The Impact of global and local Composition on the Stability of Triple Helical DNA

Jens Völker

October 1993

Thesis presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry

UNIVERSITY OF CAPE TOWN

ABSTRACT

It is common practise in antisense technology to view third strand binding to be controlled by the same principles which are found to determine the stability of the double helix. In contrast to this view based on a general consideration of the various forces contributing to the binding energy of the third strand it was proposed that the dominant contributions will originate from electrostatic interactions. These electrostatic contributions can be subdivided into sequence independent repulsive forces between the negatively charged backbones and into sequence dependent attractive forces between the positively charged protonated Hoogsteen cytosines and the backbone phosphates. The observable changes in the stability of triple helices should be a reflection of the number (global composition) and distribution (local composition) of cytosines in the third strand.

To this aim two families of 38-mer oligonucleotides were synthesized, which have as a common design feature a linear array of 10 homopurine bases followed by 10 homopyrimidine bases as Watson & Crick complementary strand to the homopurine region and ending in a 10 homopyrimidine residue stretch which binds to the W&C helix via Hoogsteen base-pairing. This arrangement of homopurine and homopyrimidine sections with connecting pyrimidine linkers allows the formation of intramolecular triple helices of predetermined stoichiometry and strand orientation.

Physical (UV-spectroscopy, CD-spectroscopy and fluorimetry) and biochemical techniques (P_1-nuclease digestion) have been used to show that the oligonucleotides undergo a stepwise folding process from a random coil into a hairpin with 3'dangling tail and then into a intramolecular triple helix. This folding occurs as a function of pH and/or ionic strength. The effect of local and global composition on the stability of the three conformational transitions has been evaluated from a comparison of the melting temperatures and the behavior of the phase boundaries of the different oligonucleotides.

As the result of this thesis the following general rules emerge: The stability of the third strand depends on the particular combination of sequence, pH and ionic strength. At physiological conditions (pH 7.1, 150 mM Na^+) thymines and cytosines contribute equally to the stability (global effect) provided that the cytosines are spaced by more than one thymine. (local effect). Below pH 7.1 (150 mM Na^+) the stability increases linearly with the number of cytosines and at pH above pH 7.1 (150 mM Na^+) it decreases. At

ionic strength below 400 mM Na^+ (pH 6.75) the stability increases with the number of cytosine while above 400 mM Na^+ (pH 6.75) it decreases.

Based on these results a rational approach for the design of oligonucleotide third strands and the choice of appropriate environmental conditions for the formation of a particular triple helix becomes feasible.

Acknowledgments

A Ph.D.-thesis can not be completed without the help and input of many people, be it that they give advice, provide technical expertise, discuss concepts or simply provide moral support. My thanks belong to everyone who over the years has helped me to finally reach this goal in my scientific career. While I am indebted to many, the contributions of some stand out and need to be highlighted especially:

First and foremost I wish to express my gratitude to my supervisor **Prof. Dr. Horst Klump**. Without his enthusiasm for my project and his encouragement whenever it was needed, this thesis would have been impossible to complete.

My special thanks belong to **Dr. Dennis Maeder**, who was always prepared to listen to my ideas, and to discuss the merits or otherwise of each in detail. Dennis also taught me all I know of computers and was tireless in correcting the many flaws in my use (or abuse) of the English language. Any lack of proficiency in either subject is entirely due to my own deficiencies.

I would like to thank **Prof. Dawie Botes** for the synthesis of the oligonucleotides and also for his advice with regard to the purification of these oligos. To develop the purification methodology was a long and very tedious process, but, I think, worth the effort in the end!

I wish to thank all past and present members of my laboratory for their camaraderie and friendship over the past five years. Especially **Paul Hüsler**, who always managed to make me try harder. Paul - despite our countless scientific and philosophical arguments over the years (or perhaps because of them) I think we have been a good team!

My special thanks also go to **Sipho Hlati**, who tirelessly put up with my whims regarding the contents and layout of the figures, even though some of it did not agree with his artistic sense.

I further wish to thank the members of the workshop Herman Alk, Uli Mutzek and James Duncan for converting my conceptual ideas into working systems and for all the repair work carried out over the years.

Last but not least I wish to express my gratitude to my parents **Dieter** and **Inge Völker**. Without their financial and moral support over the years none of this would have been possible.

I gratefully acknowledge financial support received by the ***Foundation for Development and Research*** **FRD** (Pretoria, South Africa) and the ***Gottlieb Daimler and Carl Benz Stiftung*** (Ladenburg, Germany).

dedicated

to

my father

TABLE OF CONTENTS

TABLES AND FIGURES

FIGURES

TABLES

ABBREVIATIONS

%C	-	Percent cytosines bases in the third strand
%C^+GC	-	Percent C^+•G•C base-triplets in the triple helix
%GC	-	Percent G•C base-pairs in the hairpin stem
A	-	Adenine
bp	-	Base-pair
bt	-	Base-triplet
C	-	Cytosine
C^+	-	Protonated Cytosine
CD	-	Circular Dichroism
DTT	-	Dithiothreitol
EDTA	-	Ethylendiaminetetra-acetic acid
Fe-EDTA	-	Iron-ethylenbdiaminetetra-acetic acid complex
FTIR	-	Fourier Transform Infra Red Spectroscopy
G	-	Guanine
HPLC	-	High Performance Liquid Chromatography
I	-	Inosine
log	-	logarithm to the base 10
Na^+	-	Sodium cation
[Na^+]	-	Concentration of sodium cations
NMR	-	Nuclear Magnetic Resonance
PAGE	-	Polyacrylamide gel electrophoresis
T	-	Thymine
U	-	Uracil
UV	-	Ultra-violet

INTRODUCTION

1.1 Polynucleotide Triple Helices: The early days

The discovery of Triple Helices:

Shortly after the discovery of the double helical structure of native DNA (Watson & Crick (1953)) an enzyme (polynucleotide phosphorylase) was isolated from *Azotobacter Vinlandii* (Gruneberg-Manago et al (1956)), which allowed the enzymatic synthesize of homopolynucleotides (RNA polymers) of almost any desired repetitive sequence. Biophysical studies of complementary homopolymers revealed that the resulting structure resembled the DNA double helix, (e.g poly r(A) was found to combine with poly r(U) to form poly r(A) • poly r(U) (Warner (1957))). Continuous variation titrations (Job titrations) allows the stoichiometry of the resultant complexes to be determined. In addition to the 1:1 complex, the titration of poly r(A) with poly r(U) in the presence of 0.01 M magnesium ions revealed a novel nucleic acid structure composed of one poly r(A) strand and two poly r(U) strands, poly r(U) • poly r(A) • poly r(U)[1] (Felsenfeld et al (1957)). This three-stranded nucleic acid structure received considerable attention, because it was assumed that this new triple helical structure might somehow be involved in transcription (Rich (1960), Zubay (1962)) and/or chromatin condensation (Pettijohn & Hecht (1973)). At that time the process of transcription was not yet understood.

[1] Throughout this thesis the following notation will be used: The polymer / oligomer strand in the Hoogsteen position of the triple helix is listed first, followed by the purine strand of the Watson & Crick base pair. The third position denotes the Watson & Crick pyrimidine strand.

Formation of Triple Helices between poly A and poly U/T

After deoxyribo-homopolymers became readily available (Radding et al (1962), Byrd et al (1965)), it was established that poly d(A) and poly d(T) can also associate to form the analogous triple helix poly d(T) • poly d(A) • poly d(T). Finally it was established that hybrid structures consisting of a combination of deoxyribose and ribose homopolymers could form triple helices as well (Riley et al. (1966)). Of all the possible combinations of poly r(A) (or poly d(A)) with poly r(U) and/or poly d(T) (or poly d(U) and/or poly r(T)) only the triple helix poly d(T) • poly d(A) • poly r(U) was never observed. This can be understood if we assume that the poly d(A) • poly r(U) double helix disproportionates to the triple helix poly r(U) • poly d(A) • poly r(U) rather than bind the deoxy strand (poly d(T)) in the Hoogsteen position.

Formation of Triple Helices between poly G/I and poly C

In contrast to the poly r(A) / poly r(U) system Job titrations of poly G with poly C do not reveal double strand and/or triple strand formation. The reason for this is that poly(G) homopolymers (both the ribose and deoxyribose polymers) were found to form very stable structures in solutions consisting of an association of four poly(G) strands (Fresco & Massoulie (1963)). The introduction of monomers or oligomers $r(G)_n$, for which the association of the guanine strands is sufficiently weakened to allow the formation of the oligo r(G) • poly r(C) double helix, also led to the detection of two new triple helices. At neutral or slightly basic pH oligo r(G) and poly r(C) form the oligo r(G) • oligo r(G) • poly r(C) triple helix. Below neutral pH a different triple helix is formed with protonated cytosines in the Hoogsteen position (poly $r(C^+)$ • oligo r(G) • poly r(C)) (Lipsett (1964)). Several years later Thiele et al.(1978) succeeded in the formation of the

poly d(G) • poly d(G) • poly d(C) triple helix at neutral and alkaline pH and the poly d(C^+) • poly d(G) • poly d(C) triple helix at acidic pH (Marck & Thiele (1978)).

Since the guanine analogue inosine does not form a poly$(I)_4$ structure of comparable stability, the formation of the corresponding triple helices was detected much earlier. Inman (1964) reports the formation of the following triple helices at neutral pH: poly r(I) • poly r(I) • poly r(BrC) and poly r(I) • poly r(I) • poly r(C). In acidic conditions the formation of the poly r(C^+) • poly r(I) • poly r(C) complex was also observed by Thiele & Guschlbauer (1968, 1969). The homopolypurine poly r(I) differs from the other sequences in that it can combine with poly r(A) to form the unusual triple helix poly r(I) • poly r(A) • poly r(I), which consists of purine strands only. (Rich (1958), Howard & Miles (1977), Arnott & Bond (1973)).

Incorporation of Strictly Alternating Dinucleotides into Triple Helices:

Morgan & Wells (1968) demonstrated that the dinucleotide repeat sequence poly r(CU) can be incorporated into a triple helix with poly d(GA) • poly d(CT) if the pH is lowered below neutral. Double helical poly d(GA) • poly d(CT) can undergo a conformational change at acidic pH. The new conformation was at first considered to be a four stranded helix ((Johnson & Morgan (1977). Subsequently it was shown that the double helix poly d(GA) • poly d(CT) in fact disproportionates to the triple helix poly d(C^+T) • poly d(GA) • poly d(CT) and single stranded poly d(GA) (Lee et al. (1979). Provided that the cytosine residues are 5' methylated, this disproportionation can already occur at pH 7.0 (Lee et al. (1984)). All attempts to incorporate alternating purine-pyrimidine sequence into triple helical complexes have failed so far (Wells et al. (1970)). A general rule emerged that only a homopurine or a homopyrimidine polymer as the third strand can associate with double helical homopurine • homopyrimidine complexes.

The exclusion of alternating purine - pyrimidine sequences from triple helix formation implies that triple helical arrangements could hardly be of any biological importance. Consequently in the late 1960's and early 1970's the interest in triple helices waned and they were considered an oddity peculiar to homopurine • homopyrimidine polynucleotides with no biological relevance. The early investigations of homopolymer RNA and DNA complexes including triplexes have been reviewed by Michelson et al (1967) and Felsenfeld & Wells (1967).

The Triple Helix Poly A • Poly A • Poly U

Fairly recently a new triple helix has been proposed (Broitman et al. (1987)), which consists of two strands poly r(A) and one strand poly r(U). This triple helix had not been detected earlier, because of the unusual size dependence of the third strand poly r(A). A third strand poly r(A) longer than 160 bases or an oligo r(A) strand shorter than 30 bases fails to form the poly r(A) • poly r(A) • poly r(U) triple helix. There is obviously severe sterical hindrance for third strand homopurines. The following possible homopurine • homopurine • homopyrimidine complexes [poly(A) • poly(G/I) • poly(C)], [poly(G) • poly(A) • poly(T/U)] have not been observed thus far. So have the homopyrimidine • homopurine • homopyrimidine triple helices [poly(U/T) • poly(G/I) • poly(C)], and [poly(C) • poly(A) • poly(T/U)].

The Third Strand Binding Code:

Based on all the observations listed above Letai et al in 1988 proposed a "third strand recognition code". This code is listed in **Table I**. The "third strand recognition code" shows some resemblance to Chargaff's rules, in that adenine and thymine/uracil recognize only AT base-pairs and cytosine and guanine recognize only GC base-pairs.

However unlike the complementarity rules for double helices, which always require a pyrimidine to bind to a purine, the "third strand binding code" allows for the binding of both purine and/or pyrimidine bases to the Hoogsteen position of the Watson & Crick purine. The synthetic guanine analogue inosine serves as a "wild card" for this "code", since it recognizes both AT and GC base-pairs.

Table I: Third-Strand Binding Code[2]

	Third Strand Residue				
Watson-Crick Core	A	T/U	I	G	C
A • T/U	✓	✓	✓	×	×
G • C	×	×	✓	✓	✓

Summary

The properties of polynucleotide triple helices can be summarized as follows:

1.) Triple helix formation has only been shown for homopurine • homopyrimidine sequences.

2.) Triple helices can accommodate ribose and deoxyribose backbones in either of the double helix and the third strand position.

3.) Provided that the "third strand binding code" is adhered to, both homopurine or homopyrimidine polymers can bind as third strands.

---

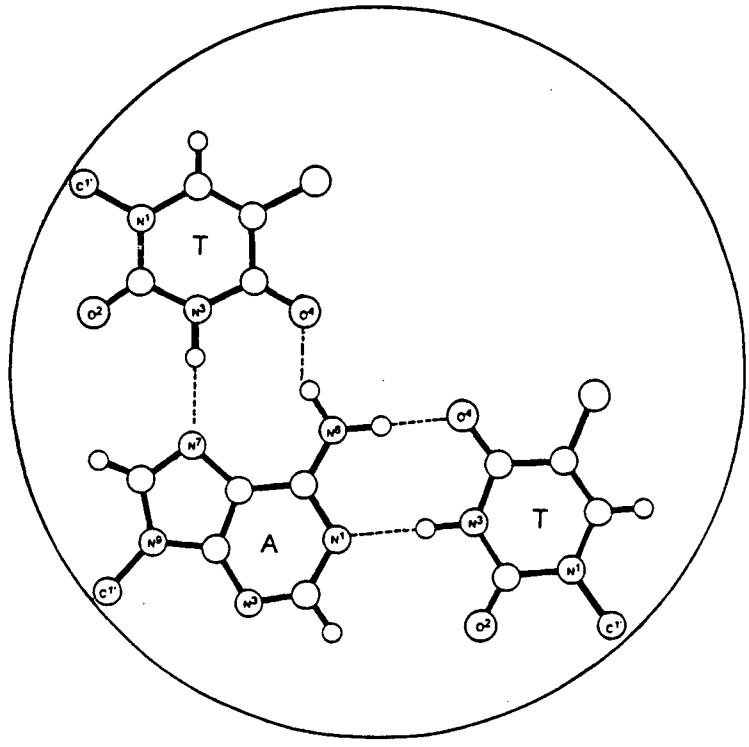
[2] ex Letai et al. (1988)

1.2 Properties of Polynucleotide Triple Helical Structures

How does the Third Strand Poly(U/T) Recognize a Poly(A) • Poly(U/T) Duplex ?

Model building showed that the third strand could wind itself around the major groove of a Watson and Crick double helix and could link itself via hydrogen-bonds to the hydrogen-bond donors and acceptors sites protruding into this groove (Rich (1958)). The models could not discriminate between two possible orientations for the third base. In the first single crystal x-ray structure solved for a base-pair it was shown that the pyrimidine is hydrogen-bonded to the hydrogen-bond donors and acceptors of the purine which commonly face the major groove in the double stranded Watson & Crick DNA. The 1-Methylthymine forms two hydrogen-bonds between its N3H and C4O positions and the N7 and C6NH_2 positions of 9-methyladenine (Hoogsteen (1959). This type of base-pairing will henceforth be referred to as Hoogsteen base-pairing. The alternative hydrogen-bonding scheme requires that the C2O group of thymine hydrogen-bonds to the C6NH_2 group of adenine and the thymine N3H hydrogen-bonds to the N7 position of adenine ring. Because this base-pairing scheme can be considered to derive from flipping over the thymine ring of the Hoogsteen basepair, it is commonly called reversed Hoogsteen base-pair (**Figure 1.1**). The major differences between Hoogsteen and reversed Hoogsteen base-pairing is the orientation of the pyrimidine strand with respect to the purine strand. In Hoogsteen base-pairing the third strand is orientated in parallel to the purine strand of the double helix, while in reversed Hoogsteen base-pairing the orientation is antiparallel. Based on infrared spectroscopy Miles has proposed that in the poly r(U) • poly r(A) • poly r(U) triple helix the poly r(U) strand binds via reversed Hoogsteen base-pairing to the major groove of the double helix (Miles (1964)).

