Triple Helical Structures Involving Inosine: There Is a Penalty for Promiscuity†

Martin Mills,‡ Jens Völker,§ and Horst H. Klump*,‡

Department of Biochemistry, University of Cape Town, Private Bag, Rondebosch 7700, Republic of South Africa, and Department of Chemistry, Rutgers University, New Brunswick, New Jersey 08855-0939

Received January 25, 1996; Revised Manuscript Received May 6, 1996⊗

ABSTRACT: Inosine has the ability to act as a "wild-card" binding nonspecifically to both A·T and G·C base pairs. This has obvious implications for the design of oligonucleotide site-directed probes. In this paper we present a series of oligonucleotides with a 5′pur$_9$-pyr$_9$-pyr$_9$ motif which are designed to fold up sequentially into intramolecular triple helices. One or more inosines are incorporated into the Hoogsteen strands in place of T's and/or C's. Once folded into the triplex, the inosine-containing third strand is incorporated in parallel orientation to the purine strand of the duplex. The influence of inosine on the triplex−duplex equilibrium, characterized by the melting temperature (T_m), and on the phase boundaries, as a function of pH and/or ionic strength, has been assessed by means of UV and CD spectroscopy. There are two distinguishable influences of third-strand inosines which affect binding, namely, backbone distortion due to bulkiness (I for T and I for C^+) and/or loss of intramolecular ion pairs between protonated cytosines and the backbone phosphates (I for C^+). A single thymine replacement drops the T_m by 25.0 (±2.1) °C, and replacing a single protonated cytosine drops the T_m by 32.1 (±1.0) °C at pH 6.0. On introducing two inosines in place of thymines, the T_m at pH 6.0 of the triple helix to hairpin transition is lowered by 35.5 (±1.4) °C; on introducing two inosines in place of cytosines, the T_m drops by 44.5 (±1.0) °C, and on replacing a cytosine and a neighboring thymine with inosines, the T_m of the same transition is lowered by 29.2 (±1.6) °C. Replacing more than two thymines or cytosines, respectively, eliminates the binding of the Hoogsteen strand at room temperature altogether. Under no circumstances does inosine replacement stabilize the triple helix: it is a poor substitute and its role as a wild-card is limited.

It was not long after the discovery of the DNA double helix (Watson & Crick, 1953) that researchers observed the formation of a polynucleotide triple helix formed from one strand of poly(A)[1] and two strands of poly(U) (Felsenfeld *et al.* 1957). Several three stranded complexes were subsequently discovered which, on detailed investigation, revealed that a homopolypyrimidine third strand binds via Hoogsteen hydrogen bonding (Hoogsteen, 1959) to the major groove sites of a homopurine−homopyrimidine Watson−Crick double helix (Radhakrishnan & Patel, 1994 and references therein; Plum *et al.*, 1995). Common to all these triplexes is a ribo- or deoxyribopolypyrimidine third strand.

X-ray diffraction studies of fibers have revealed an RNA triple helix formed between one strand of poly(A) and two strands of poly(I) (Arnott & Bond, 1973). In this complex, comparable to the U·A·U triplex, one strand of poly(I) binds to the poly(A) strand in the conventional antiparallel orientation in a kind of Watson−Crick (WC) base pairing with the second strand of poly(I) binding in analogy to Hoogsteen base pairing. The I·A·I triple helix diameter of 23.7 Å is very close to 23.2 Å of normal A′-RNA (Arnott & Selsing, 1974). The following polynucleotide triplexes involving inosine strands also form under physiological conditions:

Table 1: Third-Strand Binding Code

	third strand residue				
Watson−Crick core	A	T/U	I	G	C
A·T/U	+	+	+	−	−
G·C	−	−	+	+	+

poly(I·A·U) and poly(I·G·C) (Letai *et al.*, 1988) contain poly(I) as the Hoogsteen (HG) strand, poly(C^+·I·C) (Thielle & Guschlbauer, 1968) contains it as one of the WC strands, and in poly(I·I·C) (Fresco & Massoulie, 1963) the inosine serves as the HG and WC strand, respectively.

There is considerable interest in studying triple helix forming oligonucleotides. Their sequence specificity can be used to direct desirable chemical reactions to target sites on double stranded DNA, thus mimicking enzymes in specificity and range of action (Moser & Dervan, 1987). Current applications lie in chromosome analysis and gene mapping (Hélène *et al.*, 1989; Strobel & Dervan, 1990), and in the oligonucleotide based control of gene expression (Francois *et al.*, 1989; Maher *et al.*, 1989). The ultimate goal of "triplex technology" is to target any DNA sequence and thus provide a viable means of therapy at the gene level. A significant contribution was made in 1988 by Letai *et al.* with their proposal of the "third-strand binding code" derived from affinity chromatography (Table 1). The "code" shows some analogy to the Watson−Crick base-pairing scheme in that adenine or thymine residues in single strand sequences exclusively bind to A·T base pairs in Watson−Crick helices and guanine or protonated cytosines recognize only G·C base pairs.

† We gratefully acknowledge the support of the South African Foundation for Research Development (FRD).

* Corresponding author.
‡ University of Cape Town.
§ Rutgers University.

⊗ Abstract published in *Advance ACS Abstracts,* July 15, 1996.

[1] Abbreviations I, inosine; A, adenine; T, thymine; C, cytosine; G, guanine; UV, ultraviolet; UV-melting, temperature-dependent UV absorbance spectroscopy; CD, circular dichroic spectropolarimetry.

S0006-2960(96)00193-6 CCC: $12.00 © 1996 American Chemical Society

The guanine analogue inosine is shown in this experiment to bind to both A·T and G·C base pairs. This might be exploited in the design of oligonucleotide probes in which inosine can act as a "wild-card". Furthermore, inosine does not require protonation for binding to a duplex as cytosine does, so replacing protonated cytosine in a third strand sequence would remove the pH dependence of triplex formation.

