Electrostatic Effects in DNA Triple Helices†

Jens Völker‡ and Horst H. Klump*

Department of Biochemistry, University of Cape Town, Private Bag, Rondebosch 7700, Republic of South Africa

Received May 5, 1994; Revised Manuscript Received August 24, 1994®

ABSTRACT: Electrostatic effects dominate many aspects of nucleic acid behavior in a sequence independent manner. Sequence dependent electrostatic effects are introduced when a polypyrimidine, which contains one or more protonated cytosines, binds in the major groove (Hoogsteen side) of a complementary Watson−Crick double helix. Depending on the number of cytosines in the third strand (global effect) and on their relative position (local effect), the cytosines either enhance or decrease the binding affinity of the third strand, because adjacent protonated cytosines destabilize the third strand binding compared to cytosines separated by intervening thymines. This local effect (crowding) can reverse the effect of global composition. To investigate the extent of the local and global electrostatic effects further, two families of oligonucleotides have been synthesized. They share as a common design feature that they all fold sequentially into isosterical intramolecular triple helices by way of hairpin intermediates. This is confirmed by P_1 nuclease probing, CD spectroscopy, and UV spectroscopy. The thermal stability of these conformations depends on the sequences, pH, and the ionic strength and can be summarized as follows: The energy of third strand binding depends on the protonated cytosine content in the Hoogsteen strand. It increases with increasing cytosine content (global composition) below pH 7.1 (150 mM Na^+), decreases above pH 7.1, and is independent of the cytosine content at pH 7.1. At pH 6.75 the energy of binding increases with increasing cytosine content below 400 mM Na^+, decreases above 400 mM Na^+, and is independent of the global composition at 400 mM Na^+.

Triple helical RNA complexes have been known since 1957 (Felsenfeld et al., 1957). Subsequently, it was established that ribo- as well as deoxyribohomopyrimidine polymers can associate with complementary double helical homopurine−homopyrimidine complexes to form triple helices (Riley et al., 1966; Lipsett, 1964; Thiele & Guschlbauer, 1968; Morgan & Wells, 1968; Arnott & Bond, 1973a; Arnott & Selsing, 1974; Felsenfeld & Miles, 1967; Michelson et al., 1967). Triple helices which contain homopurine strands as third strands also have been observed (Rich, 1958; Howard & Miles, 1977; Thiele et al., 1978; Marck & Thiele, 1978; Arnott & Bond, 1973b; Broitman et al., 1987; Letai et al., 1988; Cheng & Pettitt, 1992).

The homopyrimidine third strand binds via Hoogsteen base pairing (Hoogsteen, 1959) to the purine bases in the major groove of the Watson−Crick double helix with thymines and/or uracils recognizing only A·T base pairs and protonated cytosines recognizing only G·C base pairs (Arnott & Bond, 1973a; Arnott & Selsing, 1974; Arnott & Bond, 1973b; Letai et al., 1988; Arnott et al., 1976). The pyrimidine third strand is oriented in parallel to the Watson−Crick homopurine strand (Arnott et al., 1976; Hattori et al., 1976; Moser & Dervan, 1987). Initially it was suggested that a change from a B-DNA to an A-DNA conformation occurs on triple helix formation (Arnott et al., 1976), but more recent evidence points to the formation of an altered B-DNA conformation instead (Macaya et al., 1992a; 1992b; Park & Breslauer, 1992; Raghunathan et al., 1993; Howard et al., 1992).

Oligonucleotide-based triple helix formation has been considered as a general approach for the recognition of double stranded DNA (Moser & Dervan, 1987). It will have implications for example in chromosome mapping and the inhibition of transcription *in vitro* and *in vivo* (Strobel & Dervan, 1990; Strobel & Dervan, 1991; Strobel et al., 1991; Maher, 1992; Riordan & Martin, 1991, Helene & Toulme, 1990). Because of the importance of these potential applications, a large number of studies have attempted to shed light on the thermodynamic (Manzini et al, 1990; Xodo et al., 1990; Pilch et al., 1990, Pilch et al., 1991; Xodo et al., 1991; Plum et al., 1990, Völker et al., 1993) and kinetic (Singleton & Dervan, 1992a; Singleton & Dervan, 1992b; Maher et al., 1990; Rougee et al., 1992;) aspects of triple helix formation and to determine the effect of mismatches (Griffin & Dervan, 1989, Macaya et al., 1991; Mergny et al., 1991; Horne & Dervan, 1991; Wang et al., 1992; Radhakrishnan & Patel, 1992; Radhakrishnan et al., 1991) and chemically modified bases (Lee et al., 1984; Povsic & Dervan, 1990; Ono et al., 1991; Griffin et al., 1992a; Griffin et al., 1992b; Milligan et al., 1993; Stilz & Dervan, 1993; Egholm et al., 1992; Froehler & Ricca, 1992) on the stability of triple helices. There are discrepancies in the current literature between the calorimetrically determined enthalpy values and those derived from UV-melting curves (Manzini et al., 1990; Xodo et al., 1990; Pilch et al., 1990, Pilch et al., 1991; Xodo et al., 1991; Plum et al., 1990, Völker et al., 1993).

In analogy to double-stranded DNA (Klump, 1988; Breslauer et al., 1986), it has become customary to expect that the binding enthalpy of the third strand is determined by the

† This work was supported by the Foundation for Research and Development FRD (Pretoria, RSA) and the University of Cape Town (Rondebosch, RSA). J.V. acknowledges the generous support by the Gottlieb Daimler und Carl Benz Stiftung (Ladenburg, Germany).

* To whom correspondence should be addressed. Phone: Cape Town 650 2407. Fax: Cape Town 685 5931.