A



B

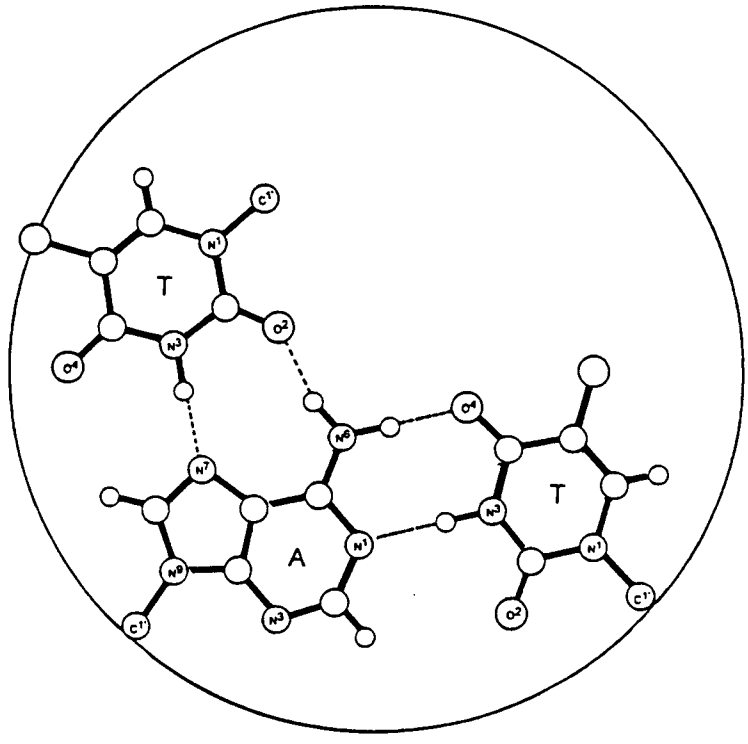


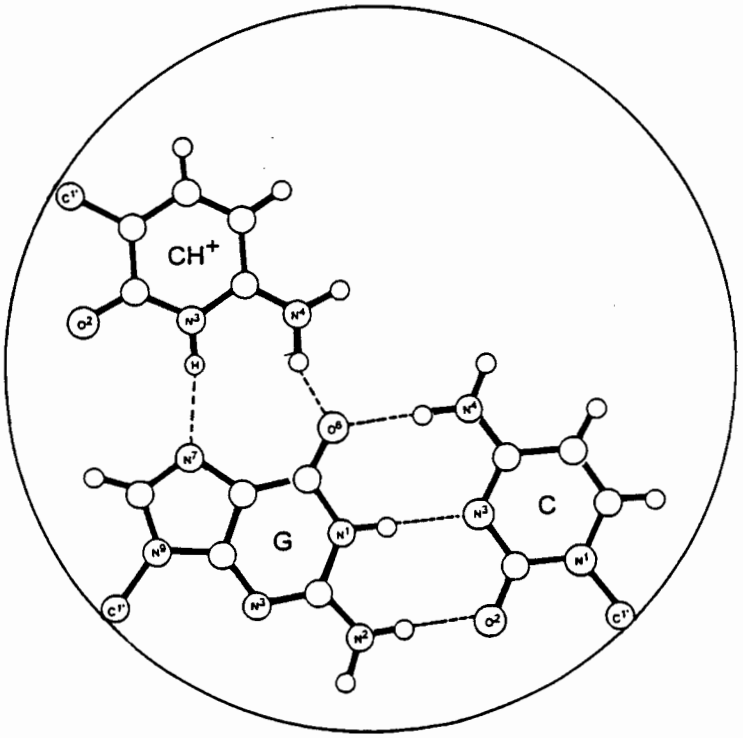
Figure 1.1: A) Model of the hydrogen-bonding scheme of the Hoogsteen triplet T•A•T
B) Model of the hydrogen-bonding scheme of the reverse Hoogsteen triplet T•A•T

However x-ray fiber diffraction studies of the poly r(U) • poly r(A) • poly r(U) (Arnott & Bond (1973b)) confirm that the third strand poly r(U) is bound via Hoogsteen base-pairing to the double helix.

Third Strand Cytosines Bind via Hoogsteen Hydrogen-bonding

From **Figure 1.2** it can be seen that cytosine is indeed capable of forming both Hoogsteen and reverse Hoogsteen base-pairs with a Watson & Crick guanine, provided that the N3 position of the cytosine is protonated. Hence the requirement for mildly acidic conditions for the formation of the poly(C) • poly(G) • poly(C) triple helix (Lipsett (1964)). The protonated cytosine in the Hoogsteen orientation is isosteric with a Hoogsteen thymine/uracil, while the protonated cytosine in the reversed Hoogsteen orientation is not isosteric to a reversed Hoogsteen thymine/uracil. Hence a triple strand consisting of both thymines/uracils and cytosines bound via reverse Hoogsteen hydrogen-bonding to the double helix cannot have a smooth backbone conformation, and is therefore unlikely to form. Since homopolymers with a repeating uracil - cytosine dinucleotide sequence are incorporated as third strands (Morgan & Wells (1968)) it is very likely that both pyrimidines bind via the same base-pairing scheme. Subsequently it has been shown that the introduction of an C8-amino group at the guanine substantially increased the T_m of poly(C) • ($8NH_2$-GMP) • poly(C) triple helix (Hattori et al. (1976)). This is only possible if the cytosine binds in the Hoogsteen orientation, since only then is it possible for the C2O carbonyl group to form an extra hydrogen-bond to the guanine C8-NH_2 group and give rise to the increased stability.

A



B

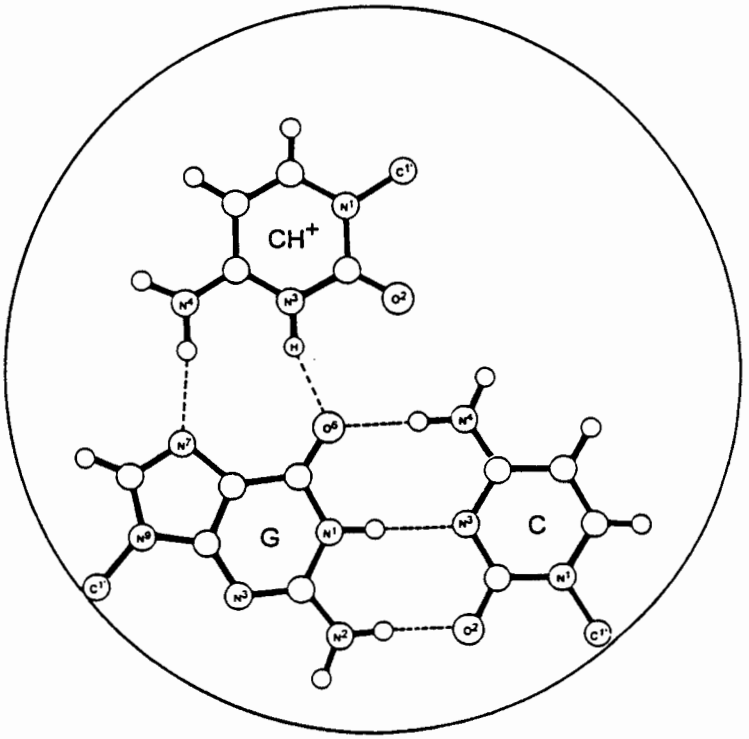


Figure 1.2: A) Model of the hydrogen-bonding scheme of the protonated Hoogsteen triplet C^+•G•C
B) Model of the hydrogen-bonding scheme of a protonated reverse Hoogsteen triplet C^+•G•C

Homopurines can attach as Third Strand

The hydrogen-bonding scheme of the three bases, which form the planar triplet is open to debate as far as the particular hydrogen-bonding of the "Hoogsteen" homopurine strand is concerned. Four different homopurine triple helical conformers are theoretically conceivable, since the purines can occur in the syn as well as anti conformation in helical structures, and because in principle each conformation can bind via Hoogsteen and reversed Hoogsteen hydrogen bonding to the Watson & Crick purine strand. Some of these conformers require unusual hydrogen-bonding types, such as hydrogen-bonds between third strand purines and the Watson & Crick pyrimidines, potentially even bridged by water molecules. One model even requires a particularly weak hydrogen-bond between a CH group of one base and a ring nitrogen in a complementary base. Despite these unusual hydrogen-bonds required for some of the third strand base-pairing schemes they cannot be rejected out off hand. The sole homopurine • homopurine • homopyrimidine triple helix for which x-ray fiber diffraction data exist is the unusual all purine triple helix poly r(I) • poly r(A) • poly r(I) (Arnott & Bond (1973a)). In this triple helix the third strand inosine assumes the anti conformation and binds via Hoogsteen type hydrogen bonds to the N7 and C6NH_2 positions of poly r(A). Hence the poly r(I) third strand must bind parallel to the poly r(A) strand. The poly r(A) • poly r(I) double helix is proposed to be in an A-DNA type conformation in order to accommodate the third strand in the major groove. The similarities between the x-ray fiber diffraction pattern of poly r(I) • poly r(A) • poly r(I) and poly r(U) • poly r(A) • poly r(U) have been interpreted to imply that polypurine third strands are conformationally similar to polypyrimidines. Despite the fact that this triple helix is formed from three bulky purine strands, the diameter of the poly r(I) • poly r(A) • poly r(I) triple helix is very similar to that of A'-DNA.

Triple Helices have an A-DNA Type Conformation.

The X-ray fiber diffraction studies of the all ribo poly r(U) • poly r(A) • poly r(U) (Arnott & Bond (1973b)) and the all deoxyribose poly d(T) • poly d(A) • poly d(T) (Arnott & Selsing (1974)) allowed to establish the helix parameters of these triple helical structures. These triple helices have a helical repeat of 12 base-triplets per turn and a pitch of 3.26 Å. Because these values are more compatible with the helix parameters of an A-DNA structure than that of a B-DNA structure it was concluded that the double helix component of triple helices must form an A-DNA like structure. When the triple helix is member of the A-DNA family, the sugar puckers has to be *C3'endo*. Unfortunately the resolution of the fiber diffraction pattern was insufficient to allow an experimental determination of the sugar pucker. From a subsequent model building study based on the limited x-ray fiber diffraction data Arnott et al (1976) concluded that a triple helix can only form if the third strand binds to the major groove of an A-DNA double helix, even though a number of sterically unfavorable contacts occurred. The major groove of a conventional B-DNA double helix was considered to be to narrow to accommodate the third strand. Interestingly the diameter of the triple helix in the model is of the same order of magnitude as that of the host A-DNA double helix. Because there were no more precise data available at that time this model has been universally accepted.

Evidence for B-DNA Conformation in Deoxyribose Triplexes

Recently Park & Breslauer (1992) have shown that the minor groove binding ligand netropsin can bind to a poly d(T) • poly d(A) • poly d(T) triple helix, albeit with a reduced binding density. This is associated with a reduction of the thermal stability of the

triple helix. The interdependence between the minor groove binding ligand and the major groove binding third strand has been termed "groove-cross-talk". The minor groove binder netropsin is shown to bind to B-DNA only (Zimmer & Wähnert (1986) and references therein). Consequently the binding of netropsin to the minor groove of this triple helix is inconsistent with the assumption that the triple helix is in an A-DNA conformation. Recent Fourier Transform Infrared Spectra (FTIR) obtained for all combinations of deoxyribose and ribose polymers poly(T/U) • poly(A) • poly(T/U) triple helices by Liquier et al (1991) have shown that the preferred sugar pucker in the all deoxyribose triple helix is *C2'-endo*. This is consistent with a B-DNA structure. The FTIR spectra of the all ribose triple helix poly r(U) • poly r(A) • poly r(U) on the other hand has contained bands characteristic of the *C3'-endo* sugar pucker of an A-type conformation. The former result has been confirmed independently by Howard et al (1992). Recently Raghunathan et al (1993) have remodeled the structure of the poly d(T) • poly d(A) • poly d(T) triple helix based on the parameters of rise and pitch determined by Arnott & Selsing (1974) and a *C2'-endo* sugar pucker. A stereochemically satisfactory triple helix model with the general B-DNA geometry has been obtained.

Interactions with the Intercalating Ligand Ethidium Bromide

Additional evidence for the formation of an altered B-DNA conformation comes from the observation that the intercalating drug ethidium bromide interacts differently with double stranded and triple stranded DNA molecules. Lee et al (1979) showed that ethidium bromide does not bind to the triple helical poly d(C^+T) • poly d(GA) • poly d(CT) at low pH, while it does bind to the double stranded poly d(GA) • poly d(CT). They attributed this lack of intercalation to the triple helix conformation and to the presence of positive charges in the third strand cytosines. Subsequently it was

noted that ethidium bromide does intercalate in the unprotonated triple helix poly d(T) • poly d(A) • poly d(T), albeit with a reduced binding density (Scaria & Shafer (1991)).

Physical characterization of the Poly r(U) • Poly r(A) • Poly r(U) Triple Helix :

Among the triple helices poly r(U) • poly r(A) • poly r(U) has been studied first and foremost. Miles (1960, 1964) could show that because of hydrogen-bond formation the infrared spectra of the double helix and the triple helix differed. Based on the differences in the infra red spectra Miles & Frazier (1964) then showed that at low salt concentrations the presence of the structure depended on the input ratio of the single stranded poly r(A) and poly r(U). At a ratio of 1:1 only the double helix is formed, while at a ratio of 1:2 the triple helix is formed preferentially. Subsequently it was established that the double stranded poly r(A) • poly r(U) can undergo a temperature dependent disproportionation reaction to result in the triple stranded poly r(U) • poly r(A) • poly r(U) and free poly r(A). This reaction occurs at ionic strength above 200 mM Na^+. (Miles & Frazier (1964)). Stevens & Felsenfeld (1964) and Blake et al. (1967) showed that one can discriminate between the triple helical state and the double helical state of poly r(A) • poly r(U) by choosing the appropriate wavelength in the UV-range.

Temperature dependent order/order and order/disorder transitions: disproportionation reaction

By following the temperature dependent change in absorbance (at the appropriate wavelength) at varying ionic strengths it became evident that the final state of the poly r(A) • poly r(U) system depends on the boundary conditions (Blake et al. (1967)). If the double helical complex poly r(A) • poly r (U) at low ionic strength (<200 mM Na^+) is heated above the T_m the double helix dissociates into single strands (2→1

transition). At intermediate ionic strength a temperature increase leads to a disproportionation reaction, i.e. two poly r(A) • poly r(U) double helices rearrange to give a triple helical complex poly r(U) • poly r(A) • poly r(U), leaving free poly r(A) in solution (2→3 transition). By raising the temperature further the triplex then dissociates to its constituent single strands in a single step (3→1 transition). Because of a decreased repulsion of the backbones, the disproportionation reaction occurs at decreasing temperature as the ionic strength is increased. The melting temperature of the disproportionation reaction decreases to such an extent that it can occur below room temperature for some salt concentration (**Figure 1.3**).

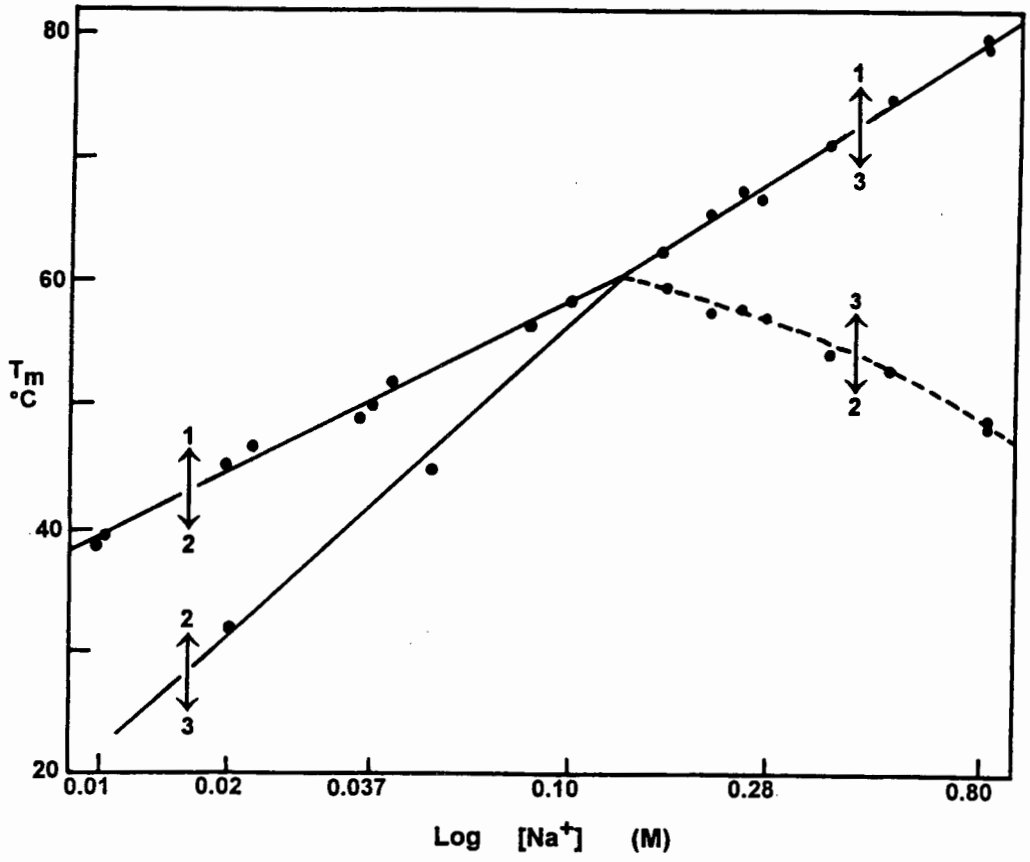


Figure 1.3: Phasediagram : Plot of transition temperature as a function of cation concentration for the various conformations of poly r(A) • poly r(U).
1 ↔ 2 - duplex ↔ coil transition
3 ↔ 2 - triplex ↔ duplex transition
3 ↔ 1 - triplex ↔ coil transition,
2 ↔ 3 - duplex ↔ triplex disproportionation reaction
(Adapted from Krakauer & Sturtevant (1968))

Temperature dependent order/order and order/disorder transitions: addition reaction

At very low ionic strength an input ratio of poly r(A) to poly r(U) of 1:2 does not result in the formation of the triple helix as might be assumed. Thermal denaturation studies will detect only the melting of the double helix (2→1 transition). At intermediate ionic strength the triple helix does form, but the third strand will dissociate upon raising the temperature (3→2 transition), followed by the melting of the double helix (2→1 transition) at even higher temperature. Above 200 mM Na^+ the triple helix will dissociate directly to its component single strands without any intermediate step (3→1 transition). The different order/order and order/disorder transitions of nucleic acids have been compared to melting processes and hence can be considered as phase transitions. The compiled results can be summarized in a phasediagram (**Figure 1.3**). " *Phasediagrams have two useful applications. From the diagram the conformational state of any polynucleotide in the system at any arbitrary set of external conditions (temperature, cation concentrations) can be taken simply by inspection. From the behavior of the phase boundaries the physical effects responsible for the phase-transition can be assessed*" Klump (1988). Massoulie (1968a,1968b) established that the transition temperatures for the various conformational transitions of poly r(A) and poly r(U) are independent of the pH, provided that none of the single stranded components reacts to form a competing secondary structure.