A different table was published by Griffin and Dervan (1989) in which 20 base pairs were examined for binding specificity compatible with the pyrimidine−Hoogsteen triple helix motif. Good binding in these experiments corresponds to effective iron−EDTA cleavage. According to their results, inosine is a very ineffective third strand binder for any of the four Watson−Crick duplex combinations. The result was recently confirmed by Shimizu *et al.* (1994).

In this paper we will re-examine the impact of inosines on the third strand binding by using a self-folding oligonucleotide model system (Häner & Dervan, 1990; Sklenár & Feigon, 1990; Völker *et al.* 1993; Plum & Breslauer, 1995).

A self-folding triple helix formed from a single DNA strand was first reported by Sklenár and Feigon in 1990, and these structures have been investigated with one- and two-dimensional NMR techniques (Macaya *et al.*, 1991; Radhakrishnan & Patel, 1994). By choosing pH and counterion concentration accordingly, three different conformational states of the sequences can be separated; besides a random coil the hairpin formation can be distinguished from the triplex formation as seen in CD spectra and UV-melting experiments (Völker *et al.*, 1993; Hüsler & Klump, 1994; Plum & Breslauer, 1995). By using the intramolecular folding of a single strand, one can (a) be almost sure of excluding alternative secondary structure formation (Häner & Dervan, 1990), (b) maintain the prospective strands in high local concentration, (c) obtain the correct stoichiometry, and (d) enforce the orientation of the third strand with respect to the purine strand in the preformed target duplex.

The sequences investigated here are based on the sequence JV-ITS (Völker *et al.*, 1993) which forms a stable intramolecular triplex. All the sequences used here form an identical Watson−Crick hairpin helix as the core with an extension which can fold on to form a triple helix in which canonical bases are systematically replaced by inosine. The technique of UV-melting was used to determine the melting temperature (T_m) of the triplex−hairpin, hairpin−coil, and triplex−coil transitions. Phase diagrams (T_m *vs* pH and T_m *vs* log[Na^+]) were constructed according to Klump (1987). The impact of inosine replacement on the third strand binding is reflected in the T_m in the first place and consequently in the shifting of the phase boundaries. CD spectroscopy was used to discriminate between the three possible conformational states, coil, double helix, and triple helix, and to monitor the thermally induced interconversion between the three structural motifs of the molecule (Plum & Breslauer, 1995).

MATERIALS AND METHODS

(A) DNA Preparation. The seven oligonucleotides listed in Table 2 were synthesized on an Autogen 6500 (Miligen, Burlington, MA) by conventional phosphoramidate chemistry and purified by HPLC on a Pharmacia Mono Q anion exchange column (Völker *et al.*, 1993). Purity was checked by gel electrophoresis. Only samples with less than 3% failure sequences have been used throughout this study.

Table 2: Parent Oligonucleotide MCT and Derivative Sequences Incorporating Inosine

Name	Sequence
MCT	GAGAGAGAG CCTT CTCTCTCTC CCTT CTCTCTCTC
MCIC	GAGAGAGAG CCTT CTCTCTCTC CCTT CTCTCICTC
MICI	GAGAGAGAG CCTT CTCTCTCTC CCTT CTCICICTC
MIT	GAGAGAGAG CCTT CTCTCTCTC CCTT CTCTITCTC
MITI	GAGAGAGAG CCTT CTCTCTCTC CCTT CTITITCTC
MII	GAGAGAGAG CCTT CTCTCTCTC CCTT CTCTIICTC
MHPIN	GAGAGAGAG CCTT CTCTCTCTC CCTT

(B) UV-Melting. Thermal denaturation monitored by UV absorbance was carried out in a Pye-Unicam SP1800 spectrophotometer with a custom-made heating block interfaced to an IBM PC through an Oasis digital converter. Typical samples contained about 1.5 μM oligonucleotide, giving an absorbance of 0.7 at 260 nm. Absorbance values at 260 nm were recorded every 0.3 °C at a heating rate of 1 °C/min. Samples were heated to 85 °C prior to the first scan and allowed to cool over 40 min to avoid alternative folding intermediates. Scans of each sequence were reproducible, so the reactions are completely reversible. The melting temperatures (T_m) were extracted from the digitized data by a computer based curve fitting procedure (Bevington, 1969; Hüsler & Klump, 1994) and processed with the Quattro-Pro software package. Solutions were composed by mixing two stock buffer solutions of 50 mM [Na^+] (20 mM sodium acetate, 20 mM sodium cacodylate, 10 mM sodium chloride) and 5 M [Na^+] (20 mM sodium acetate, 20 mM sodium cacodylate, 5 M sodium chloride). The pH was adjusted with minute amounts of concentrated hydrochloric acid to avoid undue dilution. All chemicals were analytical grade dissolved in double distilled water.

(C) CD Spectroscopy. CD spectra (360−205 nm in 0.5 nm steps) were obtained on an Aviv Model 60DS spectropolarimeter (Aviv Associates, Lakewood, NJ) in an aqueous buffer solution (100 mM [Na^+]) in a 0.1 cm path length cuvette (Hellma 110QS). An extinction coefficient of 3.27 $\times$ 10^5 M^{-1} cm^{-1} was calculated for MCT according to Cantor *et al.* (1970). Phosphate analysis (Snell & Snell, 1949) of MCIC and MICI yielded extinction coefficients of 2.68 $\times$ 10^5 ($\pm$0.8 $\times$ 10^3) M^{-1} cm^{-1} and 2.6 $\times$ 10^5 ($\pm$1.6 $\times$ 10^3) M^{-1} cm^{-1}, respectively, quoted at 90 °C on a per strand basis. The spectra were normalized in units of molar ellipticity for comparison.