‡ Current address: Department of Chemistry, Rutgers, The State University of New Jersey, New Brunswick, NJ 08903.

® Abstract published in *Advance ACS Abstracts,* October 15, 1994.

0006-2960/94/0433-13502$04.50/0

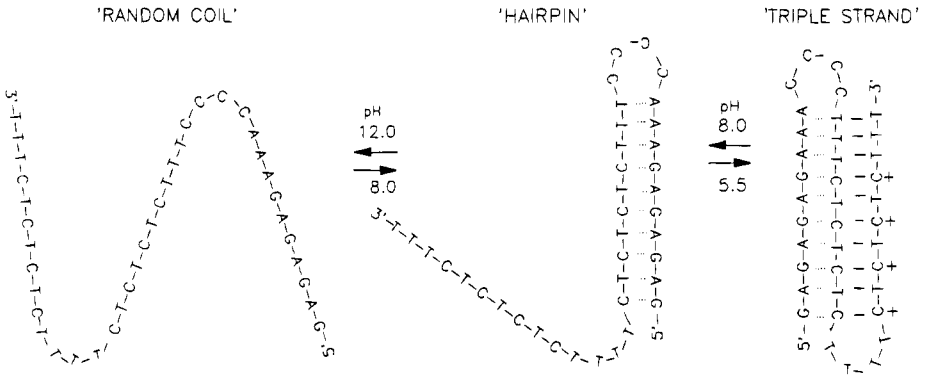

FIGURE 1: The proposed pH-dependent folding pathway from a random coil into a hairpin with 3′ extension into an intramolecular triple helix shown for the oligonucleotide ITS-4G_1. (•••) denotes the Watson–Crick hydrogen bonds, while (—) denotes the Hoogsteen hydrogen bonds.

Table 1: The Two Families of Oligonucleotides[a]

NAME	SEQUENCE
[A]	**Global Variation**
ITS-ATT	5'-A-A-A-A-A-A-A-A-A-©©©©-T-T-T-T-T-T-T-T-T-T-©©©©-T-T-T-T-T-T-T-T-T-3'
ITS-2G_3	5'-G-A-A-A-G-A-A-A-A-A-©©©©-T-T-T-T-T-C-T-T-T-C-ŦŦŦŦ-C-T-T-T-C-T-T-T-T-T-3'
ITS-3G_2	5'-G-A-A-A-G-A-A-G-A-A-©©©©-T-T-C-T-T-C-T-T-T-C-ŦŦŦŦ-C-T-T-T-C-T-T-C-T-T-3'
ITS-4G_1	5'-G-A-G-A-G-A-G-A-A-A-©©©©-T-T-T-C-T-C-T-C-T-C-ŦŦŦŦ-C-T-C-T-C-T-C-T-T-T-3'
[B]	**Local Variation**
ITS-3G_0	5'-G-A-A-A-G-G-A-A-A-A-©©©©-T-T-T-T-C-C-T-T-T-C-ŦŦŦŦ-C-T-T-T-C-C-T-T-T-T-3'
ITS-3G_1	5'-G-A-A-A-G-A-G-A-A-A-©©©©-T-T-T-C-T-C-T-T-T-C-ŦŦŦŦ-C-T-T-T-C-T-C-T-T-T-3'
ITS-3G_2	5'-G-A-A-A-G-A-A-G-A-A-©©©©-T-T-C-T-T-C-T-T-T-C-ŦŦŦŦ-C-T-T-T-C-T-T-C-T-T-3'

[a] ITS-4G_1 corresponds to JV-ITS in *J. Mol. Biol.* (1993) *230*, 1278–1290.

sum of the nearest neighbor interaction energies (Cheng & Pettitt, 1992; Kiessling et al., 1992). Hoogsteen base pairing requires the same number of hydrogen bonds for thymines as for protonated cytosines. Sequence-dependent differences must come from the effect of nearest neighbor stacking interactions. In contrast to this current belief the experimental results presented here show that electrostatic effects are predominantly responsible for the magnitude of the binding enthalpy of the third strand. These effects combine the interactions of the three negatively charged phosphate backbones (sequence independent effect) with the interaction of the protonated cytosines and the backbones (global effect), and the effect of adjacent protonated cytosines upon each other (local effect). The local electrostatic effect will obviously be strongly dependent on the sequence in the third strand, while backbone electrostatic effects are not.

To discriminate between the global and local electrostatic contributions to the stability of triple helical DNA we have synthesized two sets of oligonucleotides which fold sequentially first into hairpins with single stranded 3′ extensions and then into intramolecular triple helices (Figure 1) (Völker et al., 1993). The folding of oligonucleotides into intramolecular triple helices of this type was first described by Sklenar & Feigon (Sklenar & Feigon, 1990; Macaya et al., 1992b). The sequences used in this study were designed to allow either changes in the global composition of the oligonucleotides while trying to exclude local interactions (Table 1A) or to allow for localized changes at constant global composition (Table 1B). Intramolecular folded sequences provide an excellent experimental system for the study of the stability of triple helices since the three strands always occur in the correct stoichiometric amounts and the orientation of the strands with respect to each other is defined by the choice of the sequences. In addition the formation of competing structures is disfavored. In the following we shall derive the importance of the electrostatic

interactions from a comparison of the melting behavior of the different conformations of the six oligonucleotides listed in Table 1A,B.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. The oligonucleotides listed in Table 1A,B were synthesized by the conventional phosphoramidate procedure on an Autogen 6500 DNA synthesizer (Milligen/Biosearch, Burlington, MA), modified to include double coupling reactions in order to suppress the occurrence of aborted sequences. The crude tritylated synthesis product was then purified by anion-exchange HPLC on a Pharmacia Mono Q HR5/5 column in denaturing conditions (pH 12) before recovery by ethanol/acetone precipitation and detritylation. The oligonucleotides were found to be sufficiently pure to result in a single band on a 20% denaturing polyacrylamide gel regardless of the detection procedure (Stains-all dye, UV-shadowing or ^{32}P radiolabeling).