Thermodynamics of Triple Helix Formation

All these poly r(A) • poly r(U) systems mentioned above have been investigated by calorimetry and the transition enthalpy of the order/order and order/disorder transitions have been determined. From the enthalpy data of the association reaction of poly r(A) and poly r(U) to the poly r(A) • poly r(U) double helix (Rawitscher et al (1963)) and the thermal disproportionation reaction observed by UV-spectroscopy, Stevens & Felsenfeld (1964) have extracted the enthalpy of disproportionation. ΔH_{disp} is of the order of 1.3 - 2.5 kcal/Mbp and the enthalpy of addition of a single strand poly r(U) to a double helix ΔH_{add} is of the order of 5.7 ± 0.7 kcal/Mbp. Subsequently the individual reactions have been investigated directly by calorimetry by Ross & Scruggs (1965), Krakauer & Sturtevant (1968), Neumann & Ackermann (1969) and Klump (1988). The results of these various calorimetric studies are listed in **Table II**. Because triple helices were for a long time considered of little biological relevance, a detailed calorimetric determination of the enthalpy of triple helix formation for other polynucleotide triple helices is still lacking. The available calorimetric results indicate that the enthalpy of adding a third strand poly r(U) to a poly r(A) • poly r(U) double helix is about half the enthalpy of the double helix formation (about 4 kcal/ Mbt).

Table II: Thermodynamic data for the poly r(U) • poly r(A) • poly r(U) Triple Helix [2]

	Reaction	DH_{cal}	Comment
[A]			
1	(A•U) ↔ ½(A) + ½(U•A•U)	+ 3.9 kcal/Mbp	@ T_m
2	(U•A•U) ↔ (A) + 2(U)	+13.0 kcal/Mbp	@ T_m
3	(A•U) ↔ (A) + (U)	+ 8.1 kcal/Mbp	Corrected for Temp
4	(A•U) + (U) ↔ (U•A•U)	- 4.2 kcal/Mbp	calculate from 1 & 2
[B]			
5	(A•U) + (U) ↔ (U•A•U)	- 3.9 kcal/Mbp	measured T independ
6	(A•U) ↔ ½(A) + ½(U•A•U)	+ 2.3 kcal/Mbp	measured T depend.
7	(A) + (U) ↔ (A•U)	- 6.5 kcal/Mbp	@ 37°C
[C]			
8	(A•U) ↔ ½(A) + ½(U•A•U)	+ 5.3 kcal/Mbp	@ 45°C
[D]			
9	(A•U) ↔ ½(A) + ½(U•A•U)	+ 1.4 kcal/Mbp	@ T_m
10	(U•A•U) ↔ (A) + 2(U)	+12.7 kcal/Mbp	@ T_m
11	(A•U) ↔ (A) + (U)	+ 8.2 kcal/Mbp	@ T_m ~0.1 M Na^+
12	(U•A•U) ↔ (A•U) + (U)	+ 1.3 kcal/Mbp	@ T_m
13	(U•A•U) ↔ (A•U) + (U)	+ 3.9 kcal/Mbp	@ T_m=45 calculated
14	(A•U) ↔ ½(A) + ½(U•A•U)	+ 1.7 kcal/Mbp	@ T_m=45 calculated

[A] Neumann & Ackermann (1969) [B] Ross & Scruggs (1965)
[C] Klump (1988) [D] Krakauer & Sturtevant (1968)

[3] in the cases where more than one value for ΔH_{cal} was given at marginally different ionic strength the average was calculated

Summary

1.) The homopyrimidine third strand binds via Hoogsteen hydrogen bonding to a homopurine • homopyrimidine target sequence. This results in the homopyrimidine third strand to be in a parallel orientation w.r.t. the homopurine strand.

2.) The homopurine third strand can potentially bind in the syn or anti conformation by either Hoogsteen or reverse Hoogsteen hydrogen-bonding. This means that there are two hydrogen-bonding schemes were the third strand is in the parallel orientation w.r.t. the Watson & Crick homopurine strand and two were the orientation is antiparallel.

3.) Triple helix formation can occur either by a disproportionation reaction of a homopurine • homopyrimidine double helix or by the addition of an independent third strand to a homopurine • homopyrimidine target sequence, provided the ionic strength is sufficiently high.

4.) Triple helix formation results in a conformational transition of the double helix target sequence. It has been proposed that this is either an A-DNA type or an altered B-DNA type conformation.

1.3 Triple Helices in Plasmids - H-DNA

Intramolecular Triple Helices in Plasmids - H-DNA:

Interest in triple helical DNA conformation has reawakening recently, because advances in DNA technology have indicated that triple strands can potentially occur *in vivo*. This was sparked by the observation that supercoiled plasmids containing homopurine • homopyrimidine inserts show an enhanced sensitivity towards single strand specific nucleases and single strand specific chemical reagents (Wells (1988)). Mapping of the DNA region sensitive to these single strand specific agents reveals a specific cleavage pattern in the homopurine • homopyrimidine region. Characteristically half of the homopurine region and the center of the homopyrimidine region are very susceptible to the single strand specific reagents. This unusual pattern of cleavage sites is considered to be due to the formation of an intramolecular triple helix - also termed H-DNA (Lyamichev et al (1985, 1986) at pH 4.5.

To rearrange the Watson & Crick structure into the H-DNA, half of the pyrimidine strand dissociates from its complementary homopurine strand and binds as third strand to the other half of the homopurine • homopyrimidine region. This leaves half the homopurine region and the center of the homopyrimidine region single stranded, and hence susceptible to single strand specific modification reagents (**Figure 1.4**). (Christophe et al (1985), Lyamichev et al (1985, 1986). An important precondition for the formation of the H-DNA is the occurrence of a symmetrical homopurine and homopyrimidine sequence of sufficient length.

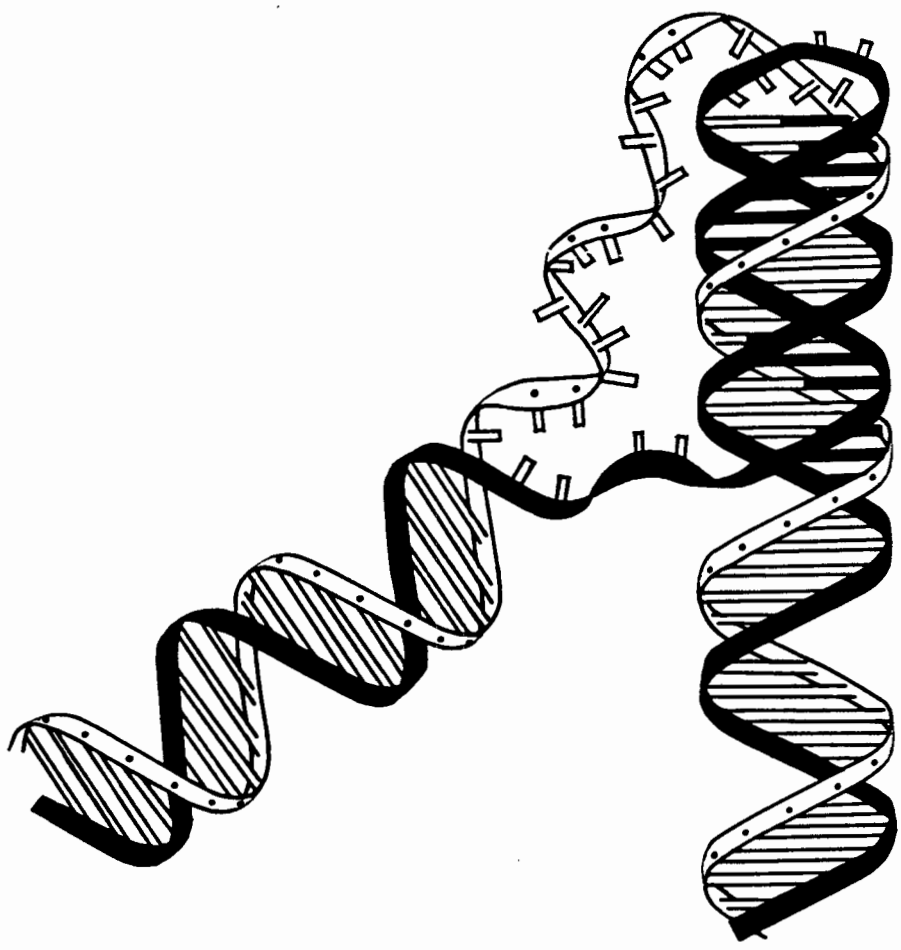


Figure 1.4: Schematic model of the intramolecular triple helix conformation (H-DNA) proposed to occur in supercoiled plasmids at low pH. (adapted from Hanvey et al (1988))

Potential Triple Helix Forming Sequences in Eukaryotes: A Statistical Analysis

Homopurine • homopyrimidine stretches of 10 base-pairs or more are statistically over-represented (e.g. in the human β globin gene and in the viruses of eukaryotes), but they are not over-represented in prokaryotes and their viruses (Behe (1987), Beasty & Behe (1988)). The increased presence of particular DNA sequence motifs in a genome reasonably may be assumed to reflect important biological functions. Consequently it has been proposed that the homopurine • homopyrimidine regions may play an important biological role in eukaryotes.

Summary

1.) Homopurine • homopyrimidine sequences are over-represented in eukaryotic genoms, and they are especially frequent upstream of genes.

2.) Symmetric homopurine • homopyrimidine inserts in supercoiled plasmids take up an altered conformation proposed to be an intramolecular triple helix upon lowering the pH (H-DNA).

1.4 Oligonucleotide Triple Helices

Oligonucleotide Based Enzymes

Triple strand forming oligonucleotides to which a chemically reactive group has been tethered do provide a convenient, easily adjustable and highly sequence specific approach for the recognition and modification of double stranded DNA sequences. In principle they can mimic DNA modifying enzymes such as restriction endonucleases.

All enzymes must have two separate functions in order to catalyze their specific reaction properly: first specific recognition and binding of a substrate is required and second the alteration and release of the product must be favored. The hydrogen bonding capacity of selected oligonucleotides provides the means of sequence specific recognition via binding as a third strand to any homopurine • homopyrimidine sequence regardless of sequence complexity or length. Chemically active groups attached to these oligonucleotides may then provide the required catalytic activity. Since both functions, the recognition and the catalytic function, are separate entities in these artificial enzymes it should be possible in principle to direct a desired chemical reaction with high specificity to a particular sequence within a double stranded DNA molecule - enlarging the pool of naturally occurring enzymes by adding sequence specificity and range of catalytic activity. Moser & Dervan (1987) first showed the viability of this technique by tethering an Fe-EDTA group to a homopyrimidine oligonucleotide. Once this oligonucleotide has bound to the target sequence the Fe-EDTA is induced to produce shortlived diffusible OH radicals by treating the complex with either DTT or H_2O_2. These shortlived OH radicals then cleave the DNA in the immediate vicinity of the Fe-EDTA group. In effect the Fe-EDTA-homopyrimidine oligonucleotide acts as an

artificial restriction endonuclease (**Figure 1.5**). Using this technique it was shown that the cleavage of large genomic DNA molecules (48 kbase λ-phage genome or the 340 kbase yeast chromosome III) can be directed to a unique sequence. (Strobel et al, (1988), Strobel & Dervan (1990)).

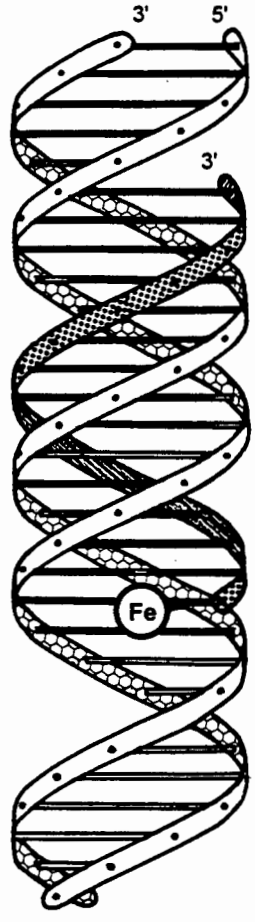


Figure 1.5: Schematic model for an oligonucleotide triple helix based artificial DNA restriction enzyme. (adapted from Moser & Dervan (1987))

Oligonucleotide Based Chemical Reagents

Based on the third strand binding capacity any chemically reactive group attached to an oligonucleotide can be used to target the desirable chemical reaction to a complementary site of double stranded DNA. Examples of oligonucleotide linked DNA cleavage

reagents mentioned in the recent literature, besides the Fe-EDTA group above, are those of oligonucleotides linked to Cu-1,10-phenanthroline (Francois et al (1989a,b), to ellipticine (Perrouault et al (1990), and even to a mutant MNase enzyme (Pei et al (1990). Guanines in the vicinity of triple stranded target sites can be sequence specifically alkylated and cleaved with piperidine by oligonucleotides tethered to N-Bromoacetaldehyde (Povsic & Dervan (1990), Povsic et al (1992)). Psoralen linked to oligonucleotides can cause sequence specific cross-links in double stranded target DNA (Takasugi et al (1991), Giovannangeli et al (1992)). It is also possible to crosslink the third strand oligonucleotide to the homopurine • homopyrimidine target sequence, if the third strand oligonucleotide contains the modified base N_4,N_4-ethano-5-methyldeoxycytidine (Saw et al (1991). The attachment of the dipyridophenazine complex of Ruthenium(II) to a triple strand forming oligonucleotide can provide a sequence specific molecular "light switch", since the dipyridophenazine complex of Ruthenium(II) fluoresces only when intercalated into the DNA (Jenkins & Barton (1992)).

Another field of applications of triple stranded oligonucleotides includes the use of biotinylated oligonucleotides to affinity purify the double stranded target DNA from complex DNA mixtures with the help of streptavidin coated magnetic beads (Ito et al (1992)), and the nonenzymatic ligation of DNA. Binding of two oligonucleotides to adjacent homopurine • homopyrimidine regions orientates the 3' hydroxyl and 5' phosphate groups of the two oligonucleotides in the appropriate orientation sufficiently close together to allow nonenzymatic ligation to occur.(Luebke & Dervan (1989). Alternatively a triple stranded oligonucleotide binding to two terminal homopurine • homopyrimidine regions of different double stranded DNA molecules will also place the '3 hydroxyl and 5' phosphate groups of the two double strands in close proximity to allow nonenzymatic ligation to occur (Luebke & Dervan (1991, 1992).

Triple Helices Inhibit Protein - DNA Recognition

On binding to a homopurine • homopyrimidine target sequence oligonucleotides prevent the binding of many DNA binding proteins to that same sequence (Maher et al (1989)). This seems to be entirely an effect of competition between the oligonucleotide and the protein for the same binding site, since a third strand bound adjacent to a protein binding site does not seem to inhibit the binding of a protein (Huang et al (1992)). Oligonucleotide directed triple helix formation does not only compete with the binding of a specific restriction endonuclease or restriction modification enzyme (e.g. methylases) to a target sequence (Hanvey et al (1989), Francois et al (1989)), but the binding of an oligonucleotide to a promoter region can also inhibit transcription of genes *in vitro* (Maher III et al. (1992), Cooney et al (1988), Ross et al (1992)) and possibly *in vivo* (Orson et al (1991). The inhibition of transcription initiation and the competition with specific binding of proteins to DNA *in vitro* makes oligonucleotide directed triple helix formation a candidate for regulatory functions *in vivo*. In that respect it is also very handy that homopurine • homopyrimidine sequences are significantly over-represented in eukaryotes and in eukaryotic viruses and are frequently found in regions upstream of coding sequences (Behe (1987), Beasty & Behe (1988)). Unfortunately there has been no convincing experimental evidence linking triple helix formation and gene control *in vivo* as yet.