RESULTS AND DISCUSSION

(A) The Unmodified Parent Sequence MCT Can Sequentially Fold into a Triple Helix by Way of a Double Helical Hairpin Intermediate. Figure 1 shows the proposed unfolding pathway of the intramolecular triple helix MCT induced by rising the temperature under given conditions of pH and sodium concentration. As will be shown below, the forma-

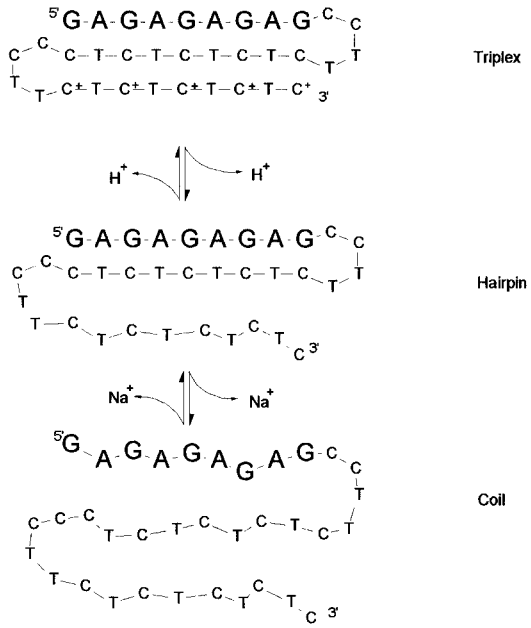

FIGURE 1: The proposed thermally induced unfolding pathway of the oligonucleotide MCT. The equilibrium between the conformations is influenced by solution conditions, as indicated by changes in proton and sodium ion binding.

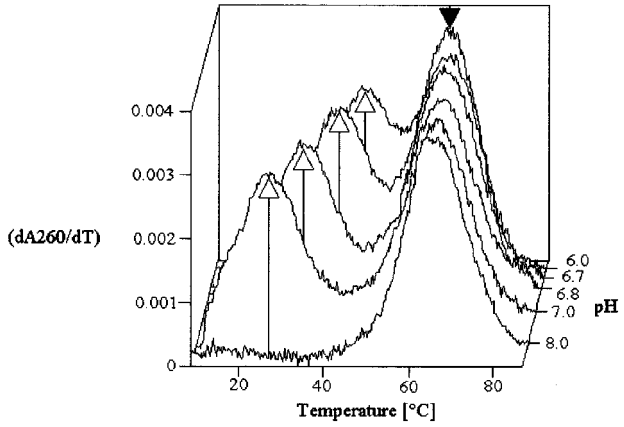

FIGURE 2: A plot of the first derivatives of melting profiles ((dA_{260}/dT) *vs* T) obtained for a pH series for oligonucleotide sequence MCT (1.5 μM single strand in 100 mM [Na^+] buffer: 20 mM sodium acetate, 20 mM sodium cacodylate, and 60 mM NaCl). Each scan was repeated at least twice.

tion of the triplex requires the formal charge on the cytosine residues of the Hoogsteen sequence. At pH 7.5 the melting temperature of the triplex has dropped below 20 °C.

The effect of the solution pH on the stability of the conformers is further exemplified in Figure 2 which shows a compilation of derivatives of melting profiles ((dA_{260}/dT) *vs* T) at five different pH values. At pH 8.0 a monophasic transition appears with a T_m of about 65 °C. As was confirmed by CD spectroscopy and chemical footprinting methods (Völker *et al.*, 1993), no Hoogsteen strand sequence is bound at this pH and the initial conformation is that of the extended hairpin conformation. Between pH 7 and 6.7, a new low-temperature transition, which is assigned to the triplex−hairpin transition, appears as a clearly separate peak set aside from the pH invariant high-temperature transition which is assigned to the hairpin−helix to coil transition. Dropping the pH further lifts the T_m of the triplex−hairpin transition. At pH 6.0 it starts to merge with the hairpin−coil transition. This is in good agreement with the previously

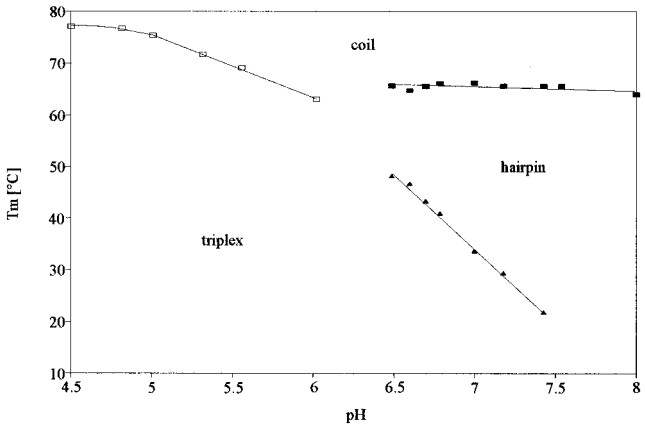

FIGURE 3: Phase diagram of MCT (T_m *vs* pH) in 100 mM [Na^+] buffer (20 mM sodium acetate, 20 mM sodium cacodylate, and 60 mM NaCl) showing the phase boundaries of the triplex−coil (empty squares), triplex−hairpin (filled triangles), and hairpin−coil (filled squares) transitions.

published observations on the pH dependency of the thermal unfolding of the related oligonucleotide sequence JV-ITS (Völker *et al.*, 1993). It is obvious from this pH dependency that the first step is accompanied by a release of the positive charges from the Hoogsteen cytosines, resulting in an increase of T_m with decreasing pH. The second step (at T_m = 65 °C) is pH independent, no further change in the degree of protonation is occurring, and the T_m is invariant under the set of solvent conditions applied.

Figure 3 gives the complete phase diagram of the reference sequence MCT, *i.e.*, the plot of T_m *vs* pH for the triplex−hairpin transition (filled triangles), the hairpin−coil transition (filled squares), and the triplex−coil transition (empty squares), in a pH range where there is no stable hairpin intermediate in the unfolding pathway. In the pH range between 6.5 and 6 the triplex−hairpin and the hairpin−coil transitions are becoming increasingly superimposed such that no effort is made to deconvolute the two transitions to extract T_m's and no individual data are shown in the diagram. The sodium concentration is kept constant throughout this series at 100 mM [Na^+]. The full lines represent the phase boundaries between the three conformers where they can be confidently resolved. A complete set of phase diagrams has been obtained for the other six sequences. Only the slopes of the boundaries are listed in Table 3. A detailed discussion of variation in slope with variation in sequence for the triplex−hairpin transitions is given below (Figure 8). The hairpin−coil transition is taken as a kind of internal standard transition because it is pH independent and does not reflect sequence variation in the single strand dangling ends of the hairpins. The average slope of the dT_m/d(pH) plots, representing the phase boundary between the hairpin and coil, is zero (Figure 4). It should be pointed out that the phase boundary between triplex and coil decreases in slope to zero when the pH approaches 5, which is close to the pK_a of free cytosine (pK_a = 4.5).