Buffers. Spectroscopic measurements were done in a standard buffer of 20 mM sodium cacodylate/20 mM sodium acetate/10 mM sodium chloride (50 mM Na^+) at pH 8.0, pH 6.75, or pH 4.5 (buffer A). To adjust the ionic strength appropriately without diluting the samples a stock solution was prepared consisting of buffer A and NaCl to make it 5 M Na^+. Measurement of T_m vs pH were done in a standard buffer of 50 mM Na_3PO_4 (150 mM Na^+) (buffer B). The pH of buffer B was adjusted by the addition of small aliquots of dilute HCl (0.5 or 0.1 M).

P_1-Nuclease Digestion. The oligonucleotides were 5′ endlabeled with γ-^{32}P ATP and T_4-kinase (Boehringer Mannheim) by standard procedure (Ausubel et al., 1987) modified to include a 3 h incubation time and a chase with an extra 20 units of T_4-kinase after $1^1/_2$ h. For the nuclease reaction approximately 5×10^5 cpm of γ-^{32}P-labeled oligonucleotide in 50 μL of 100 mM Na^+ pH 8.0 or pH 5.5 buffer was incubated with approximately 0.3 (pH 8.0) or 0.003 (pH 5.5) units P_1-nuclease, respectively. At appropriate time intervals 5 μL samples were withdrawn and added to 5 μL of 0.5 M EDTA and 20 μL of loading buffer (80% formamide, 20% glycerol) and heated to 90 °C. Then 10 μL of the heat denatured sample was loaded onto a 20% denaturing polyacrylamide gel (20:1 acrylamide:bisacrylamide) and electrophoresed at 20 V/cm overnight.

UV-Melting Curves. Melting experiments (A_{260} vs temperature) were performed on either a Hewlett-Packard HP 8450A diode array spectrophotometer equipped with a 89100A temperature station and a 89102 temperature probe or a Pye Unicam SP1700 or SP1800 spectrophotometer (Klump, 1986) and analyzed according to the procedure described by Marky & Breslauer (Marky & Breslauer, 1987).

CD Spectroscopy. CD spectra were recorded on a JASCO J40A spectropolarimeter equipped with a thermostatically controlled quartz cuvett. The spectra between 220 and 320 nm in 100 mN Na^+ buffer A at pH 8.0, pH 6.75, or pH 4.5 were obtained at the appropriate temperatures selected from the UV-melting curves.

RESULTS

Probing of Secondary Structure with P_1-Nuclease. The oligonucleotides listed in Table 1A,B are expected to fold according to the scheme depicted in Figure 1. The hairpin differs from the intramolecular triple helix in the extent and position of single stranded regions. Digestion of the cytosine-containing oligonucleotides with a single strand specific nuclease such as P_1-nuclease at pH 8.0 results in a successive loss of bases from the 3 single-stranded extension. No digestion of the hairpin stem or loop region is observed. At pH 5.5, where the intramolecular triple helix is the thermodynamically stable conformation, only the hairpin loop and to a small extent the triplex loop are digested. The results obtained for the two families of oligonucleotides (cf. Table 1A,B) are in excellent agreement with P_1-nuclease probing reported previously (Völker et al., 1993). For the analogous oligonucleotide lacking cytosines in the stem regions (ITS-ATT), any nick in the loops results in the melting and consecutive digestion of the triplex and the hairpin stems.

CD Spectroscopy. Inspection of the CD spectra of the different conformational states of the oligonucleotides (*vide supra*) reveals the same overall properties for all six sequences. The hairpin conformation (pH 8.0) always has a positive Cotton effect between 260 and 300 nm with a maximum around 280 nm and a negative Cotton effect between 230 and 260 nm with a minimum at about 248 nm. The appearance and intensity of a characteristic shoulder at 260 nm in the spectra of some hairpins is positively correlated to the extent of the A·T tract in the hairpin helix and disappears on denaturation at 70 °C. The same shoulder is prominently shown in the CD spectrum of poly[d(A)]·poly[d(T)] (Chan et al., 1990). The comparison of the CD spectra of the intramolecular triple helical conformation (pH < 6.0) show that all oligonucleotides fold into closely related structures which are characterized by a positive ellipticity between about 260 and 300 nm with a maximum around 280 nm. The positive band is reduced in intensity compared to the same band in the hairpin spectrum, and the maximum is progressively red-shifted for the more cytosine-rich triplexes. The remaining part of the spectrum is almost identical to the spectrum of the hairpin, and the shoulder observed at 260 nm for the hairpin conformation remains unchanged in the triple helix state. There is some indication of a second minimum below 220 nm which is generally considered characteristic for triplexes (Steely et al., 1986; Callahan et al., 1991).

UV-Melting Experiments. The hairpin to coil transition (pH 8.0) is characterized by a single cooperative absorbance change as a function of temperature (Figure 2 insert). On lowering the pH to 6.75, two cooperative transitions occur. The low-temperature step, which is assigned to the triple helix to hairpin transition, shows a strong pH dependency, while the high-temperature step, which corresponds to the hairpin to coil transition, remains pH independent. Reducing the pH to values below pH 6.0 results in a single cooperative transition. This transition is a superposition of the triple helix to a hairpin and hairpin to coil transitions and consequently exhibits an intermediate pH dependency. In Figure 2 the transition temperature as a function of pH (phase diagram) for ITS-4G_1 (150 mM Na^+) is shown. The phase diagram for the other cytosine-containing oligonucleotides are all very similar. Only ITS-ATT shows an aberrant behavior in that the triple helix to hairpin and the hairpin to coil transitions do not merge as a function of pH. The two transitions only superimpose at high ionic strength. The thermodynamic properties of the three conformational transitions of all oligonucleotides are summarized in Table 2A,B.