Triple helix formation of oligonucleotides which overlap a restriction enzyme recognition sequence can be used to protect these sites from the action of the restriction modification enzyme (Hanvey et al (1989), Francois et al (1989)). Removal of the third strand oligonucleotide will then cause the restriction sites originally protected to become accessible to enzymatic cleavage, while those sites initially not protected by the

oligonucleotide are then protected from enzymatic cleavage due to the chemical modification introduced by the restriction modification enzyme in the first step. Following this approach it is possible to reduce the number of sites cleaved by a particular restriction enzyme in large DNA molecules such that in the best case a unique cut occurs. This approach has been used to cleave the yeast genom and human chromosome 4 at a single site (Strobel & Dervan (1991), Strobel et al (1991)).

Applications of Oligonucleotide based Triple Strand Technology:

In recent years triple helix formation has been considered to be a useful tool for medical and pharmaceutical as well as for biotechnological applications both *in vivo* and *in vitro*. One of the proposed applications is anti-gene therapy. Anti-gene therapy has the advantage over anti-sense therapy because it allows gene activity to be controlled at the level of gene transcription instead of at the level of translation. Since most genes exist as single copy genes or in only very few copies in a cell, considerably fewer active antigene agents (oligonucleotides) inside a cell will cause an inhibition of transcription in comparison to the antisense strategy. (Maher III (1992), Riordan & Martin (1991), Charles (1991), Helene & Toulme (1990)). By this approach it should be possible to inhibit virus replication, to prevent the growth of carcinomas, and to modulate gene activity of any desired gene (for example hormone genes). Besides these pharmaceutical applications oligonucleotide directed triple helix formation can also be used in biotechnological applications by either directing desirable chemical reactions to a particular sequence, or by acting as artificial restriction endonucleases, by artificial ligases, or in the isolation and manipulation of DNA.(Helene et al (1989), Dervan(1992)).

A complete understanding of the structural and energetic aspects of third strand binding of oligonucleotide is a precondition for the successful application of triple helix technology in medicine (antigene therapy) or biotechnology (artificial enzymes). The lack of sufficient understanding has spawned a renewed interest in studying triple helix formation.

Summary

1.) Oligonucleotides binding as third strands to homopurine • homopyrimidine target sequences can act as sequence specific recognition elements for artificial DNA modifying enzymes

2.) Binding as third strands to homopurine • homopyrimidine target sequences oligonucleotides inhibit the binding of DNA modifying enzymes to this sequence

3.) Triple helix forming oligonucleotides provide the means for control of gene expression - the antigene strategy.

1.5 Properties of Oligonucleotide Triple Helices

The Structure of Homopyrimidine Oligonucleotide Triple Helices

The renewed interest in studying triple helix formation centers on oligonucleotides, because oligonucleotides unlike synthetic homo-polynucleotides can consist of complex sequences. Polynucleotides of necessity will consist of simple repeat units (monomers, dimers, trimers) only. Because of the inherent symmetry in the simple repeat sequence of polynucleotides, the orientation of the third strand with respect to the Watson & Crick homopurine or homopyrimidine strand of the double helical target site is relatively unimportant. For oligonucleotides however, because of the asymmetry of the complex sequences, the orientation of the third strand with respect to the purine strand within the double helix becomes important.

The initial proposal that pyrimidine third strands bind in a parallel orientation to the Watson & Crick homopurine strand was confirmed (Moser & Dervan (1987)). It was observed that a third strand oligonucleotide with an Fe-EDTA group attached to the 5' end cleaves the homopurine target strand at the 5' end. Subsequently the Hoogsteen base-pairing and the parallel orientation of the third strand with respect to the homopurine strand has been shown directly with the help of 1D & 2D NMR. (Rajagopal & Feigon (1989a,1989b), de los Santos et al (1989), Kan et al (1991)).

1D & 2D NMR also confirmed that the Hoogsteen cytosines are protonated at the N3 position. The requirement for protonation of cytosines was previously proposed because of the pH sensitivity of cytosine containing triple helices (Lipsett 1964, Hattori et al (1976)). Based on the fiber diffraction data it was assumed that triple helix formation would induce a structural transition to an A-DNA structure in the target double helix.

Consequently it was expected that the oligonucleotide triplexes have *C3'-endo* sugar puckers. The NMR data of oligonucleotide triple helices however do not confirm this. The early NMR experiments of oligonucleotide triple helices seem to indicate that the purine strand and pyrimidine strands have different sugar puckers. The MINSY and NOISY signals of the homopurine strand were interpreted to indicate *C2'-endo* (S-type) sugar puckers (characteristic for B-DNA). The cross-peaks assigned to the homopyrimidines were considered to indicate *C3'-endo* (N-type) sugar pucker (Feigon & Rajagopal (1989a, 1989b) de los Santos et al (1989)). An improved NMR analysis led to the proposal of S-type sugar puckers in the Watson and Crick strands of an intramolecular triple helix. In the Hoogsteen strand the NMR results indicate two types of sugar puckers. The third strand cytosines preferentially form *C3'-endo* (N-Type) sugar puckers and third strand thymines preferentially form *C2'-endo* (S-Type) sugar puckers (Macaya et al (1992a, 1992b)).

Ligand interactions support evidence of altered B-DNA like conformation

Support for the observation that oligonucleotide triple helices have a more B-DNA like conformation comes from the observation that the minor groove binding ligands netropsin, duocarmycin A and dystamycin are capable of binding without displacement of the third strand (Umemoto et al (1990), Lin & Patel (1992), Durand et al (1992)). All three minor groove binding ligands are known to bind to B-DNA only (Zimmer & Wähnert (1986) and references therein).

Interactions with the Intercalating Ligand Ethidium Bromide

Ethidium bromide was also shown to interact with oligonucleotide triple helices (Mergny et al (1991)). The intercalation of ethidium bromide into the oligonucleotide triplexes

causes a destabilization of the triple helix and occurs with a reduced binding density compared to the double helix.

A subsequent modeling study by Sun et al (1991) indicates that the reduced binding density may be due to electrostatic repulsion between the positively charged intercalator and the positive charge of the protonated cytosines. The modeling study demonstrates that the destabilisation of the triple helix may be caused by the higher energy penalty associated with unwinding a three stranded helix as compared with that of a two stranded helix. Any DNA conformation inhibiting this unwinding process will resist the intercalation of ethidium bromide. Especially intramolecular triple helices will resist unwinding because of the orientation of the three strands.

It is conceivable that there are intercalating drugs which bind preferentially to triple helices but not to double helices. A derivative of the intercalator benzo[*e*]pyridoindole has in fact been shown to intercalate with triple helices, but not with double helices (Mergny et al (1992)).

Thermodynamics of Oligonucleotide Triple Helix Denaturation

In order to understand the thermodynamics of triple helix formation it is desirable to gain insight into the forces which are responsible for the structural stability. Unfortunately there are significantly less thermodynamic data available for oligonucleotide triple helix denaturation than for double helix denaturation. Because of the scarcity of data it is difficult to assess the somewhat contradictory results found in the literature. All currently published calorimetric data of triple strand formation are summarized in **Table III**. Plum et al (1990) determined the thermodynamic parameters ΔH, ΔS and ΔG of an oligonucleotide triple helix formed from three independent oligonucleotide single strands

by both calorimetry and UV-melting experiments. Surprisingly they observed that the model-independent calorimetric ΔH_{cal} for the triple helix to duplex transition is considerably smaller than the corresponding model-dependent ΔH_{vH} obtained by UV-melting and by shape analysis of calorimetric curves. The difference in ΔH_{vH} and ΔH_{cal} indicates that this transition is not a two state process. Only the model-independent calorimetric ΔH values will give an accurate measure of the enthalpy of melting. Support for this notion comes from the observations by Ohms & Ackermann (1990) that the melting of intramolecular ribo-oligonucleotide triplexes A_xU_y (y>x) occurs preferentially in a multistate manner.

Manzini et al (1990) and Xodo et al (1990) have investigated the thermodynamics of the triple helix to hairpin plus single strand and the triple helix to homopurine and palindromic homopyrimidine single strand transition by shape analysis of calorimetric and UV-melting curves. According to their results ΔH_{vH} and ΔH_{cal} for the triple helix formation are of the same order of magnitude. A discrepancy between the model-dependent ΔH_{vH} and model-independent ΔH_{cal} similar to that reported by Plum et al (1990) for a intermolecular triple helix has been observed on melting of an intramolecular triple helix (Völker et al (1993)). All the ΔH_{cal} values reported by Plum et al (1990) and Völker et al. (1993) are consistently lower than the corresponding ΔH_{vH} and of similar magnitude as the ΔH_{cal} values published for the polymer triple helix poly r(U) • poly r(A) • poly r(U). (Krakauer & Sturtevant (1968), Neumann & Ackermann (1969), Ross & Scruggs (1965)). The values of Xodo et al (1990) and Manzini et al (1990) are nearly twice as large. At present there is no simple explanation for these discrepancies. So there is an urgent need for more calorimetric experiments.

There are ΔH_{vH} values reported in the literature which are consistent with the calorimetric results. Pilch et al (1990) report ΔH_{vH} values of between 2 kcal/M bt and

3.4 kcal/M bt derived from UV-melting curves of the $d(T)_{10} \bullet d(A)_{10} \bullet d(T)_{10}$ and the $d(C^{+}_{3}T_4C^{+}_{3})_{10} \bullet d(G_3A_4G_3)_{10} \bullet d(C_3T_4C_3)_{10}$ oligonucleotide triple helices respectively. It has to mentioned that this study was done at extreme ionic strength of 2 M NaCl and/or 50 mM $MgCl_2$.

Table III: Thermodynamic data of Oligonucleotide Triple Helix melting obtained by Calorimetry

Transition	ΔH/bt kcal/M	ΔS/bt cal/M K	Reference
T GAAGGAGGAGA T CTTCCTCCTCT T T + CTTCCTCCTCT	7.5 ± 0.5	----------	Manzini et al (1990)
T GA G GA G GA GA T CT C CT C CT CT T T + $M_{CT}M_CM_{CT}M_CM_{CT}M_{CT}$	5.1 ± 0.5	16.5 ± 1.6	Xodo et al (1991)
T CTCTTCTTTCT T T CTCTTCTTTCT T + GAGAAGAAAGA	6.2 ± 0.6	---------	Xodo et al (1990)
GCTAAAAAGAGAGAGATCG CGATTTTTCTCTCTCTAGC + TTTTTCTCTCTCTCT	2.0 ± 0.1	6.5 ± 0.5	Plum et al (1990)
C GAGAGAGAAA C T CTCTCTCTTT C T C T T CTCTCTCTTT	4.0 ± 0.2	12.7 ± 0.6	Völker et al (1993)

Isothermal Equilibrium Studies

Using the triple helix affinity cleavage method developed by Moser & Dervan (1987) Singleton & Dervan (1992a, 1992b) attempted to measure the equilibrium binding constant of a single strand binding to a homopurine • homopyrimidine target region. According to the study the binding constant depends only marginally on the length of the third strand oligonucleotide, but strongly on the pH of the solution. The strong influence of the pH on the equilibrium constant is expected, because the third strand cytosines have to be protonated in order to form a second Hoogsteen hydrogen bond. It is worth mentioning that the pH also influences the kinetics of triple strand formation predominantly by changing the dissociation rate constant. It is observed that at high pH the dissociation rate is much larger than at low pH. An increase in monovalent salt concentration also seemed to reduce the overall rate constant of triple strand formation by increasing the dissociation rate predominantly (Maher et al (1990)). Using nonequilibrium UV-melting experiments Rougee et al (1992) have determined that the kinetics of triple helix formation proceeds analogous to that of double helix formation and can be described by a pseudo-first order rate equation. However compared to the rate of formation of a double helix the rate of formation of a triple helix is two orders of magnitude lower.

Structure and Stability of Triple Strand Mismatches

A substantial research effort has been put into understanding the effect of mismatches on triple helix stability and structure. Triple helix mismatches can in principle be classified into two classes: **Class I** third strand mismatches - i.e. the third strand contains bases not complementary to the target sequence, while the double helix target is a conventional homopurine • homopyrimidine sequence and **Class II** target sequence mismatches - i.e. the homopurine • homopyrimidine double helix is interrupted by a single inverted base-pair. The 16 possible combinations are listed in **Table IV.**

Table IV: Triple Helix Mismatches

	CLASS I		CLASS II	
3rd Strand	match	mismatch	mismatch	mismatch
T	A•T	G•C	T•A	C•G
C	G•C	A•T	C•G	T•A
A	A•T	G•C	T•A	C•G
G	G•C	A•T	C•G	T•A

Using the affinity cleavage approach with Fe-EDTA linked third strand oligonucleotides Griffin & Dervan (1989) investigated all 16 possible combinations. Not surprisingly the extend of cleavage was maximal for target DNA's containing the canonical base-triplets T•A•T and C^{+}•G•C, while for all mismatch combinations the degree of cleavage was substantially reduced. Surprisingly the degree of cleavage for third strand containing a G opposite a double helical T•A base-pair was nearly identical to the cleavage observed for oligonucleotides with matching canonical base-triplets. The proposed A•A•T and

G•G•C base-triplets were never observed to occur. Weak cleavage of the target DNA by oligonucleotides forming T•G•C, T•C•G and C•C•G triplets was also observed. Subsequently 1D and 2D NMR and UV-melting studies (Macaya et al (1991)) of intramolecular triple helices containing X•G•C mismatches (X = A, G, T) showed that all three mismatched third strand bases are capable of forming at least one hydrogen-bond with the homopurine strand, and that the order of thermal stability is T < G < A. The increase in thermal stability of the A^{+}•G•C triplet is caused by a protonation of the adenine at the N1 position at pH 5.2. This allows the formation of two hydrogen bonds between the adenine and guanine. This novel A^{+}•G•C mismatch was presumably not detected by Griffin & Dervan (1989) because of the differences in pH of the two studies. UV-melting studies by Mergny et al (1991) of all 16 possible base-combinations in a triple helix formed by the association from three oligonucleotides show a different result. In this case the G•G•C triplex is the least obstructing mismatch, closely followed by the G•T•A mismatch. The least stable mismatches in this study were the C•A•T and C•T•A base-triplets. Observed differences in the stability of mismatched base-triplets reported by Griffin & Dervan (1989), Macaya et al (1991) and Mergny et al (1991) may be due to the vastly different environmental conditions and differences in nearest neighbor base-triplets. As expected mismatches are more easily tolerated at the termini of the added third strand than in the center of the sequence.

The effect of a single base bulge loop in the third strand is less pronounced in a pyrimidine oligonucleotide than in a purine oligonucleotide. This can be explained by assuming that it is energetically less costly to unstack and loop out a pyrimidine base than a purine base. (Mergny et al (1991)). To include a third strand abasic site into a triple helix has a surprisingly large destabilizing effect. The abasic site is accommodated better if it occurs opposite a pyrimidine • purine inversion in a homopurine • homopyrimidine run, than opposite a continuos homopurine • homopyrimidine sequence

(Horne & Dervan (1991)). Triple helices do also form if the Watson & Crick strand contains a mismatch in the homopurine • homopyrimidine recognition site (Sun et al (1991)).

The Molecular Structure of the G•T•A Triplet

The molecular structure of the G•T•A mismatch has received considerable attention, because of the unexpected tolerance of this mismatch. Griffin & Dervan (1989) proposed a model in which the guanine C6 amino proton forms a hydrogen-bond with the C4 keto oxygen of thymine. Molecular modelling showed that either of the two amino protons can potentially form this hydrogen-bond. In either case the backbone conformation is distorted to accommodate the bulky purine base. Experimentally it can be shown by two dimensional (Wang et al (1992), Radhakrishnan (1991, 1992b) and three dimensional NMR (Radhakrishnan (1992a)) that for the G•T•A mismatch sandwiched between T•A•T triplets it is one particular amino proton which is involved in the hydrogen bond. NMR also showed that while the Watson & Crick T•A base-pair is not substantially distorted by the third strand guanine, the guanine itself is tilted with respect to the plane of the Watson & Crick base-pair. This can potentially give rise to a second hydrogen-bond between the other amino-proton and the thymine carbonyl oxygen of an adjacent base-triplet. The tilt of the guanine is accompanied by a rotation away from the center of the helix, giving rise to a *C2-exo* (N-Type) sugar pucker. (Radhakrishnan (1992b)

Triple Helix Formation with Homopurine Oligonucleotides

The "triple helix binding code" defined by Letai et al (1988) (**Table I**) predicts the formation of triple helices with purine bases in the third strand. However the replacement of the matching pyrimidine base with its complementary purine base destabilizes the

triple helix strongly (Griffin & Dervan (1989). Despite this destabilizing effect Cooney et al (1988) have observed that the transcription of the *c-myc* gene is inhibited in the presence of an oligonucleotide containing predominantly purine bases. It was postulated that this is due to the formation of a purine triple strand at a purine • pyrimidine rich sequence upstream of the promoter. The sequence of this purine rich oligonucleotide is such that the triple strand can only form with the third strand orientated antiparallel to the homopurine region. Subsequently Beal & Dervan (1991) confirmed that a purine rich oligonucleotide with a Fe-EDTA group attached can bind as third strands to a homopurine • homopyrimidine region. The observed cleavage pattern of the Fe-EDTA generated free hydroxyl radicals is consistent with an antiparallel orientation of this strand with respect to the Watson & Crick homopurine strand. Support for the antiparallel arrangement of the two purine strands comes from the investigation of intramolecular triple helices. In intramolecular triple helices the order of the purine and pyrimidine strands and therefor also the orientation of these strands is predetermined by the sequence. An intramolecular G•G•C Triple helix consisting of two G_8 and one C_8 strands is formed only if the two G_8 sections in the folded state are in an antiparallel orientation with respect to each other (Chen (1991)). The purine • purine • pyrimidine triple helix is reported to be considerably more stable then a corresponding pyrimidine • purine • pyrimidine triple helix in the presence of Mg^{2+} and at neutral pH (Pilch et al (1991)). 2 D NMR of an intramolecular purine • purine • pyrimidine triple helix showed that the antiparallel third strand can accommodate adenines opposite A•T base-pairs as well as thymines.(Radhakrishnan et al (1991)) In this case the thymines are considered to bind via reverse Hoogsteen base-pairing to the Watson & Crick adenine.