(B) CD Spectroscopy Is Used To Assign the Conformational States Associated with Each Transition. The thermodynamically most stable conformational state and its conversion into another conformation can be best characterized by CD spectroscopy as temperature-dependent CD spectra, recorded under preselected experimental conditions (temperature, pH, and [Na^+]), give clear and independent evidence for the presence of the triplex, hairpin, or random

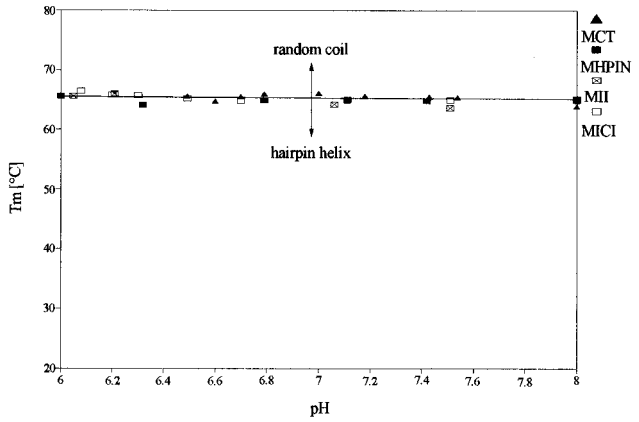

FIGURE 4: T_m *vs* pH for the hairpin−coil transition for MCT, MHPIN, MII, and MICI (100 mM [Na^+]).

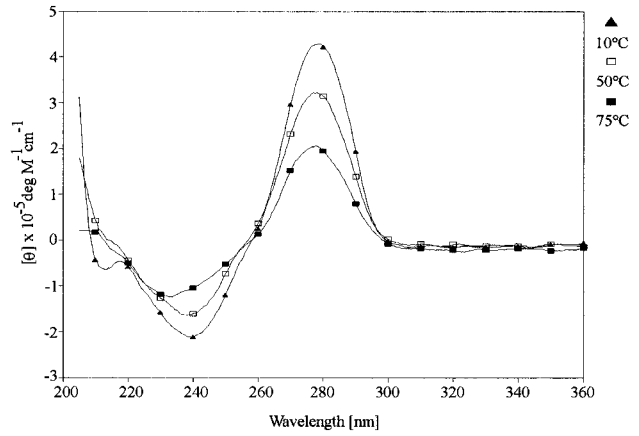

FIGURE 5: Temperature dependence of CD spectra of MCT ($1.78 \times 10^{-5} \pm 0.1 \times 10^{-5}$ M) in aqueous buffer solution (100 mM [Na^+], pH 6.7) at 10 °C (triangles), 50 °C (empty squares), and 75 °C (filled squares).

coil as the dominant conformation (Figures 5 and 9). A negative band at 246 nm and a second one appearing at 213 nm, present only at low temperature (10 °C spectrum, Figure 5) and pH conditions (>pH 6), make up the established indicators for the presence of the triplex as the prevailing conformation (Mancini *et al.*, 1990; Pilch *et al.*, 1990; Xodo *et al.*, 1990; Johnson *et al.*, 1991; Scaria & Schafer, 1991; Roberts & Crothers, 1992; Völker *et al.*, 1993; Plum & Breslauer, 1995).

Conditions which favor the hairpin as the dominant stable conformation such as neutral pH and intermediate temperature (25−50 °C) are changing the features of the CD spectra. A large positive band at 276 nm, a small negative band at 242 nm, and a small positive band at 221 nm are indicative for the B-DNA type conformation which is expected for the hairpin helix (Bush, 1974) (*cf.* spectrum 2 in Figure 5). Raising the temperature well above the T_m for the hairpin−coil transition, which occurs at about 65 °C for the sequences investigated here in 100 mM [Na^+] solution, the positive band at 276nm decreases substantially as does the negative band at 246 nm (*cf.* spectrum 3 in Figure 5). The high-temperature spectrum (75 °C) is in good agreement with the spectrum of a thermally disordered single strand. Since the characteristics of the three conformational states, triplex, duplex, and coil, are well reflected in the CD spectra, it is sufficient to use the corresponding bands to identify the three structural motifs. They can be used to assign the three motifs to their respective areas in the phase diagram and to monitor

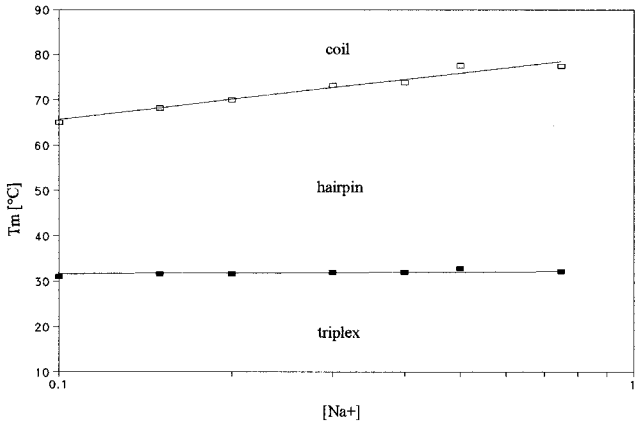

FIGURE 6: T_m *vs* log[Na^+] (pH 6.7) for triplex and hairpin conformations of MCT.

the phase transitions as a function of pH or sodium concentration and/or temperature.

(C) The Hairpin−Coil Transition Can Serve as an Internal Control for the pH and Salt Dependent Behavior of Each Phase Transition. It is clear from Figure 6 that the triplex−hairpin transition temperature (full rectangles) is independent of the sodium concentration in the interval 100−750 mM [Na^+], *i.e.*, no sodium ions are taken up or released during this conformational change. In contrast to the behavior of the triplex, the hairpin conformation gains thermal stability with increasing [Na^+] and sodium ions have to be released when the hairpin unfolds to the coiled single strand. Similar findings are reported in the literature (Record *et al.*, 1978; Breslauer *et al.*, 1986).