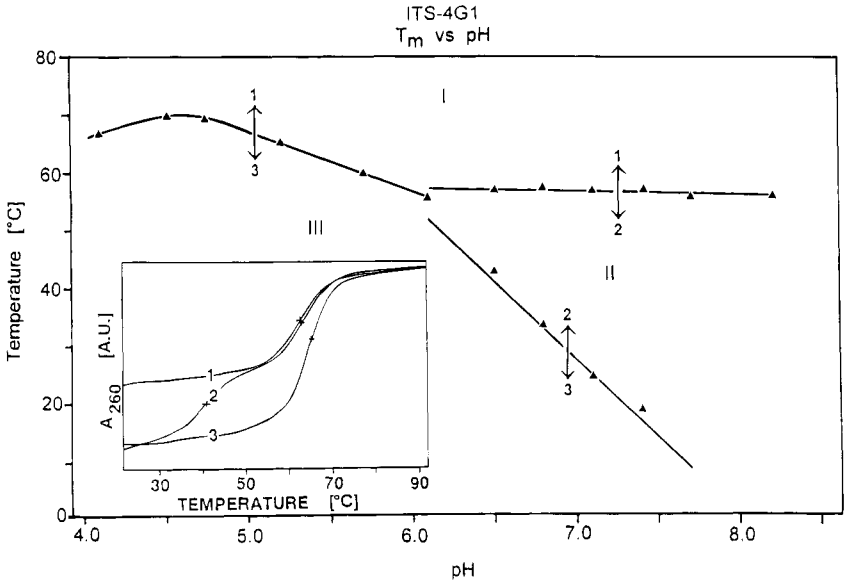

FIGURE 2: Phase diagram of T_m versus pH for the oligonucleotide ITS-4G$_1$. Region I denotes the conditions where ITS-4G$_1$ exists in the random coil state, region II where it exists in the hairpin state, and region III where ITS-4G$_1$ exists as an intramolecular triple helix. In the insert the original UV-melting curves at pH 8.0 (1), pH 6.75 (2), and pH 4.5 (3) are shown.

Table 2: Thermodynamic Parameters for the Oligonucleotides[a]

	hairpin ↔ coil (pH 8.0)		triplex ↔ hairpin (pH 6.75)				triplex ↔ coil (pH 4.5)	
	T_m (°C)	$d(T_m)/d(\log[Na^+])$ (°C)	T_m (°C)	$d(T_m)_{low}/d(\log[Na^+])$ (°C)	$d(T_m)_{high}/d(\log[Na^+])$ (°C)	$d(T_m)/d(pH)$ (°C)	T_m (°C)	$d(T_m)/d(\log[Na^+])$ (°C)
				(A) Global Variation				
ITS-ATT	50.5 ± 1.0	16.39 ± 0.56	28.3 ± 1.0	29.60 ± 1.49		−1.5 ± 0.30		21.81 ± 0.45
ITS-2G$_3$	53.8 ± 1.0	14.69 ± 0.64	32.1 ± 1.0	11.46 ± 1.25	20.66 ± 1.38	−14.1 ± 0.91	56.0 ± 1.0	17.74 ± 0.37
ITS-3G$_2$	56.8 ± 1.0	13.26 ± 0.71	37.0 ± 1.0	8.04 ± 0.68	14.43 ± 2.40	−20.7 ± 0.59	65.0 ± 1.0	15.39 ± 0.30
ITS-4G$_1$	63.0 ± 1.0	12.00 ± 0.71	41.0 ± 1.0	2.42 ± 0.25		−26.0 ± 1.15	71.0 ± 1.0	12.65 ± 0.27
				(B) Local Variation				
ITS-3G$_0$	57.8 ± 1.0	15.32 ± 0.47	31.9 ± 1.0	4.49 ± 0.16	15.50 ± 1.03	−17.3 ± 0.73	62.0 ± 1.0	13.75 ± 0.50
ITS-3G$_1$	57.7 ± 1.0	12.93 ± 0.52	36.5 ± 1.0	7.60 ± 0.56	17.75 ± 0.83	−17.7 ± 0.31	66.7 ± 1.0	14.72 ± 0.20
ITS-3G$_2$	56.8 ± 1.0	13.26 ± 0.71	37.0 ± 1.0	8.04 ± 0.68	14.43 ± 2.40	−20.7 ± 0.59	65.0 ± 1.0	15.39 ± 0.30

[a] $d(T_m)/d(\log[Na^+])_{low}$ is the slope of the plot of T_m versus $\log[Na^+]$ determined from the linear region between 50 mM Na^+ and 200 mM Na^+, while the $d(T_m)/d(\log[Na^+])_{high}$ is the slope in the linear region between 500 mM Na^+ and 1 M Na^+.

DISCUSSION

A. The Hairpin to Coil Transition: The T_m of the Hairpin Is a Straight Function of the Composition. Plotting the melting temperature of the hairpin to coil transitions at pH 8.0 (100 mM Na^+) as a function of the %GC shows that it follows the same rules as previously developed for the melting of genomic DNA (Klump, 1988; Mamur & Doty, 1962) and for dumbbell structures (Doktycz et al., 1992) (Figure 3). Only the unfolding of the hairpin of ITS-ATT (0%GC) deviates from the predicted behavior. This maybe due to the presence of an altered B-DNA conformation (B_h-DNA or A-tract DNA) (Lipanov et al., 1990; Nelson et al., 1987; Coll et al., 1987). Evidence for the formation of A-tract DNA in ITS-ATT comes from CD spectroscopy. The well-developed shoulder observed at 260 nm in the CD spectrum of ITS-ATT (spectrum not shown) is characteristic for the B_h-DNA conformation in poly[d(A)]·poly[d(T)] (Chan et al., 1990). These melting experiments ensure that the 3′ extension does not affect the thermodynamic behavior of the hairpin loop structure.