Mixed Purine-Pyrimidine Sequence in the Third Strand

The antiparallel orientation of the third strand purines with respect to the Watson & Crick purine amply explains, why the presence of a single purine base in an otherwise all pyrimidine third strand disfavors the incorporation of this sequence into a triple helix. A purine base within a pyrimidine third strand will not only distort the backbone of this strand to accommodate the bulkier purine base, but will also be prevented from forming the most favorable hydrogen bonds. If the homopurine • homopyrimidine target DNA is however flanked by a homopyrimidine • homopurine target sequence, then it is possible for an mixed purine-pyrimidine oligonucleotide to bind as third strand, provided that the purine and pyrimidine bases can align with the appropriate sequences (Jayasena & Johnston (1992a, 1992b)). Recently Giovannangeli et al (1992) reported that a co-oligomer containing a homopurine block next to a homopyrimidine block can bind to a homopurine • homopyrimidine target region in a parallel orientation to the Watson & Crick homopurine strand. This mixed sequence third strand seems to bind better than the corresponding all pyrimidine third strand at most environmental conditions. The formation of a mixed purine/pyrimidine triplex implies the formation of stable Hoogsteen G•G•C base-triplets not observed before. But in the same series of experiments it turned out that a single guanine in the midst of pyrimidine third strand destabilizes the triple helix. Currently the knowledge of triple helices involving purines or mixed purine/pyrimidine sequences in the third strand is very fragmented and insufficient to allow the proposal of an unambiguous model.

The Role of Modified Bases in the Third Strand

In order to expand the range of applications of triple strand forming oligonucleotides it is important to find oligonucleotides that can accommodate pyrimidine • purine interruptions amongst the homopurine • homopyrimidine target sequence and still bind with high affinity as third strands. It is also important that these triple helix forming oligonucleotides bind to their target sequence with reasonable specificity at neutral pH and low ionic strength, if triple helix formation should be used in gene control *in vivo*. Hence the large effort currently undertaken to find base-substitutions and nonnatural bases that can expand the third strand binding code to neutral pH. 5-Methyl cytosine is a naturally occurring modified base that helps to stabilize the oligonucleotide triple helices around neutral pH (Lee et al (1984), Povsic & Dervan (1989)). It has been established that the methylation of the 5' position raises the apparent pka of the cytosines within the triple helix by about 0.5 pH units and the melting temperature by about 10°C (Plum et al (1990). It does not change the pH dependence of the melting temperature compared to the corresponding oligonucleotides with unmodified cytosines. The increase in pka_{app} and T_m have been attributed to the replacement of water in the major groove (Xodo et al. 1990). 5'-Bromouracil has also been found to increase the stability of a oligonucleotide triple helix (Povsic & Dervan (1989). The cytosine analog 2'-O-methylpseudoisocytidine, which already contains a proton at the N3 amino position allows the formation of a triple helix at neutral pH (Ono et al (1991)). The triple helix binding code can be extended to allow recognition of both possible pyrimidine • purine interruptions (T•A and C•G) of a homopurine • homopyrimidine region by the use of the nonnatural deoxyribonucleoside 4-(3-benzamidophenyl) imidazole (D_3), which is capable of recognizing both T•A and C•G base-pairs with similar selectivity. (Griffin et al (1992), Kiessling et al (1992)) For the purine • purine • pyrimidine triple helices the use of nonnatural bases increases the

range of target sequences recognized by the third strand. 7-deaza-2'-deoxyxanthosine is the purine analog of thymine with respect to the hydrogen-bonding donor and acceptor sites. Its use in a mixed sequence with guanine in a homopurine oligonucleotide allows the recognition of A•T base-pairs under conditions where a mixed guanine/thymine oligonucleotide would not bind.(Millligan et al (1993)). The baseanalog 2-deoxynebularine expands the purine third strand binding code (Stilze & Dervan (1993)). It allows to tolerate C•G interruptions in the homopurine • homopurine run. Besides modifications of the bases, modifications of the backbone have also been used. The backbone modifications are introduced in order to prevent the degradation of these oligonucleotides by intercellular nucleases. Surprisingly some backbone modifications also increase the thermal stability of the triple helical complex. Most notable amongst these is the peptide nucleic acid PNA-T (Egholm et al (1992)). The carbocyclic analogs of thymidine and 5-methyl-2'-deoxycytidine cause an increase in pka_{app} for the triplex formation similar to, but more pronounced than, the 5-methylcytosine (Froehler & Ricca (1992)).

Summary

1.) Homopyrimidine third strand oligonucleotides bind via Hoogsteen hydrogen-bonding to the major groove of a homopurine • homopyrimidine double helix. The third strand is orientated in parallel to the Watson & Crick homopurine strand, with the Hoogsteen cytosines protonated at the N3 position.

2) The sugar pucker of the Watson & Crick homopurine and homopyrimidine strand is *C2'-endo*, while the sugar pucker of the Hoogsteen pyrimidine strand depends on the sequence. The sugar pucker of the third strand thymines is *C2'-endo*, while that of third strand cytosines is *C3'-endo*.

3.) The minor groove binding ligands netropsin, duocarmycine and dystamycine remain bound to the minor groove position on triple helix formation. They bind with reduced density and destabilize the triple helical complex - "groove-cross-talk".

4.) Ethidium bromide intercalates between adjacent base-triplets with a reduced binding density compared to double helical DNA. The presence of a positively charged cytosines seems to prevent any intercalation in the vicinity of cytosines.

5.) The enthalpy of the triple helix to duplex transition determined by calorimetry (ΔH_{cal}) is about half the amount determined by curve fitting procedures from UV-melting curves (ΔH_{vH}).

6.) Guanine rich homopurine third strands are orientated antiparallel to the Watson & Crick homopurine strand. The Hoogsteen guanines are in the anti conformation.

7.) Mixed homopurine / hompyrimidine third strands have thus far only been found if the homopurine and homopyrimidines bases occur in discrete blocks, and the double stranded target sequence contain a homopurine • homopyrimidine region followed by a homopyrimidine • homopurine region.

8.) Triple helix mismatches can be classified into two groups. Either the third strand contains bases other than the conventional Hoogsteen bases, or the homopurine • homopyrimidine region is interrupted by an inverted base-pair. An example of the former is the A^{+}•G•C mismatch, while the G•T•A base-triplet is an example of the latter.

9.) Modified bases are introduced to increase the range of potential target sequences for triple helix formation

10.) Modified backbone structures have been introduced to prevent the degradation of the oligonucleotide by nucleases.

1.6 Intramolecular Triple Helices

Classification of Oligonucleotide Triplexes

Oligonucleotide based triple helices can be formed from one, two or three independent strands. The interaction of three different oligonucleotides of equal size in the stoichiometric ratio of 1:1:1 provides the most straight forward way to form a triple helix (**Class I**) (**Figure 1.6 A**).(Pilch et al.(1990) Rajagopal & Feigon (1989a,1989b) Plum et al (1990)). Following this nomenclature a **Class II** triple helix is then formed from the interactions of two different oligonucleotide strands (**Figure 1.6 B**). There are two different ways to form a triple helix from two independent strands. A single stranded oligonucleotide, usually a oligopyrimidine, can combine with a homopurine • homopyrimidine hairpin to form a triple helix (Manzini et al (1990), Roberts & Crothers (1992)). Alternatively two oligonucleotides form a triple helix through the addition of a homopurine strand to a palindromic homopyrimidine sequence linked by a sequence which ends up as a loop region (Xodo et al (1990). In this case neither of the strands is able to form an ordered secondary structure by itself. Only when the two strands are mixed at stoichiometric ratio of 1:1 will they combine to form a triple helix.
Class II triple helices can also originate from a disproportionation reaction of the hairpin molecule under special circumstances. The disproportionation reaction is favored by high oligonucleotide concentration, high ionic strength and/or low pH (Mooren et al (1990)). Another **Class II** type triple helix is formed by the interaction of a homopurine single strand with a circular homopyrimidine oligonucleotide (Kool (1991), Prakash & Kool (1991,1992)). Triple helices formed from a single oligonucleotide strand are grouped into **Class III** (intramolecular triple helix) (**Figure 1.6 C**).

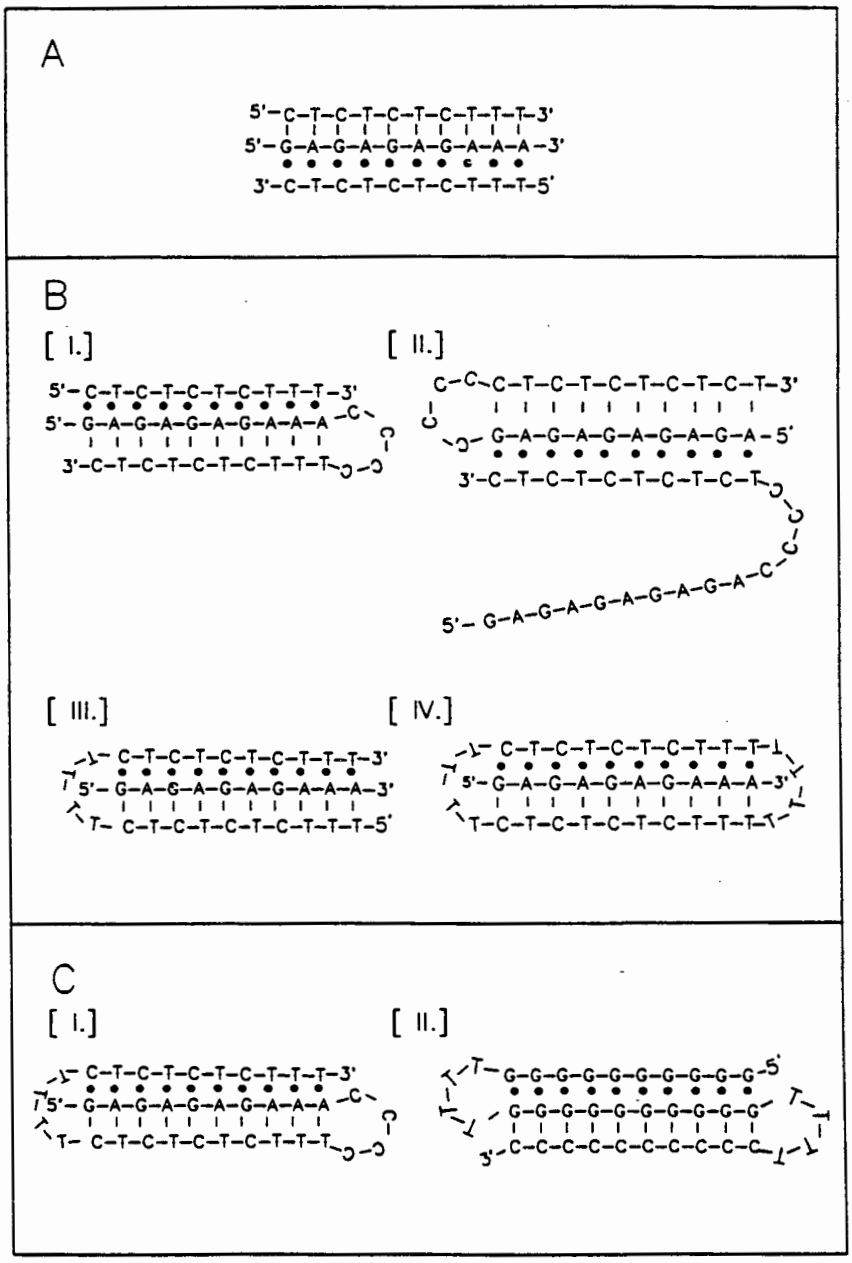


Figure 1.6: Classification of oligonucleotide triplexes based on the number of oligonucleotide molecules:

A) Triplexes from three independent oligonucleotide strands (**Class I**),

B) Triplexes from two independent oligonucleotide strands (**Class II**):

- [I.] hairpin plus homopyrimidine third strand
- [II.] disproportionation of a hairpin
- [III.] homopyrimidine palindrome plus homopurine strand
- [IV.] circular homopyrimidine palindrome plus homopurine strand

C) Triplex from one oligonucleotide strand (**Class III**):

- [I.] pyrimidine third strand
- [II.] purine third strand

Prerequisite for the intramolecular triple helix formation is a certain alignment of the building blocks. The sequence has to start or to end with one type of oligonucleotide, purine strand for purine • pyrimidine • pyrimidine sequences and pyrimidine sequence for purine • purine • pyrimidine sequences. Sequences which end up in the loop-region of the helical arrangement must consist of at least four nucleotides (**Figure 1.7**) (Sklenar & Feigon (1990), Häner & Dervan (1990), Chen (1991)).

Advantages of Intramolecular Triple Helices

It can be shown that **Class III** triple helices exhibit several technical and experimental advantages over **Class II** and **Class I** triple helices:

In an intramolecular triple helix the orientation of the third strand with respect to the Watson & Crick homopurine strand is predetermined by the sequence. Whether the third strand ends up in a parallel or antiparallel orientation with respect to the Watson & Crick homopurine strand depends on whether the third strand is a homopurine or a homopyrimidine sequence. In intermolecular triple helices (**Class I & Class II**) the orientation of the third strand can only be influenced by a judicious choice in sequence. The possibility of accommodating base mismatches, single stranded overhanging tails and hairpin loops (e.g. Pei et al (1991), Distefano et al (1991), Distefano & Dervan (1993)) in the third strand can make it difficult at times to unambiguously predetermine the orientation of the third strand in the case of **Class I** and **Class II** triple helices.