At pH >7.5 the stability of the triplex has dropped below room temperature and only the hairpin exists as the thermodynamically most stable conformation. Hence the UV-melting profile for the hairpin−coil transition (not shown) can be used to calculate the T_m and the corresponding transition enthalpy (ΔH_{vH}) following a procedure first published by Marky and Breslauer (1987); we can safely assume that the hairpin−coil transition follows a two-state mechanism with no intermediate states populated. The T_m for this sequence amounts to 65 °C (100 mM [Na^+]), and the ΔH_{vH} yields 57.6 (±5) kcal/mol of single strand. The enthalpy value is in good agreement with a computed value derived from increments of stacked base pairs, $\Delta H_{calc} = 59.9$ kcal/mol of 9mer (55% GC). These increments were derived from stacking interactions of selected polymer double helical complexes measured by calorimetry (Klump, 1988). The close correspondence of the two values supports the assumption that the hairpin−coil transition is well described by a two-state process. Plum and Breslauer (1995) have reported corresponding ΔH_{VH} values for slightly shorter hairpin helices with larger loops, but their ΔH_{cal} values are systematically lower. No explanation can be given for the discrepancy at this stage. The sequences investigated (*cf.* Table 2) are designed in such a way that they will all fold into the same hairpin helix, with the only variations present in the dangling single strand extensions which can be shown not to affect the stability of the hairpin conformation in any detectable way. The results for the hairpin−coil transition are shown in Figures 2 and 4 and listed in Tables 3 and 4. Taken together, we can safely state that the hairpin−coil transition temperature is independent of the pH and the sequence variations in the 3′-extensions. This conserved

Table 3: Summary of the pH Dependency (d(T_m)/d(pH)) Data Recorded at 100 mM [Na^+]

	triplex−hairpin	hairpin−coil
MCT	−28.9 ± 0.9	−0.7 ± 0.4
MCIC	−29.0 ± 2.1	0.3 ± 0.3
MICI	−30.8 ± 1.4	−0.9 ± 0.2
MIT	−21.5 ± 0.6	−1.4 ± 0.2
MITI	−15.7 ± 0.8	0.0 ± 0.1
MII	−19.4 ± 1.6	−0.9 ± 0.3
MHPIN		0.0 ± 0.2
av		−0.5 ± 0.6

Table 4: Summary of the Sodium Counterion Concentration Dependency (d(T_m)/d(log[Na^+])) Data

	pH 4.5	pH 5.8		pH 8.0
	triplex−coil	triplex−hairpin	hairpin−coil	hairpin−coil
MCT	10.2 ± 0.6	0.4 ± 0.6[a]	14.7 ± 0.9[a]	14.1 ± 0.4
MCIC	13.0 ± 1.0	0.8 ± 1.1	15.8 ± 0.8	11.2 ± 1.0
MICI	11.4 ± 1.0	−4.6 ± 0.5	11.7 ± 0.6	12.8 ± 0.6
MIT	13.9 ± 1.1	4.0 ± 1.7	11.3 ± 0.6	11.6 ± 0.6
MITI	17.5 ± 0.1	6.7 ± 1.2[b]	14.9 ± 1.0[b]	13.2 ± 0.9
MII	11.3 ± 1.4	6.1 ± 1.5	12.9 ± 0.6	13.5 ± 0.7
MHPIN				12.1 ± 0.4
av			13.6 ± 1.7	12.6 ± 1.0

[a] pH 6.7. [b] pH 5.0.

structural element can be used to compare the thermal transitions of the various triplexes.

(D) The Triplex−Hairpin Transition of the Unmodified Sequence MCT. The transition on which we expect the sequence variation and pH changes to have a measurable effect is the triplex−hairpin transition, which becomes apparent in the phase diagram (Figure 3) on lowering the pH to 7.4 and lower. The triplex−hairpin transition shows a strong pH dependency (dT_m/d(pH) = −28.9 ± 0.9 °C), giving the Hoogsteen/loop cytosines an apparent pK_a of about 7.3 at 25 °C (Hüsler & Klump, 1995). In other words, the Hoogsteen strand extension contains five cytosines, each carrying a formal charge. This localized positive charge density explains why, in contrast to the hairpin−coil transition, the triplex−hairpin transition shows no [Na^+] dependence, with dT_m/d(log[Na^+]) = 0.4 ± 0.6 °C (*cf.* Figure 6). The phosphate groups in the vicinity of the third-strand cytosines must be electrically neutralized by the localized positive charges.

In the following paragraphs we will discuss the impact of exchanging one or more pyrimidine residues in the Hoogsteen strand extension with the uncharged purine base inosine. Inosine is supposed to hydrogen bond with the Hoogsteen side of A·T and G·C base pairs. Since inosine is a large purine base, its presence within the pyrimidine sequence should exhibit some steric effects, and in addition, the uncharged inosine must have an effect on the pH dependency of the triplex−hairpin transition. The results are presented in Figures 7 and 8.