Contributions of the 3′ Extension to the Degree of Ion Condensation. To elucidate the effect of ionic strength (gross electrolyte effect) on the T_m of the hairpin to coil transition the $d(T_m)/d(\log[Na^+])$ values for all oligonucleotide systems have been determined. A comparison of this value for ITS-ATT at pH 8.0 (16.39 ± 0.56 °C) with that of the corresponding hairpin A_{10}-C_4-T_{10} (13.89 ± 0.36 °C) (HP-

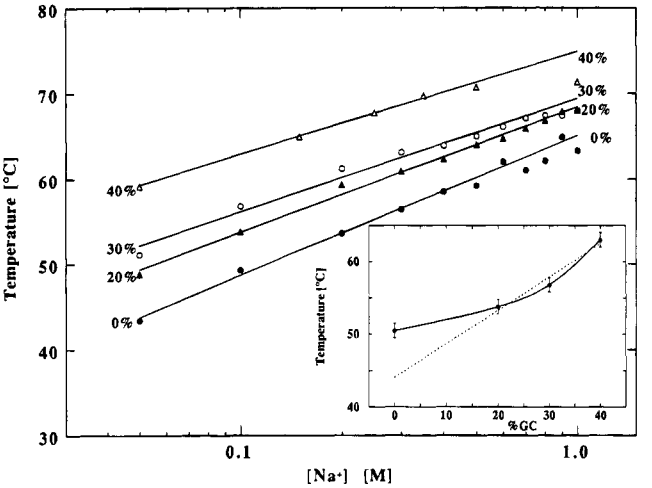

FIGURE 3: Phase diagram of T_m versus $\log[Na^+]$ (pH 8.0) for the hairpin to coil transitions of the oligonucleotides listed in Table 1A: ITS-ATT closed circle; ITS-2G$_3$, closed triangle; ITS-3G$_2$, open circle; ITS-4G$_1$, open triangle. The insert shows a plot of T_m versus %GC at 100 mM Na^+ (pH 8.0) for these oligonucleotides.

AT, unpublished observations) shows, that the single strand extension modifies the degree of ion condensation. Application of the oligoelectrolyte theory developed by Record & Lohman (Record & Lohman, 1978), calibrated with the $d(T_m)/d(\log[Na^+])$ value of HP-AT, suggests an apparent hairpin extension of four to five bases for ITS-ATT. (T_m's of a 14-mer hairpin double helix). It is reasonable to assume that the other oligonucleotide hairpins can be viewed accordingly.

The Degree of Ion Condensation Depends on the Composition of the Hairpin Stem. A comparison of the $d(T_m)/d(\log[Na^+])$ values of the different oligonucleotides reveals a linear relationship between the $d(T_m)/d(\log[Na^+])$ values and the global composition (%GC) (Figure 3). The slope is twice as steep as that observed for genomic DNA (Blake & Haydock, 1979). In accordance with the polyelectrolyte theory (Manning, 1978; Record et al., 1978), the $d(T_m)/d(\log[Na^+])$ value of the Watson–Crick double helix is directly proportional to the difference in charge density between the helix and the coil state and independent of the base composition. Composition-dependent differences, e.g., in local charge density and/or in the degree of hydration of A•T and G•C basepairs, are neglected in the theory. It is interesting to note that the change in $d(T_m)/d(\log[Na^+])$ value with the increase in cytosine content corresponds to a decrease in the intensity of the shoulder at 260 nm observed in the CD spectra which in turn correlates to the length of AT runs in the hairpin helix. It is obvious that the decrease in slope with increasing GC content corresponds directly to the decrease in the extent of A-tract DNA in the hairpin stem. A-tract DNA is known to have a reduced helical pitch and therefore higher charge density than conventional B-DNA (Nelson et al., 1987; Coll et al., 1987).

The Stability of the Hairpin Is Independent of the pH. The T_m of the hairpin to coil transition of all oligonucleotides investigated here is independent of the pH between pH 6.0 and pH 9.0. The pH range between pH 6.0 and pH 9.0 is at least 1.5 pH units outside the pK_a values of any of the bases present (Sänger, 1984), and the pH is therefore not expected to effectively influence the T_m (Record, 1967). In the pH range below pH 6.0 the hairpin to coil transition for oligonucleotides containing cytosine merges with the triple helix to hairpin transition and can not be separated from the unfolding of the third strand. Only the oligonucleotide ITS-ATT shows a separate hairpin to coil transition. This transition is slightly dependent on the pH in the pH range between pH 4.0 and pH 6.0. It is reasonable to assume that this behavior originates from the closure of the hairpin loop by an unusual C^+•C base pair at the low pH conditions (Völker et al., 1993).

B. The Intramolecular Triple Helix to Hairpin Transition: The T_m Changes Linearly with %C^+ (Global Composition). Listing the thermodynamic parameters for the intramolecular triple helix to hairpin transitions (3 ↔ 2) of a series of oligonucleotides with an increasing number of third strand cytosines (all spaced by intervening thymines) allows the determination of the impact of global composition (%C^+) on the binding of the third strand. A plot of T_m versus (%C^+) at pH 6.75 (100 mM Na^+) results in a straight line with a slope almost identical to that of the hairpin to coil transition (Figure 4 insert), provided the T_m of ITS-ATT above pH 7.0 is used. ITS-ATT is the only oligonucleotide of this series containing a C_4 loop connecting the third strand to the hairpin component. At pH values below pH 7.0, this C_4 loop, like the C_4 loop of the hairpin at lower pH values, becomes protonated and forms the unusual C^+•C basepair.