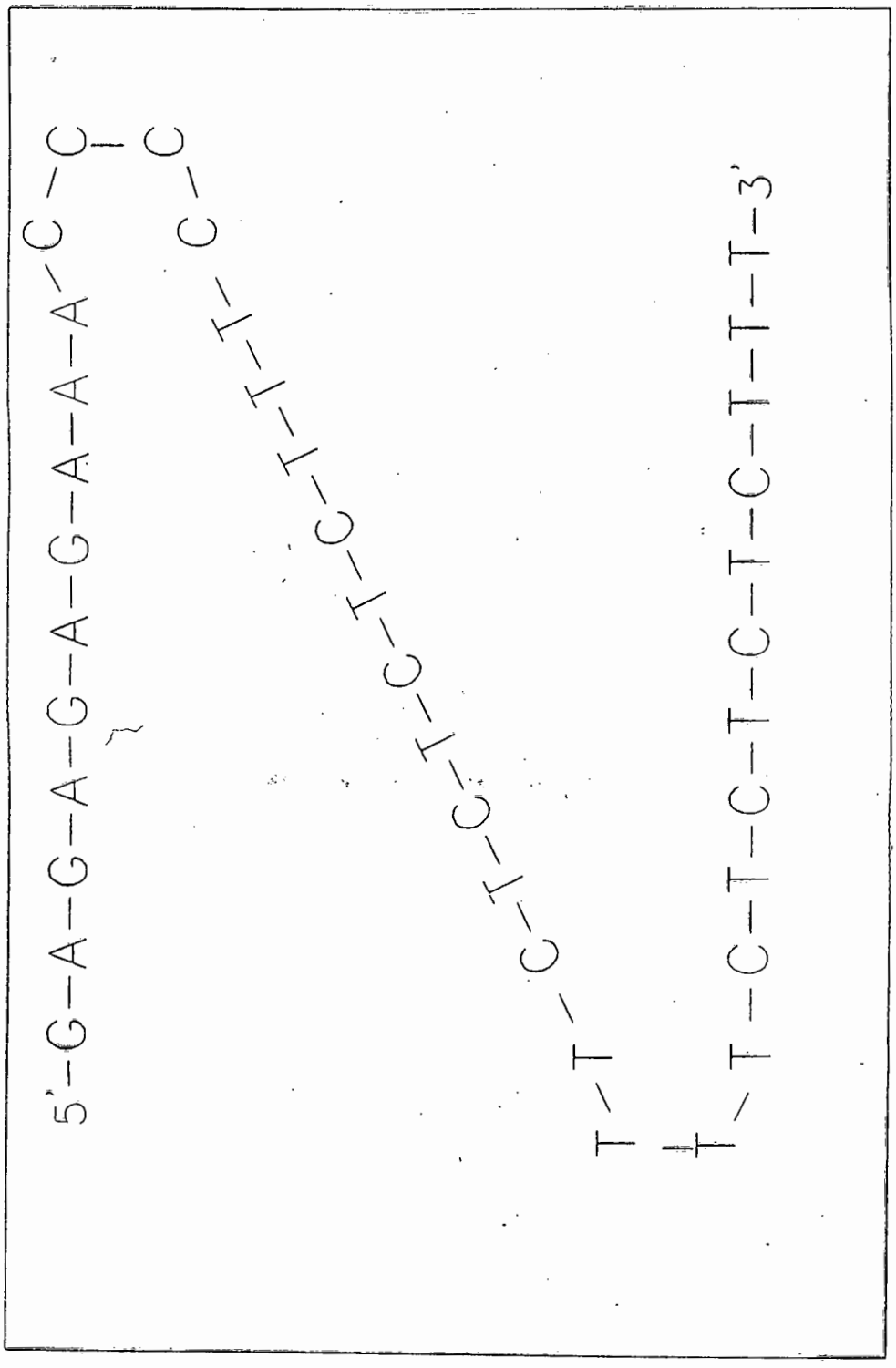


Figure 1.7: Consecutive folding steps of an intramolecular triple helix

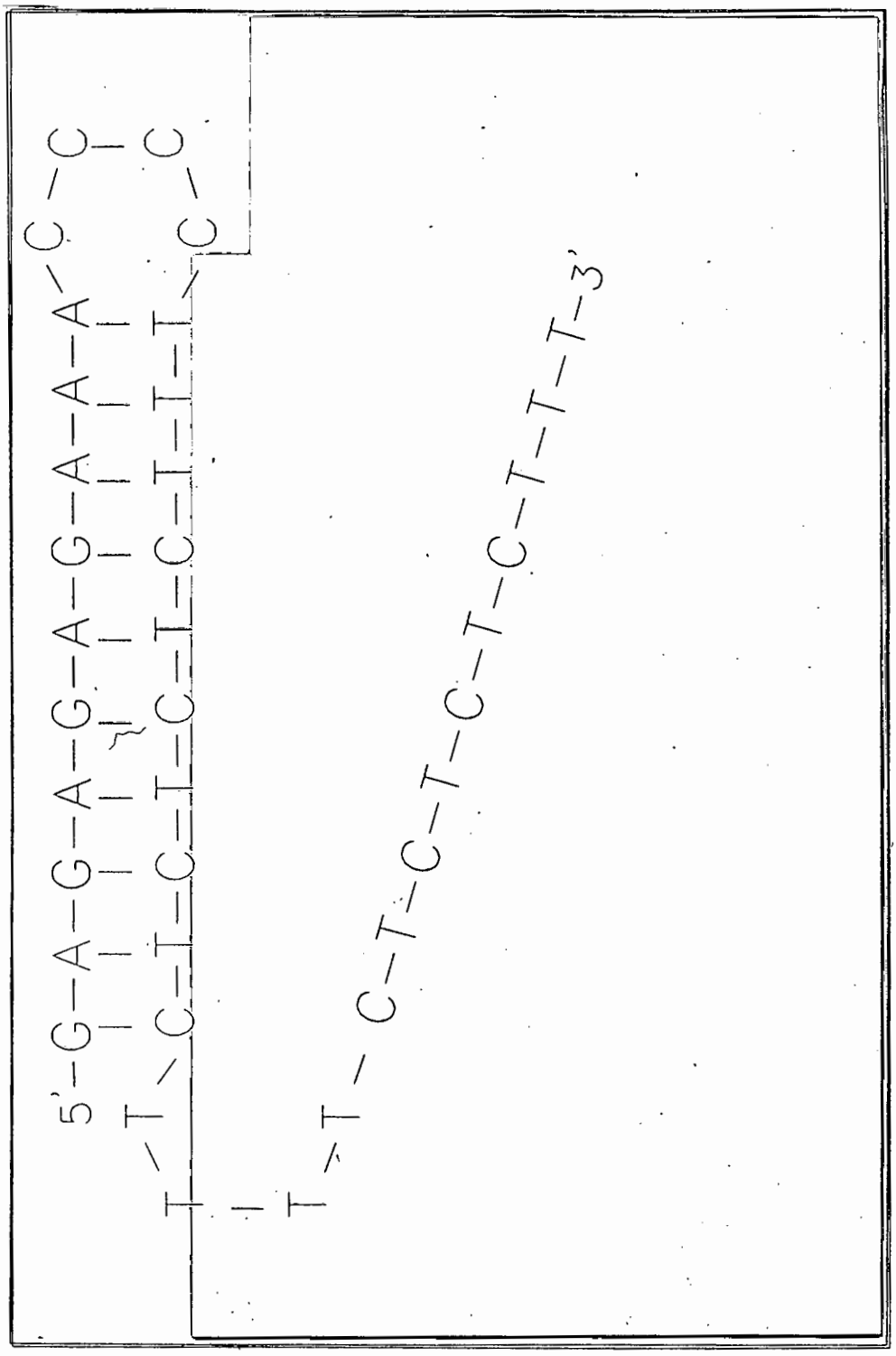


Figure 1.7: Consecutive folding steps of an intramolecular triple helix

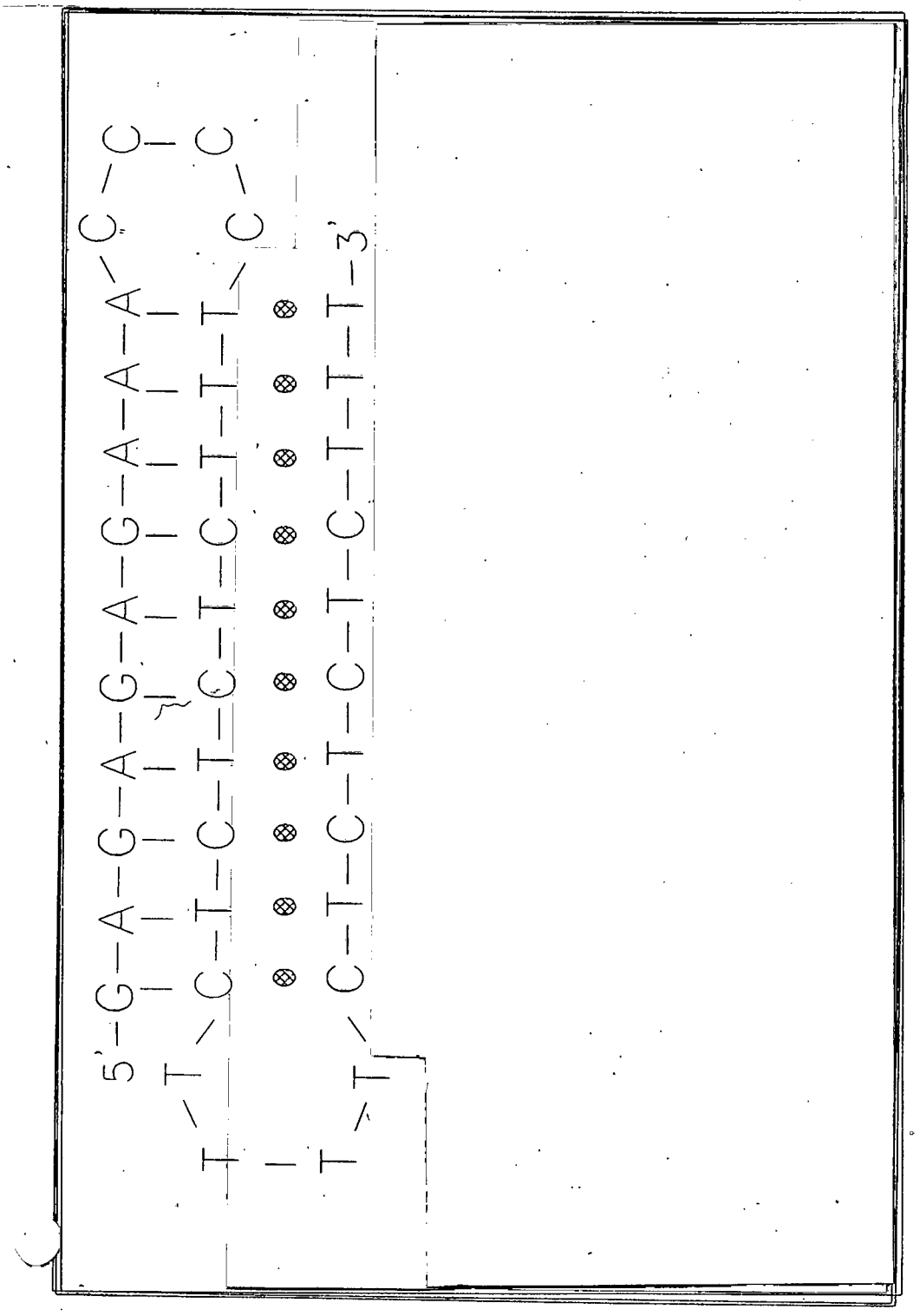


Figure 1.7: Consecutive folding steps of an intramolecular triple helix

Because of the intramolecular nature of the **Class III** triple helices the stoichiometry of the three strands with respect to each other is invariably given as 1:1:1. In the case of intermolecular triple helices the stoichiometry depends on the input ratio of the single strands, which is dependent on the exact determination of the concentration of the individual oligonucleotides. Inaccuracies in the determination of oligonucleotide concentration can potentially lead to erroneous conclusions.

The intramolecular structures are thermally more stable than corresponding intermolecular structures. The increased stability of the intramolecular triplexes is due to the apparent high local concentration of the third strand and the reduced conformational space.

The folding of an intramolecular triplex is a monomolecular process, and hence independent of the oligonucleotide concentration. This facilitates the comparisons of T_m values between different **Class III** triple helices, since it is not necessary to extrapolate the T_m to some standard reference concentration.

The formation of defective triple helical structures is grossly reduced in the case of intramolecular triplexes, since any partially complete folding is associated with a relatively large free energy penalty.

As will be shown intramolecular triple helices provide a superior experimental system for the study of temperature driven order-order and order-disorder equilibria at low and moderate oligonucleotide concentrations.

Problems with Intramolecular Triple Helices

Very high concentrations of oligonucleotides and the formation of base-pairs (unusual or conventional) in the bulge region can shift the equilibrium in favor of the intermolecular structure (**Figure 1.8B**). The hairpin loop and the triplex loop are structural elements of an intramolecular triple helix, which can also have considerable influence on the stability of the triplex. The loops will reduce the configurational entropy of the three strands - hence reduce the free energy of triple helix. But loops can also effect the enthalpy of the triple helix by the formation of stable or unstable loop structures. Stable loop structures involving unusual base-pairs are known in hairpin loops (e.g. Blommers et al (1989)) and can potentially also occur as triplex loops. Unstable loops can originate if the loop size is to small and thus exerting conformational stress on the neighboring base-triplets. The system can respond to this by drawing bases which normally are part of the triple helical section into the loop. (Macaya et al (1992)).

Thermodynamics of Intramolecular Triplex Formation

Since **Class III** triple helices provide the most straight forward experimental system, they can be viewed as a convenient system to determine the thermodynamics of triple helix formation and to monitor the effect of the environmental conditions on the thermal stability of these structures (Völker et al (1993), Durand et al (1992)). Currently our knowledge of the thermodynamics of oligonucleotide triple helix formation and the effect of environmental conditions on the stability of these triple helices is very limited. A large number of studies have attempted to evaluate the effect of mismatches (e.g. Griffin & Dervan (1990), Macaya et al (1991), Mergny et al (1991)) and nonnatural bases (e.g. Ono et al (1991), Griffin et al (1992), Kiessling et al (1992), Millligan et al (1993)) on the stability of triple helical structures.

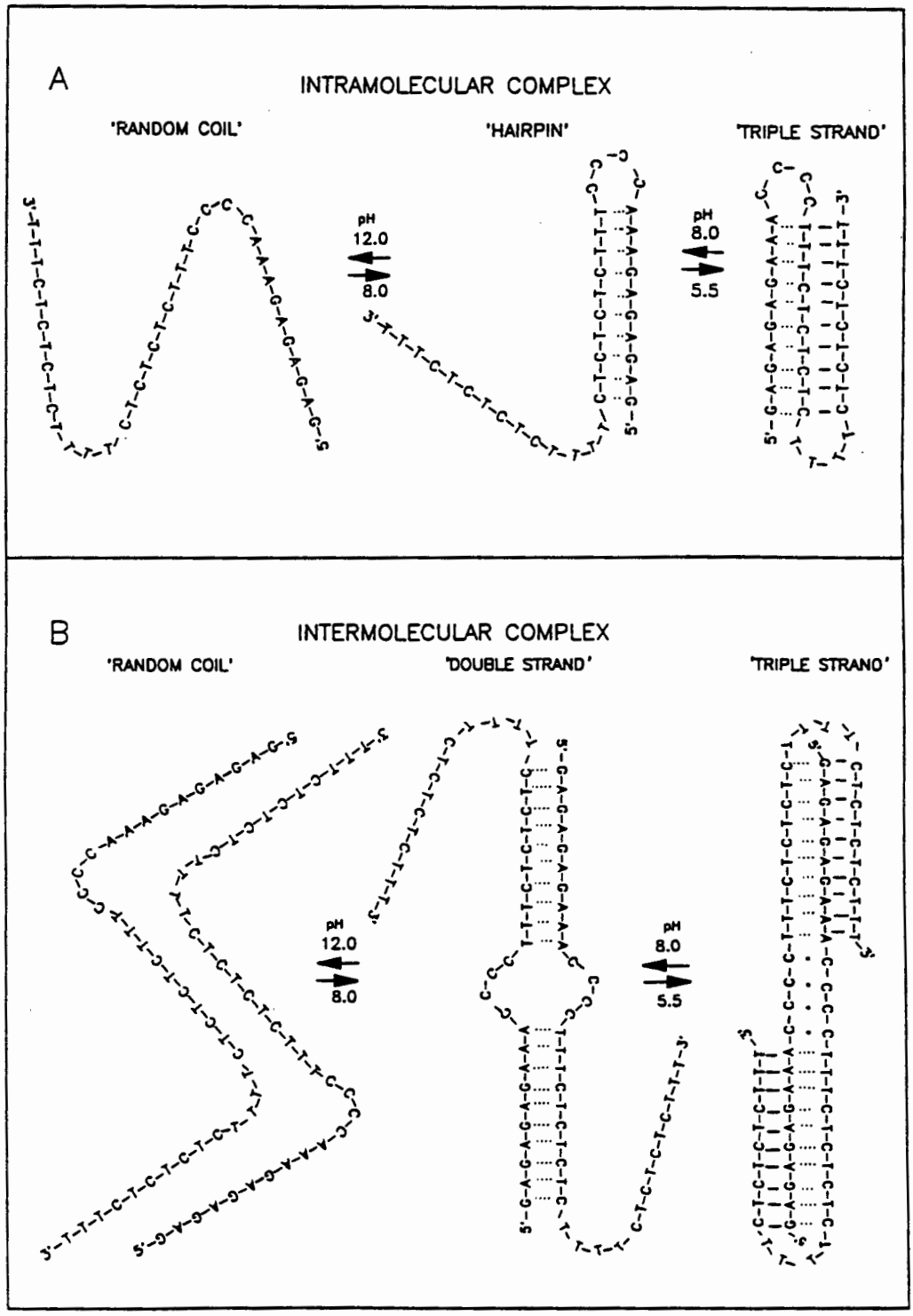


Figure 1.8: Intramolecular versus intermolecular triple helix
[A] Progressive folding of ITS-4G$_1$ into an intramolecular triple helix
[B] Progressive folding of ITS-4G$_1$ into an competing intermolecular triple helix

To asses the effect of mismatches and nonnatural bases on triple strand formation properly it is obviously required to study the effect of compositional variations of matched base-triplets on the stability of the structure. Very little is known about the effect of sequence variation in the third strand on the thermodynamics and stability of matching triple helices. Currently it is simply assumed that the effects of sequence variation on the stability of triple helices can be derived by extrapolation from the features of individual bases in homooligo- or homopolynucleotide triplexes. It will be shown that this approach is unjustified. This study was designed to gain a better quantitative and qualitative understanding of the impact of local and global composition on the stability of triple helical DNA. To this goal a systematic investigation of the effect of environmental conditions on the properties of sequence variants of intramolecular oligonucleotide triple helices was launched. The following oligonucleotides listed in **Table V** were chosen. All are capable of folding sequentially into a hairpin with dangling 3' tail and consecutively into an intramolecular triple helix given the appropriate environmental changes.