(E) The Influence of Inosine for Thymine Exchange on the Binding of the Hoogsteen Strand to the Conserved Hairpin Helix. To separate the effect on the triplex stability of the I for T mutation from the effect of the I for C mutation, we have designed two sets of oligonucleotide sequences either with a single thymine or with two thymines, spaced by a cytosine, replaced by inosine(s) and the corresponding sequences with cytosine replacements. The phase diagram

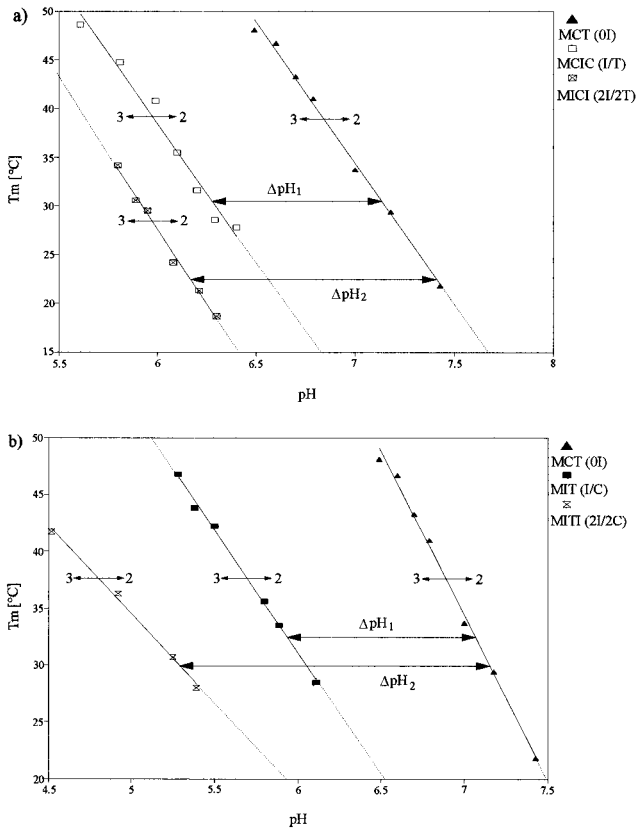

FIGURE 7: Phase diagrams (T_m *vs* pH) comparing the triplex−hairpin (3 ↔ 2) transitions of (a) MCT, MCIC, and MICI and (b) MCT, MIT, and MITI (100 mM [Na^+]).

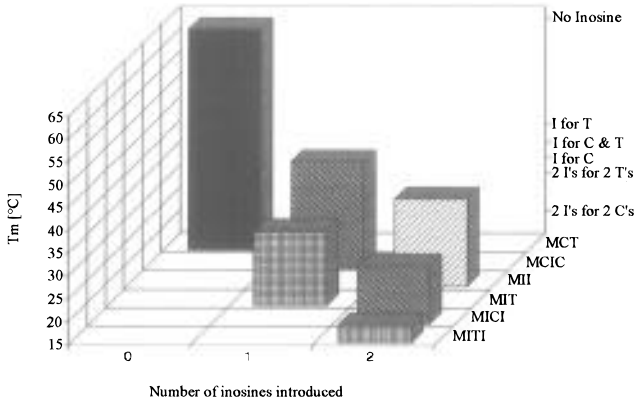

FIGURE 8: Summary of the effect on the triplex−hairpin transition of replacing cytosine and/or thymine with inosine in the third strand as compared to the sequence MCT. T_m *vs* number of inosines introduced (pH 6.0, 100 mM [Na^+]), *vs* the kind of Hoogsteen pyrimidine replaced.

(T_m *vs* pH) which shows the conditions for the triplex−hairpin transition for the two sequences with thymine replaced is given in Figure 7a. The phase boundary for the reference sequence MCT is also included. The three phase boundaries run parallel, i.e., the same number of protons are released during the phase transition. This corresponds to the fact that no formal charge is deleted from the third strand since no cytosine is replaced by inosine.

A single thymine replacement (MCIC) offsets the boundary toward a lower pH by 0.8 pH unit. A further replacement of a second thymine (MICI) offsets the phase boundary by an additional 0.4 pH unit, resulting in a total downward shift of 1.2 pH units. Compared to the reference sequence MCT, a substantial stepwise lowering in pH is required to keep

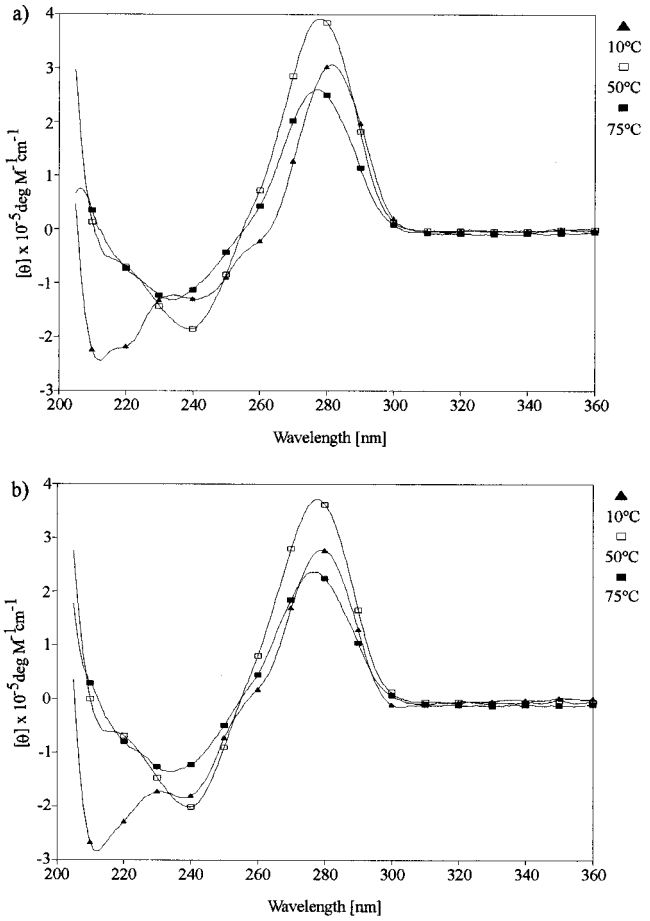

FIGURE 9: Temperature dependence of CD spectra of ($3.43 \times 10^{-5} \pm 0.1 \times 10^{-5}$ M) MCIC (a) and ($2.54 \times 10^{-5} \pm 0.1 \times 10^{-5}$ M) MICI (b) in aqueous buffer solution (100 mM [Na^+], pH 5.8) at 10 °C (triangles), 50 °C (empty squares), and 75 °C (filled squares).

the T_m for the triplex−coil transitions of the three sequences identical. Without compensating for the destabilizing effect of inosine through a drop in pH (*i.e.*, at fixed pH), the T_m drops from 65.9 °C (MCT, pH 6, 100 mM [Na^+]) to 40 °C (MCIC, pH 6, 100 mM [Na^+]; *cf.* also Shimizu *et al.*, 1994) and down to 30 °C (MICI, pH 6, 100 mM [Na^+]) (*cf.* Figures 7a and 8). We attribute this sizeable drop in transition temperature to significant backbone distortion, resulting in decreased stacking interactions. The destabilizing effect does not appear to be simply additive (Table 3 and Figures 7a and 8), but no detailed description of the local distortions can be given. The replacement of one thymine base in the third strand by inosine has no influence on the slope (dT_m/d(log[Na^+])) at pH 5.8 (Table 4), but the slope changes in the order of −5 °C on substitution of a second thymine by inosine, indicating that some destabilization must occur.