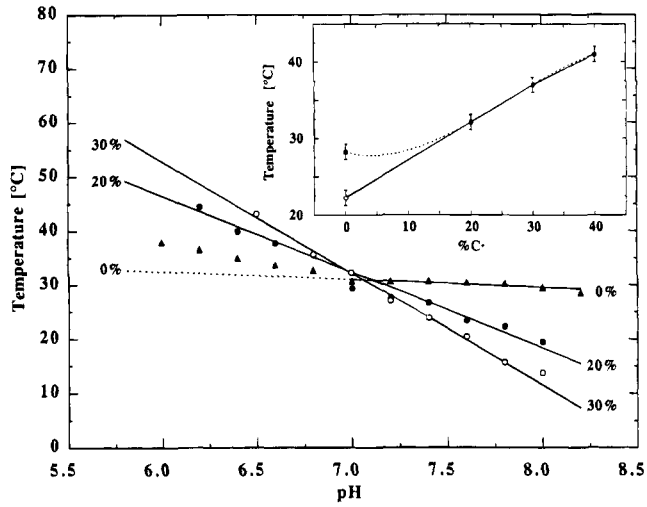

FIGURE 4: Phase diagram of T_m versus pH for the triple helix to hairpin transition for the oligonucleotides in Table 1A: ITS-ATT, closed triangle; ITS-$2G_3$ closed circle; ITS-$3G_2$, open circle. The increase in T_m for the 3 ↔ 2 transition of ITS-ATT at pH values below pH 7.0 due to the protonation of the C4 triple helix loop is clearly shown. The insert shows that a plot of T_m versus global composition (%C^+) (100 mM Na^+, pH 6.75) is linear if the T_m for 0% cytosine at pH 8.0 (open circle) is used instead of pH 6.75 (closed circle).

Unlike for the hairpin to coil transition, the increase in T_m with %C^+ observed for the triple helix to hairpin transition cannot be ascribed to the difference in the number of hydrogen bonds between C^+•G and T•A Hoogsteen base pairs. Both are linked by two hydrogen bonds (Cheng & Pettitt, 1992; Arnott et al., 1976). Small differences in stacking interactions alone are insufficient to explain the observed increase in T_m. Only electrostatic attractions between the protonated cytosines and the negatively charged phosphates in the three backbones can be responsible for the increase in T_m.

The Binding Energy of a Third Strand Depends on the pH. In order to bind in the Hoogsteen position, the third strand cytosines need to be protonated several pH units above the pK_a of free cytosine. The overall negative free energy change for the formation of a C^+•G•C base triplet arises from hydrogen bond formation, stacking interactions, and electrostatic interactions. It must be sufficiently large ($\Delta G < 0$) in order to compensate for the positive free energy of protonation ($\Delta G > 0$). From the observed direct correlation between the T_m and %C^+, one can conclude that the free energy change for the formation of a T•A•T triplet is on the average smaller than the free energy for the formation of a C^+•G•C triplet. Since the magnitude of the free energy assigned to the protonation is dependent on the pH, it is not surprising that the gross free energy change of triplex formation will vary with the pH. A plot of T_m versus pH for the oligonucleotides investigated in this study results in a series of straight lines (Figure 4). The sequence with the highest percentage of protonated cytosines shows the steepest slope $\{d(T_m)/d(pH)\}$. The $d(T_m)/d(pH)$ value is proportional to the number of protons bound to the triple helix (Record et al., 1978). Thus a plot of $\{d(T_m)/d(pH)\}$ versus global composition is expected to result in a straight line relationship. This is indeed observed. The lines in Figure 4 all

intersect at one point (pH 7.1/150 mM Na^+). At this pH the binding energy of C^+ to G·C is identical to the binding energy of T to A·T. At pH values lower than pH 7.1 the C^+ binding energy is higher than that of T, and at pH values larger than pH 7.1 the opposite results.

The Binding Energy of the Third Strand Depends on the Ionic Strength. It is noteworthy that the T_m's of the triplexes are all identical at about 400 mM Na^+ at pH 6.75. At this condition thymines bind as strongly to A·T base pairs as protonated cytosines bind to G·C base pairs. Above 400 mM Na^+ (pH 6.75), the thymine-rich third strands bind stronger than the cytosine-rich strands, and below 400 mM Na^+ the trend is reversed. The extension of the polyelectrolyte theory to protonated polyions predicts (Record et al., 1976; Record, 1967) that a linear increase in the number of protonated residues results in a linear decrease in the slope $d(T_m)/d(\log[Na^+])$. For the triple helix to hairpin transition the observed change in $d(T_m)/d(\log[Na^+])$ with third strand (%C^+) is only approximated by the theoretical predictions. This could be due to either sequence-dependent differences in ion condensation onto the hairpin stem or result from the "cross talk" of site-bound protons.

The Local Composition Modifies the Effect of Global Composition. It is generally assumed that the binding enthalpy of bases in the Hoogsteen position is only dependent on the percentage of protonated cytosines and not on their sequential position. The experiment tells otherwise. A comparison of the T_m's of the oligonucleotide ITS-$3G_0$, ITS-$3G_1$, and ITS-$3G_2$, which have the same global composition ($3C^+$) but differ in the arrangement of the central two cytosines (clustered vs spaced), reveals that the local composition indeed exerts a strong influence on the binding enthalpy of the third strand. The most pronounced electrostatic repulsion between adjacent protonated cytosines occurs for ITS-$3G_0$. The repulsion of the adjacent cytosines is strong and sufficient to overcome the effect of the electrostatic attraction between the protonated cytosines and the negatively charged backbones. The T_m is lower than that of a triple helix with fewer cytosines (ITS-$2G_3$) which are not clustered. Insertion of thymines between adjacent cytosines hence must be favorable. As the experiments show, the separation of cytosines by two thymines (e.g., ITS-$3G_2$) appears sufficient to render these local effects insignificant (Table 2B).

