still a point at issue. An electron-donating methyl group is expected to increase the pK_a of the base, thus shifting its protonation to higher pH values while reducing the hydrogen-bond donating ability of the protonated nitrogen. Differences in the pK_as of cytosine and 5-methylcytosine in nucleosides and single-stranded oligonucleotides were determined to be 0.2 pH unit (vide supra) or even less (*9*, *28*). However, the electrostatic potential within a triple-helical complex may significantly affect acidities compared to the monomer. Therefore, pH values of semiprotonation for 5-methylcytosine within the third strand of triplex-forming oligonucleotides have been determined. Differences between cytosine and 5-methylcytosine in the triplex amount to 0.1−0.2 pH unit (see Table 1) and thus correspond to differences observed for the free nucleosides. Similarly, methyl substituents at C5 of cytosine in an oligonucleotide targeted against a 15 base pair sequence were shown to increase the apparent pK_a of the triplex by 0.2 unit (*26*). In contrast, an increase in the apparent pK_a of 1.0 unit was determined for triplexes formed upon binding of an 11-mer with C → ^{m}C substitutions to an oligonucleotide duplex target (*9*). This apparent discrepancy in pK_a changes may be attributed to the methods used to measure the pK_a or to sequence-dependent effects (*26*). It should also be noted that pK_a values determined in the latter two studies describe the effect of pH on the global triplex−duplex transition rather than an intrinsic pK_a at a single atom.

Clearly, even the small increase in apparent pK_as as determined for the intramolecular triplexes with ^{m}C substitutions will result in triplex stabilization, especially at higher pH. However, additional effects are likely to contribute to the significant increase in thermal stabilization upon incorporation of 5-methylcytosine. This is also corroborated by the observed stabilizing effect of methyl substituents at pH 5.0 where nonterminal third-strand cytosines are expected to be completely protonated (see Figure 3).

Excluding extensive cancellation by opposing effects of the methyl substituent on base pair opening and true protonation equilibria, our results suggest that (i) pK_a differences between cytosine and 5-methylcytosine are largely maintained in the triple-helical complex and (*2*) the methyl substituent does not significantly affect the stability of a Hoogsteen base pair. Also, protonation of the 5-methylcytosine base is unaffected by additional methyl substituents within the third strand, as indicated by the pK_a^{app} for ^{m}C19 in m1-TFO-2 and m3-TFO-2.

Consequently, any other stabilizing interaction by the methyl group other than shifts in protonation equilibria will similarly affect protonated and unprotonated triplexes. There are two likely sources for additional stabilization, namely, increased stacking energies or hydrophobic effects. Stacking energies arise from dispersion and induced dipole interactions and thus correlate with the polarizabilities of the interacting heterocycles. In fact, substitution of a methyl group at C5 of cytosine has been shown to increase its molecular polarizability with a concomitant increase of the free energy for stacking of cytosine with adenine in water by 0.2 kcal·mol^{-1} (*37*). Methylation may also induce a release of water molecules from the helix, thereby contributing a positive entropy change (*9*). However, in line with previous studies (*27*), favorable contacts between adjacent methyl groups has only a minor effect on thermal stabilization as indicated by the similar increase in melting temperatures for m3-TFO-1 bearing two adjacent 5-methylcytosines and m3-TFO-2 (Table 1).

ACKNOWLEDGMENT

We are grateful to R. Zander for his help in the synthesis of isotope-labeled nucleosides.

REFERENCES

1. Felsenfeld, G., Davies, D. R., and Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023−2024.
2. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. M., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495−497.
3. Htun, H., and Dahlberg, J. E. (1989) *Science 243*, 1571−1576.
4. Maher, L. J., III., Wold, B., and Dervan, P. B. (1989) *Science 245*, 725−730.
5. Hanvey, J. C., Shimizu, M., and Wells, R. D. (1990) *Nucleic Acids Res. 18*, 157−161.
6. Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E., and Dervan, P. B. (1991) *Science 254*, 1639−1642.
7. Thuong, N. T., and Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 666−690.
8. Callahan, D. E., Trapane, T. L., Miller, P. S., Ts'o, P. O. P., and Kan, L.-S. (1991) *Biochemistry 30*, 1650−1655.

9. Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., and van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.
10. Plum, G. E., and Breslauer, K. J. (1995) *J. Mol. Biol. 248*, 679−695.
11. Osborne, S. E., Cain, R. J., and Glick, G. D. (1997) *J. Am. Chem. Soc. 119*, 1171−1182.
12. Leitner, D., Schröder, W., and Weisz, K. (1998) *J. Am. Chem. Soc. 120*, 7123−7124.
13. Gaffney, B. L., Kung, P.-P., Wang, C., and Jones, R. A. (1995) *J. Am. Chem. Soc. 117*, 12281−12283.
14. Sklenár, V., and Feigon, J. (1990) *Nature 345*, 836−838.
15. Kupferschmitt, G., Schmidt, J., Schmidt, Th., Fera, B., Buck, F., and Rüterjans, H. (1987) *Nucleic Acids Res. 15*, 6225−6241.
16. Sklenár, V., and Bax, A. (1987*) J. Magn. Reson. 74*, 469−479.
17. Sklenár, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson., Ser. A 102*, 241−254.
18. Völker, J., Botes, D. P., Lindsey, G. G., and Klump, H. H. (1993) *J. Mol. Biol. 230*, 1278−1290.
19. Völker, J., and Klump, H. H. (1994) *Biochemistry 33*, 13502−13508.
20. Plum, G. E. (1997) *Biopolymers 44*, 241−256.
21. Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., and Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603−6614.
22. Jayasena, S. D., and Johnston, B. H. (1992) *Nucleic Acids Res. 20*, 5279−5288.
23. Koh, J. S., and Dervan, P. B. (1992*) J. Am. Chem. Soc. 114*, 1470−1478.
24. Asensio, J. L., Lane, A. N., Dhesi, J., Bergqvist, S., and Brown, T. (1998) *J. Mol. Biol. 275*, 811−822.
25. Povsic, T. J., and Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059−3061.
26. Singleton, S. F., and Dervan, P. B. (1992) *Biochemistry 31*, 10995−11003.
27. Wang, S., and Kool, E. T. (1995) *Biochemistry 34*, 4125−4132.
28. Hall, R. H. (1971) in *The Modified Nucleosides in Nucleic Acids*, pp 192−194, Columbia University Press, New York.
29. Lefevre, J.-F., Lane, A. N., and Jardetzky, O. (1985) *J. Mol. Biol. 185*, 689−699.
30. Guéron, M., Kochoyan, M., and Leroy, J.-L. (1987) *Nature 328*, 89−92.
31. Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., and Guéron, M. (1988) *J. Mol. Biol. 200*, 223−238.
32. Cain, R. J., and Glick, G. D. (1998) *Biochemistry 37*, 1456−1464.
33. de los Santos, C., Rosen, M., and Patel, D. (1989) *Biochemistry 28*, 7282−7289.
34. Macaya, R., Wang, E., Schultze, P., Sklenár, V., and Feigon, J. (1992) *J. Mol. Biol. 225*, 755−773.
35. Asensio, J. L., Brown, T., and Lane, A. N. (1998) *Nucleic Acids Res. 26*, 3677−3686.
36. Sun, J.-s., Mergny, J.-L., Lavery, R., Montenay-Garestier, T., and Hélène, C. (1991) *J. Biomol. Struct. Dyn. 9*, 411−424.
37. Sowers, L. C., Shaw, B. R., and Sedwick, W. D. (1987) *Biochem. Biophys. Res. Commun. 148*, 790−794.

BI992630N