Table V : The oligonucleotide sequences

NAME	SEQUENCE
[A]	Global Variation
ITS-ATT	5'-AAAAAAAAAA CCCC TTTTTTTTTT *CCCC* TTTTTTTTTT-3'
ITS-2G$_3$	5'-GAAAGAAAAA CCCC TTTTTCTTTC TTTT CTTTCTTTTT-3'
ITS-3G$_2$	5'-GAAAGAAGAA CCCC TTCTTCTTTC TTTT CTTTCTTCTT-3'
ITS-4G$_1$[4]	5'-GAGAGAGAAA CCCC TTTCTCTCTC TTTT CTCTCTCTTT-3'
[B]	Local Variation
ITS-3G$_0$	5'-GAAAGGAAAA CCCC TTTTCCTTTC TTTT CTTTCCTTTT-3'
ITS-3G$_1$	5'-GAAAGAGAAA CCCC TTTCTCTTTC TTTT CTTTCTCTTT-3'
ITS-3G$_2$	5'-GAAAGAAGAA CCCC TTCTTCTTTC TTTT CTTTCTTCTT-3'

[4]ITS-4G$_1$ corresponds to JV-ITS in *J. Mol. Biol.*, (1993) **230**, 1278 - 1290

THEORETICAL BACKGROUND

2.1 General Considerations

In order to facilitate the interpretation and understanding of the experimental results and the discussion it is advisable to give a short overview of the theoretical background on which this thesis is based. The following chapter outlines briefly the theoretical considerations that were used as a basis for the design of the experiments and in the interpretation of the experimental results (and also indicates in which way this thesis differs from the customary interpretation of triple helix stability).

2.2 Forces Stabilizing Double Stranded DNA Helices

For the canonical Watson & Crick DNA helix the gross stability of a base-pair depends on three contributions: 1) the formation of interbase hydrogen-bonds; 2) stacking interactions between consecutive base-pairs; and 3) electrostatic interactions due to the negatively charged backbones. The energy of a single hydrogen-bond results from the difference in energy of the hydrogen-bond between the bases and the hydrogen-bonds between the bases and solvent water. Stacking interactions reflect a number of contributions from a) induced dipole interactions (London dispersion forces), b) van der Waals interactions and c) hydrophobic interactions. In general the stacking interactions are of the same order of magnitude as the hydrogen-bonding interactions. Both depend on the nature of the individual base-pair and its neighbors. This gives rise to sequence dependent contributions to the stability of DNA (Klump (1988), Breslauer et al. (1986)). The electrostatic backbone interactions on the other hand are independent of the nature of the base-pair or its neighbors, since each base is attached to an identically charged

section of the sugar-phosphate backbone. Normally electrostatic contributions are neglected when discussing the effects of sequence and composition on the stability of double stranded DNA. The interpretation of DNA stability in terms of hydrogen-bond formation and nearest neighbor stacking interactions has proved to be extremely successful for double stranded DNA, allowing the *ab initio* prediction of the T_m, enthalpy, entropy and free energy of any DNA sequences with reasonable accuracy from primary structure information (Klump (1988), Breslauer et al. (1986)). The large surface charge density resulting from the close association of the DNA backbones has to be reduced by the condensation of monovalent counterions onto the DNA surface to allow the double helix to form. This condensation partially neutralizes the negative charges of the backbone and thus reduces the repulsive forces between the backbones. The interaction of monovalent cations with the negatively charged phosphate backbone is commonly described in terms of either the counterion condensation theory developed by Manning (1978) or the Poison-Boltzmann Theory (Anderson & Record (1982)).

2.3 **Forces Stabilizing Triple Stranded Helices**

It is very likely that the same forces that stabilize a double helix, also contribute to the stabilization of a triple helix. It has therefor been assumed that the stability of a DNA triple helix can be derived from the summation of hydrogen-bonding and stacking interactions of the constituent base-triplets (Cheng & Pettit (1992)). As will be shown this approach is not valid. Since both the third strand cytosine and the third strand thymine form two hydrogen-bonds to their respective purine base (**Figure 1.1 & 1.2**), the enthalpy of hydrogen-bond formation will be approximately identical for the binding of a third base. Any differences in stability between different base-triplets must be either due to differences in stacking enthalpy or originate from local differences in electrostatic interactions. Most current studies of oligonucleotide triple helix formation discuss differences in stability of C•G•C and T•A•T triplets in terms of differences in stacking interactions. Contributions from electrostatic effects are mentioned, but not specified (e.g. Kiessling et al. (1992)). The importance of nearest neighbor stacking interactions in stabilizing double helical DNA clearly influences the interpretation of the contributions of sequence to triple helix stability.

2.4 **The Importance of Electrostatic Interactions**

Close inspection of homopyrimidine triple helix models should have highlighted the importance of electrostatic interactions for the stability of the complex. The electrostatic forces involved in triple helix formation can be grouped into a global and a local contribution. The binding of the third strand to the double helix occurs without a significant increase in the diameter of the helix (Arnott et al (1976), in other words the three strands are packed into the same volume prior occupied by the two strands of the Watson & Crick helix. Hence the surface charge density of the triple helix is considerably higher and consequently the electrostatic repulsion between the strands stronger than in the double helix. The repulsion between the strands is only due to the properties of the backbone and sequence independent. It can be considered as a global property. A local electrostatic component arises from the protonation of the N3 position in the third strand cytosines in order to create a second hydrogen-bond to the guanine N7 position (**Figure 1.2**). Hence the localization of the positive charges along a triple helix depends solely on the distribution of the third strand cytosines. The localized positive charge exerts an attractive force on the surrounding negatively charged phosphates in the backbones. It is impossible to calculate the attractive forces using Coulombs law, since it requires the dielectric constant for the hydrophobic core of the triple helix to be known. It can be safely stated however that the attractive electrostatic force due to cytosine protonation will modulate the repulsive electrostatic force originating from the close proximity of the negatively charged backbones. The overall electrostatic contributions to the stability of the triple helix result from the global (repulsive) and the local (attractive) electrostatic components. A comparison of the stacking interactions with the ionic interactions (e.g Chang (1981)) shows that the electrostatic forces will dominate the helix formation. Consequently the focus of this thesis will be directed to investigate these effects.

2.5 The Theory of Counterion Condensation

Counterion Condensation onto Polyelectrolytes

The counterion condensation theory (Manning (1978)) applied to triple helices with different cytosine content will allow an estimate of the importance of electrostatic forces for triple helix stability. Monovalent counterions associate with polyelectrolytes causing a partial neutralization of the polyelectrolyte charge. A fundamental requirement for Manning's theory is that the polyion is an infinitely long linear lattice molecule with point charges only, i.e the ratio of polyion diameter to polyion length must be close to zero. If the separation of the point charges by an average distance **b** (in Å) results in an overlap of the equipotential surfaces surrounding each ion, then it can be assumed that a constant potential energy surface will form around the polyion. Counterions condense onto this potential energy surface and are free to migrate along the constant potential but not perpendicular to it. Counterion condensation is thus fundamentally different from ion site binding. A large energy input is needed to remove a counterion from the equipotential energy surface to which it is condensed. Ion condensation will persist irrespective of the bulk concentration of the counterion in solution i.e. the counterion condensation causes nonideal behavior of electrolyte solutions at the limit of infinite dilution. The degree of ion condensation depends on the charge density of the polyion. Intuitively the higher the charge density, the stronger the electric field and the more counterions are attracted to the unit surface, or conversely an increase in polyion charge separation (lower charge density) will mean a decrease in ion condensation. Above a certain threshold distance separating the point charges this model is not valid any more. The polyion will then act as a linear array of individual monomer ions. Hence there must be a critical distance between charges that determines whether or not monovalent counterions will condense onto a polyion. Manning's counterion condensation theory

(Manning (1978)) relates the degree of counterion condensation to the axial charge separation, thus allowing a determination of the critical charge separation distance. Because of the small diameter and long persistence length of DNA (Record et al (1981)) and its constant helical repeat, nucleic acid molecules are ideally suited for the application of the Manning theory.

Manning (1978) has determined the critical charge separation parameter ξ for counterion condensation (also called the dimensionless charge density parameter) from the average axial charge density parameter **b** (in Å) by the equation:

$$\xi = \frac{q^2}{\epsilon * k * T * b} \qquad \textbf{(2.1a)}$$

where **q** is the protonic charge, **ε** the dielectric constant of the solvent, **k** is the Boltzman constant and **T** the absolute temperature (in K). Since in aqueous solution **ε*T** is approximately constant for all temperatures the equation **(2.1a)** simplifies to :

$$\xi = \frac{7.15}{b} \qquad \textbf{(2.1b)}$$

If the average charge density parameter **b** for a particular polyion is larger than 7.15, then no ion condensation will occur. If **b** is smaller than 7.15, then small counterions will condense onto the polyion until **b'** (the sum of the average axial charge density parameter of the polyion and the condensed counterions) is equal to 7.15. For single stranded, double stranded and triple stranded DNA the helical pitch, and therefor the separation of negatively charged phosphates, is always less than 7.15Å and hence small monovalent counterions will condense onto the DNA. The fraction of condensed monovalent counterion per polyion charge is given by:

$$\theta = 1 - (\xi)^{-1} \qquad (2.2)$$

As a consequence of the dependence of ion condensation on the average axial charge density parameter **b** any conformational transition by a polyion to a conformation with a different axial charge density will be accompanied by a corresponding change in the degree of ion condensation. Hence the thermal denaturation of conventional B-DNA with a axial charge density of **b** = 1.7Å to two single strands with axial charge density of **b** > 3.4Å will be associated with a release of counterions from the polynucleotide backbone. The amount of counterions released on denaturation is determined by the difference in the degree of counterion association by equation:

$$\eta = (\xi_c)^{-1} - (\xi_h)^{-1} \qquad (2.3)$$

where η is a parameter introduced by Manning. (ξ_h) is the dimensionless charge density parameters of the helix and (ξ_c) that of the coil. If the average axial charge density of one of the conformations is known, and η can be determined experimentally, it is possible to derive the unknown axial charge density of the other conformation. It needs to be stressed though that this quantity is a reflection of the average separation of charges, not that of atoms.

A conformational transition, such as the denaturation of a DNA double helix to two single strands is an equilibrium reaction and hence can be expressed by the equilibrium equation:

$$\text{Helix} \leftrightarrow 2\ \text{Coil} \qquad (2.4a)$$

but since this transition will always be associated with a change in the average charge distance and hence the degree of counterion condensation between the helix and the coil state the counterion dissociation needs to be accounted for in the equilibrium equation. Hence equation **(2.4b)** is a more appropriate equilibrium expression:

$$\mathbf{Helix} \leftrightarrow \mathbf{2\ Coil} + \mathbf{iNa^+} \qquad \textbf{(2.4b)}$$

where $\mathbf{iNa^+}$ is the number of sodium ions released on helix coil transition. The application of simple equilibrium thermodynamics and the substitution of the Manning parameter η for $[Na^+]^i$ followed by differentiation with respect to the logarithm of the counterion activity **Log** $[a_{Na+}]$ leads to equation **(2.5a)** (Manning (1978), Record et al. (1978)).

$$\frac{d(\log K_{obs})}{d(\log [a_{Na+}])} = -\frac{\eta}{2} \qquad \textbf{(2.5a)}$$

At the transition temperature (T_m), where the equilibrium constant K_{obs} is equal to one and hence log K_{obs} equal to zero equation **(2.5a)** can be converted to equation **(2.5b)**:

$$\frac{d(T_m)}{d(\log [a_{Na+}])} = \frac{2.3 * R * T_m^2}{\Delta H_{obs}} * \frac{\eta}{2} \qquad \textbf{(2.5b)}$$

where **R** is the universal gas constant and $\Delta \mathbf{H_{obs}}$ the observed calorimetric enthalpy for the conformational transition. Equation **(2.5b)** allows the determination of the amount of counterion released (η) on thermal denaturation by simply following the melting temperature as a function of the activity of the counterions in solution. It is experimentally more convenient to determine the counterion concentration rather than the counterion activity, and hence in many experimental studies only the $\mathbf{d(T_m)/d(\log[Na^+])}$ is reported. Substituting the change in counterion concentration $\mathbf{d(\log[Na^+])}$ for the change in counterion activity $\mathbf{d(\log[a_{Na+}])}$ in equation **(2.5b)** will introduce only a small and constant error of about 6% (Record et al (1978)). Because the error is constant all slopes ($\mathbf{d(T_m)/d(\log[Na^+])}$) are equally effected. The use of $\mathbf{d(T_m)/d(\log[Na^+])}$ values in conjunction with equation **(2.5b)** is therefor justified. The new equation then becomes:

$$\frac{d(T_m)}{d(\log[Na^+])} = \frac{2.3 * R * T_m^2}{\Delta H_{obs}} * \frac{\eta}{2} \qquad \textbf{(2.5c)}$$

The ΔH_{obs} in equation **(2.5b)** is the average model independent enthalpy of the denaturation of a base-pair, and since melting is a cooperative process it needs to be determined calorimetrically. For double stranded DNA any increase in T_m with an increase in ionic strength corresponds to an equivalent increase in ΔH_{obs} (Klump 1988), such that the term $(2.3*R*T_m^2/\Delta H_{obs})$ in equation **(2.5b)** is taken to be constant (Record et al. 1976)). This is only the case if the increase in free energy with ionic strength is due to an increase in enthalpy only and the entropy remains constant. If this is the case then $d(T_m)/d(\log[Na^+])$ becomes directly proportional to the amount of counterions released on melting. However it is necessary to confirm that any increase in T_m depends solely on a corresponding increase in ΔH_{obs}.

2.6 The Effect of Protonation of Bases in Polynucleotides

The binding of protons to bases in polynucleotides introduces a localized positive charge into an otherwise negatively charged helix. This results in the reduction of the average axial charge density of the polyion. Proton binding will therefor act to reduce the amount of counterion condensation. The major difference between the binding of protons and that of conventional counterions is that protons are sitebound.

Assuming that protonation simply causes a reduction in the overall charge density of the polyion Record et al (1976) have derived an equation that relates the degree of ion condensation to protonated polynucleotides and the degree of ion condensation to unprotonated polyelectrolyte.

$$\xi_x' \equiv \xi_x * (1 \pm k_x) \qquad (2.6)$$

where ξ_x is the dimensionless charge density parameter of the unprotonated polyion of conformation x, k_x the number of protons bound to the protonated polymer and ξ_x' the dimensionless charge density parameter of the protonated polyion. Applying equilibrium thermodynamics to the protonated polyelectrolyte leads to the equation:

$$\frac{d(T_m)}{d(\log[Na^+])} = \frac{2.3 * R * T_m^2}{\Delta H_{obs}} * \left(k + \frac{\eta}{2} \right) \qquad (2.7)$$

where **k** is the difference in the number of protons bound to the helix and the coil state:

$$k = k_h - k_c \qquad (2.8)$$

and all other variables are defined as in equation **(2.5b)**. The term **(k+η/2)** in equation **(2.7)** is equivalent to the term **(η/2)** in equation **(2.5b)** with **k** as a correction factor to account for the reduction in charge density caused by protonation. Consequently it is expected that a increase in the number of protons bound to a particular conformation results in a corresponding decrease in the amount of counterions bound to this conformation. Any deviations from the expected behavior must be a reflection of the difference in the properties between the sitebound protons and the condensed counterions. Intramolecular triple helices with different GC content should provide an excellent experimental system to investigate these differences.

2.7 Counterion Condensation onto Oligoelectrolytes

The polyelectrolyte theory as outlined above applies only to long polymers, where the ratio of polyion diameter to polyion length is close to zero. For oligonucleotides the ratio of diameter to length is much larger and can approach the value of one. The potential energy surface of oligonucleotides can be modeled by ellipsoids rather than a linear rod. Hence polyelectrolyte theory will not give an adequate description of condensation of counterions onto oligoions. Because oligonucleotides are short end effects have to be considered as they reduce the overall charge density of the oligoelectrolyte. Record and Lohman (1978) adapted the polyelectrolyte theory to oligoelectrolytes by introducing a reduced dimensionless charge density parameter ξ_{end} acting over a distance **r** from the oligonucleotide end. ξ_{end} can be chosen such that it can be expressed as a function **f** of the polymer dimensionless charge density parameter ξ i.e:

$$\theta_{end} = 1 - (\xi_{end})^{-1} = f(1 - (\xi)^{-1}) \qquad (2.9)$$

Elson et al (1970) studied the melting behavior of a series of hairpin oligonucleotides $(AT)_n$ of increasing length. Based on these results Record and Lohman estimated that the $d(T_m)/d(\log[Na^+])$ for the oligonucleotide hairpins varies linearly as a function of the inverse of oligonucleotide length $\mathbf{N^{-1}}$ (**N** = number of nucleotide residues per strand) for oligonucleotides of sufficient length. A subsequent Monte Carlo study by Olmsted et al (1989) determined that a minimum length of 24 nucleotides is required to satisfy this linear dependence of $d(T_m)/d(\log[Na^+])$ on $\mathbf{N^{-1}}$. For shorter oligonucleotides significant deviations from linear curve $d(T_m)/d(\log[Na^+])$ versus $\mathbf{N^{-1}}$ can be expected. Olmsted et al (1991) also note that while a linear decrease in $d(T_m)/d(\log[Na^+])$ with oligonucleotide length does occur in hairpin helices, Monte Carlo modeling of short helices made from two separate strands predicts a completely different behavior.