The extent of the electrostatic repulsions depends on the pH and ionic strength. At pH values closer to the pK_a of free cytosine (pH 4.5), electrostatic repulsions between adjacent cytosines become less pronounced. (The T_m for the triplex to hairpin transition of ITS-$3G_0$ is now larger than that of ITS-$2G_3$.) Conversely, an increase in ionic strength increases the electrostatic repulsion between adjacent cytosines. (The T_m of ITS-$3G_0$ becomes progressively smaller than that of ITS-$2G_3$.) At any combination of pH and ionic strength the T_m of ITS-$3G_0$ always remains lower than that of ITS-$3G_1$ or ITS-$3G_2$.

C. The Intramolecular Triple Helix to Coil Transition: The Triple Helix to Coil Transition Can Be Described by the Polyelectrolyte Theory. From the extension of the polyelectrolyte theory to protonated polymers, it can be predicted that the degree of ion condensation is proportional to the difference in overall charge density between the triple helix and the coil state, which in turn is linearly dependent on the number of cytosines in the third strand. The thermal denaturation of the triple helix at the pK_a of free cytosine (pH 4.5) does not result in a net loss or gain of protons. In contrast to the results obtained for the triple helix to hairpin transition the $d(T_m)/d(\log[Na^+])$ value of the triple helix to coil transition at pH 4.5 is in agreement with the predicted straight-line relationship between $d(T_m)/d(\log[Na^+])$ and global composition (%C^+GC).

The Stability of the Intramolecular Triple Helix Reaches a Maximum at pH 4.5. The triple helix to hairpin transition of all cytosine containing oligonucleotides merges with the hairpin to coil transition below pH 6.0 resulting in a single transition. The T_m of the triple helix to coil transition increases with decreasing pH up to pH 4.5 and decrease again at pH values below pH 4.5 (Figure 2). This pH forms an inflection point in the plot of T_m versus pH, indicating that the pK_a of the cytosines in the triple helix is the same as that of free cytosine (Record et al., 1978, Sänger, 1984).

The T_m Changes Linearly with Global Composition. A plot of the T_m values of the triple helix to coil transition (pH 4.5/100 mM Na^+) versus the global composition (%C^+GC) results in a straight-line relationship. The slope of this plot is nearly twice that of the hairpin to coil transitions at pH 8.0. The T_m of the intramolecular triple helix to coil transition cannot be estimated from the T_m's of the individual component transitions. A formal description of this phenomenon is currently impossible because of the inherent difficulty in extrapolating reliable T_m values from the experimentally accessible pH range of the component transitions to pH 4.5.

CONCLUDING REMARKS

The electrostatic interactions between protonated cytosines and the backbones on the one hand and between adjacent protonated cytosines on the other hand have to be considered in addition to stacking and hydrogen bonding interactions to account for the observed local and global composition-dependent changes in triple helix stability.

REFERENCES

Arnott, S., & Bond, P. J. (1973a) *Nature New Biol. 244*, 99–101.

Arnott, S., & Bond, P. J. (1973b) *Science 181*, 68–69.

Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521.

Arnott, S., Bond, P. J., Selsing, E., & Smith, P. J. C. (1976) *Nucleic Acids Res. 3*, 2459–2470.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K., Eds. (1987) In *Current Protocols in Molecular Biology Unit 3.10.2*, J. Wiley & Sons, New York.

Blake, R. D., & Haydock, P. V. (1979) *Biopolymers 18*, 3089–3109.

Breslauer, K. J., Frank, R., Blöcker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.

Broitman, S. L., Im, D. D., & Fresco, J. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5120–5124.

Callahan, D. E., Trapane, T. L., Miller, P. S., Ts'o, P. O. P., & Kan, L.-S. (1991) *Biochemistry 30*, 1650–1655.

Chan, S. S., Breslauer, K. J., Hogan, M. E., Kessler, D. J., Austin, R. H., Ojemann, J., Passner, J. M., & Wiles, N. C. (1990) *Biochemistry 29*, 6161–6171.