CD spectra of MCIC and MICI are presented in Figures 9a,b. The characteristic bands (279 nm (+) and 213 nm (−)) which indicate the triplex formation are clearly shown in the 10 °C spectrum (filled triangle). A slight shoulder between 250 and 260 nm is the only additional change in the CD spectrum as compared to Figure 5 (10 °C, filled triangle) when inosine is introduced into the 3′ pyrimidine sequence.

(F) The Influence of Inosine for Cytosine Exchange on the Binding of the Hoogsteen Strand to the Conserved Hairpin Helix. Figure 7b shows the effect on the triplex stability of replacing cytosine(s) with inosine(s) (MIT and MITI) as compared to the reference sequence MCT. With increasing loss of formal positive charges (protonated cytosines) from the third strand, the slope of the phase boundary decreases indicating that fewer protons are released upon the triplex to hairpin transition. A 7 °C decrease in slope is observed for a single cytosine replacement, which is in very good agreement with the decrease in slope of 6.6 °C predicted for losing a single formal charge in the third strand (Völker, Ph.D. thesis, Cape Town, 1993). To observe a phase transition at 25 °C, the pH of the solution has to be lowered by 1.0 pH unit when one cytosine is mutated to inosine (MIT) and by 1.7 pH units when two cytosines are replaced. If the pH is kept constant, the T_m for the phase transition decreases by 32 °C (pH 6, 100 mM [Na^+]) for a single cytosine replacement, and a further drop of about 12 °C is observed on the second replacement. In general, it seems that the loss of the formal charge has a greater impact on the triplex stability than the distortion by the bulky purine residue.

As seen for thymines, the combined influence of two replacements does not appear to be simply additive (*cf.* Figures 7b and 8). It should be pointed out that the triple helix (9-mer) does not exist above room temperature when as few as two cytosines are exchanged by inosines. In other words ion pair formation seems to be the major stabilizing factor in intramolecular triplexes formed from mixed sequence.

(G) The Effect of a Combined Mutation. Replacing a neighboring cytosine and thymine with inosines (MII) lowers the T_m for the triplex hairpin transition at pH 6, 100 mM [Na^+], by about 29 °C, indicating that this triplex is slightly more stable than sequence MIT which has an isolated cytosine replaced (*cf.* Figure 8). The slope (dT_m *vs* d(pH)) of this sequence (MII), however, is very similar to the slope of the latter (MIT), underlining the dominant effect of the loss of a formal charge on T_m. The observed relative increase in stability as compared to MICI may be due to an increase in stacking interaction between the two consecutive inosines.

CONCLUSIONS

Although able to bind to both A·T and G·C base pairs in the Hoogsteen mode, inosine replacement of either pyrimidines destabilizes the triple helix conformation. This reflects poorly on its potential use as a "wild-card" in the Hoogsteen position, and a modified base would perhaps be more suited to this role. These results support the findings of Griffin and Dervan (1989) and Shimizu *et al.* (1994). Since triple helix stability is sensitive to backbone distortion and/or decrease in stacking interactions within the third strand, these factors will have to be taken into account when designing and implementing a modified base in site-directed oligonucleotide probes.

Replacing third-strand thymines and/or cytosines in a canonical intramolecular triple helix reduces the binding energy of the Hoogsteen third strand due to distortion of the backbone and reduced stacking interaction.

Substituting cytosine with inosine reduces triplex stability further in favor of the hairpin helix because of the loss of an ion pair between the formally protonated cytosine and the sugar phosphate backbone.

Two consecutive inosines, even if they replace a neighboring thymine and a protonated cytosine, lower the binding

energy of the third strand to a lesser extent than the replacement of two separate thymines or two separate cytosines. It seems that some stacking is gained from the neighboring inosines, making the backbone distortion a bit less pronounced.

ACKNOWLEDGMENT

We thank the lab of the late Professor Dawie Botes (UCT) for synthesizing the oligonucleotides and the Department of Chemistry, Rutgers University, for use of their CD spectrophotometer.

REFERENCES

Arnott, S., & Bond, P. (1973) Triple-Stranded Polynucleotide Helix Containing Only Purine Bases. *Science 181*, 68−69.

Arnott, S., & Selsing, E. (1974) Structures for the polynucleotide complexes poly(dA)·poly(dT) and poly(dT)·polyd(A)·poly(dT). *J. Mol. Biol. 88*, 509−521.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Breslauer, K. J., Frank, R., Blöcker, H., & Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746−3750.

Bush, C. A. (1974) Ultraviolet spectroscopy, circular dichroism and optical rotatory dispersion. In *Basic Principles in Nucleic Acid Chemistry*, Vol. 2, pp 91−169, Academic Press, New York.

Cantor, C. R., Warshaw, M. M., & Shapito, H. (1970) Oligonucleotide interactions. III. Circular dichroism studies of the conformations of deoxyoligonucleotides. *Biopolymers 9*, 1059−1077.

Felsenfeld, G., Davies, D. R., & Rich, A. (1957) Formation of a three stranded polynucleotide molecule. *J. Am. Chem. Soc. 79*, 2023−2024.

Francois, J. C., Saison-Behmoaras, T., Thuong, N. T., & Hélène, C. (1989) Inhibition of restriction endonuclease cleavage via triple helix formation by homopyrimidine oligonucleotides. *Biochemistry 28*, 9617−9619.