Cheng, Y.-K., & Pettitt, B. M. (1992) *Prog. Biophys. Mol. Biol. 58*, 225–257.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., & Benight, A. S. (1992a) *Biopolymers 32*, 849−864.
Egholm, M., Nielsen, P. E., Buchard, O., & Berg, R. H. (1992) *J. Am. Chem. Soc. 114*, 9677−9678.
Felsenfeld, G., & Miles, H. T. (1967) *Annu. Rev. Biochem. 36*, 407−448.
Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023−2024.
Froehler, B. C., & Ricca, D. J. (1992) *J. Am. Chem. Soc. 114*, 8320−8322.
Griffin, L. C., & Dervan, P. B. (1989) *Science 245*, 967−971.
Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P., & Dervan P. B. (1992a) *J. Am. Chem. Soc. 114*, 7976−7982.
Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P., & Dervan, P. B. (1992b) *Biochemistry 31*, 2829−2834.
Hattori, M., Frazier, J., & Miles, T. D. (1976) *Biopolymers 15*, 523−531.
Helene, C., & Toulme, J.-J. (1990) *Biochim. Biophys. Acta 1049*, 99−125.
Hoogsteen, K. (1959) *Acta Crystallogr. 12*, 822−823.
Horne, D. A., & Dervan, P. B. (1991) *Nucleic Acids Res. 19*, 4963−4965.
Howard, F. B., & Miles, H. T. (1977) *Biochemistry 16*, 4647−4650.
Howard, F. B., Miles, H. T., Liu, K., Frazier, J., Raghunathan, G., & Sasisekharan, V. (1992) *Biochemistry 31*, 10671−10677.
Kiessling, L. L., Griffin, L. C., & Dervan, P. B. (1992) *Biochemistry 31*, 2829−2834.
Klump, H. H. (1986) *FEBS Lett. 196*, 175−179.
Klump, H. H. (1988) Conformational transitions in Nucleic Acids, in *Biochemical Thermodynamics* (Jones, M. N., Ed.) 2nd ed., Elsevier, New York.
Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603−6614.
Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., & Fresco, J. R. (1988) *Biochemistry 27*, 9108−9112.
Lipanov, A. A., Chuprina, V. P., Alexeev, G., & Skuratovskii, I. Ya. (1990) *J. Biomol. Struct. Dyn. 7*, 811−826.
Lipsett, M. N. (1964) *J. Biol. Chem.* 239, 1256−1260.
Macaya, R., Schultz, P., & Feigon, J. (1992a) *J. Am. Chem. Soc. 114*, 781−783.
Macaya, R., Wang, E., Schultz, P., Sklenar, V., & Feigon, J. (1992b) *J. Mol. Biol. 225*, 755−773.
Macaya, R. F., Gilbert, D. E., Malek, S., Sinsheimer, J. S., & Feigon, J. (1991) *Science 254*, 270−274.
Maher, L. J. III. (1992) *BioEssays 14*, 807−815.
Maher, L. J., III. Dervan, P. B., & Wold, B. J. (1990) *Biochemistry 29*, 8820−8826.
Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179−246.
Manzini, G., Xodo, L. E., Gasparotto, D., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1990) *J. Mol. Biol. 213*, 833−843.
Marck, C., & Thiele, D. (1978) *Nucleic Acids Res. 5*, 1017−1028.
Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
Marmur, J., & Doty, P. (1962) *J. Mol. Biol.* 5, 109−118.
Mergny, J. L., Sun, J.-S., Rougée, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J., & Hélène, C. (1991) *Biochemistry 30*, 9791−9798.
Michelson, A. M., Massoulié, J., & Guschlbauer, W. (1967) *Prog. Nucleic Acid Res. Mol. Biol. 6*, 83−141.
Milligan, J. F., Krawczyk, S. H., Wadwani, S., & Matteucci, M. D. (1993) *Nucleic Acids Res. 21*, 327−333.
Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol. 37*, 63−80.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645−650.
Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221−226.
Ono, A., Ts'o, P. O. P., & Kan, L.-S. (1991) *J. Am. Chem. Soc. 113*, 4032−4033.
Park, Y.-W., & Breslauer, K. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6653−6657.
Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990) *Nucleic Acids Res. 18*, 5743−5750.
Pilch, D. S., Levenson, C., & Shafer, R. H. (1991) *Biochemistry 30*, 6081−6087.
Plum, E. G., Park, Y.-W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436−9440.
Povsic, T. J., & Dervan, P. B. (1990) *J. Am. Chem. Soc. 112*, 9428−9430.
Radhakrishnan, I., & Patel, D. J. (1992) *J. Am. Chem. Soc. 114*, 6913−6915.
Radhakrishnan, I., Gao, X., & Patel, D. J. (1991) *Biochemistry 31*, 2514−2523.
Raghunathan, G., Miles, H. T., & Sasisekharan, V. (1993) *Biochemistry 32*, 455−462.
Record, M. T., Jr. (1967) *Biopolymers 5*, 993−1008.
Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103−178.
Record, M. T., Jr., & Lohman, T. M. (1978) *Biopolymers 17*, 159−166.
Record, M. T., Jr., Woodbury, C. P., & Lohman, T. M. (1976) *Biopolymers 15*, 893−915.
Rich, A. (1958) *Nature 181*, 521−525.
Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359−389.
Riordan, M. L., & Martin, J. C. (1991) *Nature 350*, 442−443.
Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T., & Helene, C. (1992) *Biochemistry 31*, 9269−9278.
Sänger, W. (1984) in *Principle of Nucleic Acid Structure*, Springer-Verlag, New York.
Singleton, S. F., & Dervan, P. B. (1992a) *J. Am. Chem. Soc. 114*, 6957−6965.
Singleton, S. F., & Dervan, P. B. (1992b) *Biochemistry 31*, 10995−11003.
Sklenar, V., & Feigon, J. (1990) *Nature 345*, 836−838.
Steely, H. T., Gray, D. M., & Ratliff, R. L. (1986) *Nucleic Acids Res. 14*, 10071−10090.
Stilz, H. U., & Dervan, P. B. (1993) *Biochemistry 32*, 2177−2185.
Strobel, S. A., & Dervan, P. B. (1990) *Science 249*, 73−75.
Strobel, S. A., & Dervan, P. B. (1991) *Nature 350*, 172−174.
Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Houseman, D. E., & Dervan, P. B. (1991) *Science 254*, 1639−1642.
Thiele, D., & Guschlbauer, W., (1968) *FEBS Lett. 1*, 173−175.
Thiele, D., Marck, C., Schneider, C., & Guschlbauer, W. (1978) *Nucleic Acids Res. 5*, 1997−2012.
Völker, J., Botes, D. P., Lindsey, G. G., & Klump, H. H. (1993) *J. Mol. Biol. 230*, 1278−1290.
Wang, E., Malek, S., & Feigon, J. (1992) *Biochemistry 31*, 4838−4846.
Xodo, L. E., Manzini, G., & Quadrifoglio, F. (1990) *Nucleic Acids Res. 18*, 3557−3564.
Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.