Fresco. J. R., & Massoulie J. (1963) Polynucleotides .v. helix−coil transition of polyriboguanylic acid. *J. Am. Chem. Soc. 85*, 1352.

Griffin, L. C., & Dervan, P. B. (1989) Recognition of Thymine·Adenine Base Pairs by Guanine in a Pyrimidine Triple Helix Motif. *Science 245*, 967- 971.

Häner, R., & Dervan, P. B. (1990) Single-stranded DNA Triple-Helix formation. *Biochemistry 29*, 9761−9765.

Hélène, C., Thuong, N. T., Saison-Behmoaras, T., & Francois, J. C. (1989) Sequence-specific artificial endonucleases. *Trends Biotechnol. 7*, 310−315.

Hoogsteen, K. (1959) The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr. 12*, 822−823.

Hüsler, P. L., & Klump, H. H. (1994) Unfolding of a Branched Double-Helical DNA Three-way Junction with Triple-Helical Ends. *Arch. Biochem. Biophys. 313*, 29−38.

Hüsler, P. L., & Klump, H. H. (1995) Prediction of pH-Dependent Properties of DNA Triple Helices. *Arch. Biochem. Biophys. 317*, 46−56.

Johnson, K. H., Grav, D. M., & Sutherland, J. C. (1991) Vacuum UV CD spectra of homopolymer duplexes and triplexes containing A-T or A-U basepairs. *Nucleic Acids Res*, *19*, 2275−2280.

Klump, H. H. (1987) Energetics of order/order transitions in nucleic acids. *Can. J. Chem. 66*, 804−811.

Klump, H. H. (1988) Conformational transitions in nucleic acids. In *Biochemical Thermodynamics* (Jones, M. N., Ed.) 2nd ed., pp 100−144, Elsevier, New York.

Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., & Fresco, J. R. (1988) Specificity in formation of triple stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry 27*, 9108−9112.

Macaya, R. F., Gilbert, D. E., Malek, S., Sinsheimer, J. S., & Feigon, J. (1991) Structure and stability of X·G·C mismatches in the third strand of intermolecular triplexes. *Science 254*, 270−274.

Maher, L. J., III, Wold, B., & Dervan, P. B. (1989) Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. *Science 245*, 725−730.

Manzini, G., Xodo, L. E., Gasporotto, D., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1990) Triple Helix formation by oligopurine−oligopyrimdine DNA fragments; electrophoretic and thermodynamic behavior. *J. Mol. Biol. 213*, 833−843.

Marky, L. A., & Breslauer, K. J. (1987) Calculating Thermodynamic Data for Transitions of any Molecularity from Equilibrium Melting Curves. *Biopolymers 26*, 1601−1620.

Moser, H. E., & Dervan, P. B. (1987) Sequence-specific cleavage of double helical DNA by triple helix formation. *Science 238*, 645−650.

Pilch, D. S., Levensen, C., & Schafer, R. H. (1990) Structural analysis of the $d(A)_{10}$-$2(dT)_{10}$ triple helix. *Proc. Natl. Acad. Sci. U.S.A. 87*, 1942−1946.

Plum, G. E., & Breslauer, K. J. (1995) Thermodynamics of an Intramolecular DNA triple Helix: A Calorimetric and Spectroscopic Study of the pH and Salt Dependence of Thermally Induced Structural Transitions. *J. Mol. Biol. 248*, 679−695.

Plum, G. E., Park, Y. W., Singleton, S. F., & Breslauer, K. J. (1995) Nucleic acid hybridization: triplex stability and energetics. *Annu. Rev. Biophys. Biomol. Struct. 24*, 319−350.

Radhakrishnan, I., & Patel D, J. (1994) DNA Triplexes: Solution Structures, Hydration Sites, Energetics, Interactions, and Function. *Biochemistry 33*, 11405−11416.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association and release, screening, and ion effects on water activity. *Q. Rev. Biophys. 11*, 103−178.

Roberts, R. W., & Crothers, D. M, (1992) Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science 258*, 1463−1466.

Scaria, P. V., & Schafer, R. H. (1991) Binding of ethidium bromide to a DNA triple helix: evidence for intercalation. *J. Biol. Chem. 266*, 5417−5423.

Shimizu, M., Inoue, H., & Ohtsuka, E. (1994) Detailed Study of Sequence-Specific DNA Cleavage of Triplex-Forming Oligonucleotides Linked to 1,10-Phenanthroline. *Biochemistry 33*, 606−613.

Sklenár, V., & Feigon, J. (1990) Formation of a stable triplex from a single DNA strand. *Nature (London) 345*, 836−838.

Snell, F. D., & Snell, C. T. (1949) *Colorimetric analysis*, 3rd ed., Vol. 2, p 671, Van Nostrand, New York.

Strobel, S. A., & Dervan, P. B. (1990) Single-site enzymatic cleavage of yeast chromosome by oligonucleotide-directed triple-helix formation. *Science 249*, 72−75.

Thielle, D., & Guschlbauer, W. (1968) Evidence for a three stranded complex between poly I and poly C. *FEBS Lett. 1*, 173−175.

Völker, J. (1993) The impact of local and global composition on the stability of triple helix formation, Ph.D. Thesis, University of Cape Town, South Africa.

Völker, J., Botes, D. P., Lindsey G. G., & Klump, H. H. (1993) Energetics of a stable intramolecular DNA triple helix. *J. Mol. Biol. 230*, 1278−1290 [JV-ITS has the sequence d(5′-GAGAGAGAAACCCCTTTCTCTCTCTTTTCTCTCTCTTT-3′)].

Watson, J. D., & Crick, F. H. C. (1953) Genetic implications of the structure of deoxyribonucleic acids. *Nature (London) 171*, 964−967.

Xodo, L. E., Manzini, G., & Quadrifoglio, F. (1990) Spectroscopic and calorimetric investigation on the DNA triplex formed by d(CTCTTCTTTCTTTTCTTTCTTCTC) and d(GAGAAGAAA-GA) at acidic pH. *Nucleic Acids Res. 18*, 3557−3564.

BI960193W