Since for hairpins a decrease in helix length corresponds to a proportional decrease in the degree of counterion condensation, it is possible to calculate the degree of ion condensation for a given oligonucleotide hairpin from the polynucleotide value and the end effect parameter. Record and Lohman (1978) derived equation **(2.10a)** by relating the $d(T_m)/d(\log[Na^+])$ of a hairpin helix with a four membered loop to that of a corresponding polynucleotide: (equation 6 in Record and Lohman (1978))

$$\frac{S_N}{S_\infty} = \frac{N - 2.02 * r * (1-f)}{N - 4} \qquad \textbf{(2.10a)}$$

where S_N is the $d(T_m)/d(\log[Na^+])$ value of an oligonucleotide of length **N** and S_∞ that of the corresponding polymer. The term **r * (1-f)** represents the correction term for the end effect acting over a distance **r** from the oligonucleotide end, and **f** is a function relating ξ_{end} to ξ. The term **r * (1-f)** can be determined by fitting equation **(2.10a)** to the S_N values of series of model oligonucleotide hairpins of increasing size **N**. Record and Lohman (1978) calculated the term **r * (1-f)** to be 5.64. It is not possible to further factorize the term **r * (1-f)** into independent parameters **r** and **f**. Introducing the value of 5.64 into equation **(2.10a)** simplifies the equation to:

$$\frac{S_N}{S_\infty} = \frac{N - 11.4}{N - 4} \qquad \textbf{(2.10b)}$$

where the term **N-4** in the denominator derives from the fact that the number of bases in the stem of the hairpin is reduced by 4 bases in the loop.

As was shown above equation **(2.5b)** allows to calculate the amount of counterions released on denaturation and can be used calculate the average charge density of polyelectrolyte conformations, presumed that of one of the conformations undergoing the

transition is known. Failing that, a comparison of $d(T_m)/d(\log[Na^+])$ values of different polyelectrolytes can still give indications of differences in charge density. This is possible even if little is known about the average charge density in either conformation of any of the polyelectrolytes involved in the transition, provided that one of the conformations remains the same in all cases. Because of pronounced end effects in oligoelectrolytes equation **(2.5b)** cannot be applied to oligonucleotides directly. Equations **(2.10a+b)** show that the difference in counterion condensation in oligoions is the sum of a length dependent term and the length independent end effects. Hence for oligonucleotides with the same length identical contributions from end effects can be expected. Consequently any differences of $d(T_m)/d(\log[Na^+])$ values for oligonucleotides of the same length must reflect differences in the average charge density of these oligonucleotides. It is assumed that this formal approach is applicable to intramolecular triple helices as well.

A comparison of the slope $d(T_m)/d(\log[Na^+])$ of oligonucleotide triple helices should therefor give an accurate reflection of electrostatic forces involved with triple helix formation. Since the global electrostatic contributions due to the negatively charged backbones can be expected to be the same in all oligonucleotide triple helices, any deviations in charge density must be due to the local electrostatic contributions originating from the protonated cytosines. The reduction in the overall charge density caused by the protonation of the cytosines should be accurately described by equation **(2.7)**. If this equation does not give an accurate description of the degree of counterion condensation to a particular triple helix, we have to conclude that the effect of sitebound protons cannot simply be described by the behavior of condensed counterions. Because of the sequence specific distribution of the protonated cytosines in a oligonucleotide triple helix any effect caused by sitebinding of the protons should show up more clearly than in polynucleotides with a regular repeating distribution of protonated bases.

2.8 The Application of Ligand Binding Theory to Protons in Triple Helices

In order for triple helix formation to occur in mixed sequences it is necessary that the Hoogsteen cytosines are protonated at the N3 position to be able to form a second hydrogen bond to the guanine N7 position. The binding of a proton to the cytosines is equivalent to the binding of a ligand to cytosines.

In general the binding of a ligand to nucleic acids influences the equilibrium of the helix - coil transition, if the ligand binds preferentially to one or the other of these two conformations. The equilibrium of the helix and coil states in the presence of a ligand can be defined as follows:

$$\text{Helix} \leftrightarrow 2\ \text{Coil} + \Delta n \tag{2.11}$$

where Δn is the difference in the number of ligand molecules bound to the helix state (n_h) and the coil state (n_c).

$$\Delta n = n_h - 2n_c \tag{2.12}$$

Applying equilibrium thermodynamics to this process and differentiating the equation with respect to the logarithm of the ligand concentration it can be shown (Record et al (1976,1978)) that the change in helix - coil equilibrium caused by a change in ligand concentration is proportional to the difference in the amount of ligand bound to the helix and the coil state Δn :

$$\frac{d(\log K_{obs})}{d(\log [n])} = -\Delta n \tag{2.13a}$$

At the T_m the equilibrium constant K_{obs} is equal to one, and the equation **(2.13)** changes to :

$$\frac{d(T_m)}{d(\log[n])} = \frac{2.303 * R * T_m^2}{\Delta H_{obs}} * \Delta n \qquad \textbf{(2.13b)}$$

where **log [n]** the logarithm of the ligand concentration **[n]**, **R** is the universal gas constant, T_m is the melting temperature, and ΔH_{obs} is the observed model independent enthalpy of the transition per base-pair. ΔH_{obs} must be determined calorimetrically.

If the ligand is a proton, then **log[n]** corresponds to the pH of the bulksolution and hence equation **(2.13b)** changes to:

$$\frac{d(T_m)}{d(pH)} = \frac{2.303 * R * T_m^2}{\Delta H_{obs}} * \Delta k \qquad \textbf{(2.14)}$$

where Δk is the difference in the number of protons bound to the two conformational states. The ΔH_{obs} value in equation **(2.14)** is the model independent calorimetric ΔH_{cal}. If the change in T_m associated with a change in solution pH is enthalpic in origin, then the term $(2.3*R*T_m^2/\Delta H_{obs})$ will be constant and the change in $d(T_m)/d(pH)$ is directly proportional to the difference in protonation between the helix and coil state Δk. Proton binding to the third strand cytosines is thus conceptionally similar to the binding of conventional ligands.

The binding of protons to third strand cytosines differs from the binding of conventional ligands to DNA in that the proton is required to serve as a hydrogen-bond to bring about the binding of the third strand to the double helix. Consequently the equilibrium for the interaction of protons with the double helix and triple helix respectively:

$$\textbf{(Cytosine)} + \textbf{Targetsite} + H^+ \leftrightarrow \textbf{Triplex} \qquad \textbf{(2.15a)}$$

can be expressed as the sum of two different equilibria:

$$\textbf{Cytosine} + H^+ \leftrightarrow \textbf{Cytosine-}H^+ \qquad \textbf{(2.15b)}$$

$$\textbf{(Cytosine-}H^+\textbf{)} + \textbf{Targetsite} \leftrightarrow \textbf{Triplex} \qquad \textbf{(2.15c)}$$

where **(Cytosine)** and **(Cytosine-H^+)** represents the third strand in its unprotonated and protonated form respectively, and **Cytosine** and **Cytosine-H^+** represent the unprotonated and protonated form of individual cytosine bases. The equilibrium constant for the reaction **(2.15b)** is given by the pka of cytosine, while the equilibrium constant for the reaction **(2.15c)** is that of the binding of the third strand. The amount of protons bound to the uncomplexed third strand is defined by the pka of free cytosine, while the amount of protonation of the third strand in the triple helix is defined by the equilibrium reaction **(2.15a)**. If the solution pH is more than 1 pH unit higher than the pka value of the free cytosines, then the uncomplexed third strand cytosines will not be protonated and the magnitude of the **Δk** value (or $d(T_m)/d(pH)$ value) will depend entirely on the number of protons bound to the third strand. As the solution pH approaches the pka value the difference in protonation between the complexed and uncomplexed third strand will become progressively smaller and hence the $d(T_m)/d(pH)$ value will also become progressively smaller.

A comparison of the $d(T_m)/d(pH)$ values of oligonucleotide triple helices with local and global differences in composition at pH values larger than the pka of free cytosine can be used to determine the effect of protonation on the stability of the triple helix conformation. A requirement for this is that the term **$(2.3*R*T_m^2/\Delta H_{obs})$** in equation **(2.14)** is numerically the same for all oligonucleotides. If this is not the case, then the

$\Delta\mathbf{k}$ value has to be evaluated for each oligonucleotide individually. The $d(T_m)/d(pH)$ value can be easily determined experimentally by following the thermal denaturation of the oligonucleotides at different solution pH by measuring the UV-absorption as a function of temperature.

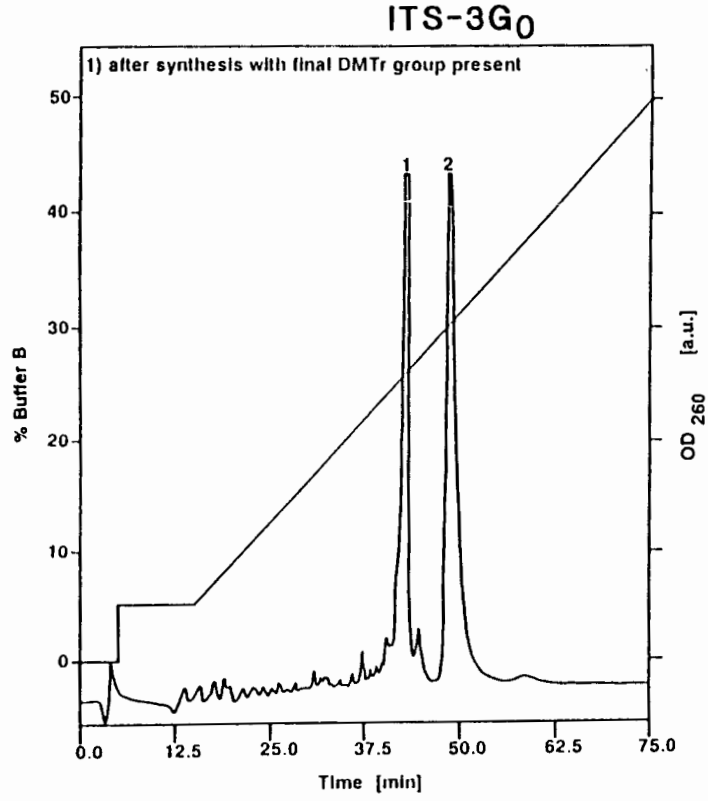
MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification

All oligonucleotides were synthesized by standard phosphoramidate procedure (Narang (1987) and references therein) on an Autogen^{TM} 6500 DNA Synthesizer (Milligen/Biosearch, Burlington Mass.). In order to drive each synthesis cycle to completion double coupling reactions were carried out during each cycle. Despite this each synthesis cycle results in a small percentage of incomplete synthesis reaction, and hence a small percentage of failure sequences. It is therefor necessary to purify the full length synthesis product from these shorter length failure sequences.

Unfortunately the conventional purification procedure of trityl selection on reversephase HPLC columns cannot be applied for the purification of these oligonucleotides because of secondary structure formation. Purification by anion exchange HPLC in denaturing conditions, which overcomes the effect of secondary structure formation, cannot be used, because of the length of the oligonucleotides. To overcome this problem a novel purification procedure was developed that combines anion exchange HPLC in denaturing conditions with nonspecific hydrophobic trityl selection. After deblocking the trityl containing synthesis product desolved in 50 mM LiCl, 10 mM NaOH (pH 12.0) (Buffer A) is loaded onto the Pharmacia Mono Q HR 5/5 ion exchange column, and then eluted by a linear gradient with buffer B (3M LiCl, 10 mM NaOH, pH 12). The tritylated oligonucleotide elutes well separated from any of the shorter length failure sequences and also from the detritylated full length oligonucleotide (**Figure 3.1A**). The tritylated oligonucleotide was then recovered by precipitation overnight at -20°C with 6 volumes of ethanol/aceton (1:3) and centrifugation at 15k rpm for 30 min.

A



B

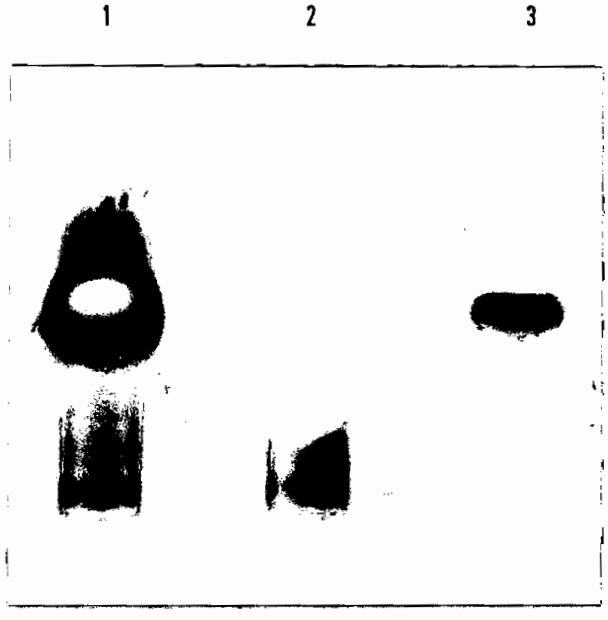


Figure 3.1: Purification of the oligonucleotide ITS-3G$_0$
[A] Trace of the HPLC run
[B] 20% denaturing PAGE gel.
lane 1 - unpurified, lane 2 - peak 1 in [A], lane 3 - peak 2 in [A]

The oligonucleotide pellet was resuspended in concentrated ammonia and then desicated to dryness. The final dimethoxytrityl group was then removed by adding 100 µl 80% acetic acid for 1 hour at room temperature, before the oligonucleotide was recovered by butanol precipitation and resuspended in a 50 mM Na^+ buffer (20 mM NaCacodylate/20 mM NaAcetate/10 mM NaCl, pH 8.0). All oligonucleotides except for ITS-$4G_1$ were purified by this method. The purity of each oligonucleotide was confirmed by denaturing gel electrophoresis on 20% polyacrylamide gels (**Figure 3.1B**) and the product was generally found to be sufficiently pure to allow ^{32}P endlabling for P_1-nuclease digestion without any further purification step.

P_1-Nuclease Digestion

Oligonucleotides were 5' endlabled with γ-^{32}P ATP and T4-kinase (Boehringer Mannheim) using the standard protocol modified to include a 3 hours incubation time and a chase with an extra 20 Units T4-Kinase after 1 1/2 hours incubation (Current Protocols in Molecular Biology). This proved necessary in order to achieve sufficiently high levels of incorporation of ^{32}P into the oligonucleotides, because the reaction of the T4-Kinase was inhibited by secondary structure formation. P_1-nuclease digestion was carried out in a 50 µl 100 mM Na^+ pH 8.0 or pH 5.5 reaction buffer containing approximately $5 * 10^5$ cpm of γ -^{32}P labeled oligonucleotide and approximately 0.3 (pH 8.0) or 0.003 (pH 5.5) units P_1-nuclease (Boehringer Mannheim) respectively. (depending on the labeling efficiency for the individual oligonucleotides.) At appropriate time intervals 5 µl samples were withdrawn from the reaction mixture and the reaction was stopped by the addition of 5 µl 0.5 M EDTA and 20 µl loading buffer (80% formamide, 20% glycerol) and by heating to 90°C. 10 µl of the heat denatured sample were loaded onto a 20% denaturing polyacrylamide gel (20:1 acrylamide : bisacrylamide, 8 M Urea) and electrophoresed overnight at 20 V/cm. The PAGE gel was then exposed

to an x-ray plate overnight and the autoradiograph was scanned with the help of a custom build densitometer.

UV-Spectroscopy and Thermal Denaturation Experiments

Thermal melting experiments (A_{260} vs temperature) were mainly carried out with the help of a Hewlett Packard HP 8450A Diode array spectrophotometer equipped with a 89100A Temperature Station and 89102A Temperature Probe. Occasionally melting experiments were performed with the help of a Pye-Unicam SP1800 UV-spectrophotometer equipped with a custom build preasurizable electrical temperature station. All melting experiments were routinely performed at a heating rate of 1°C/min, since a trial experiment with one oligonucleotide had shown that no discernible difference in the melting curves can be observed when heated at either 0.2 °C/min or 2 °C/min. All melting experiments were carried out using a 20 mM Na-Acetate/20 mM Na-Cacodylate/10 mM NaCl basic buffer (total Na^+ = 50 mM) at either pH 8.0, pH 6.75, pH 5.5 or pH 4.5. The Na^+ concentration was adjusted to the required concentration using a 5 M Na^+ (20 mM Na-Acetate/20 mM Na-Cacodylate/5 M NaCl) stock solution of the appropriate pH. The impact of pH on the melting behavior of the oligonucleotides was studied in a 150 mM Na^+ standard buffer (50 mM $NaPO_4$). The pH of the solution was adjusted by the addition of small volumes of dilute HCl (approximately 0.1 or 0.5 M) to the standard buffer. The pH was determined with the help of a Swiss-Lab pH Meter 300.

CD spectroscopy

CD-spectroscopy was carried out on a Jasco J-40A spectropolarimeter equipped with a thermostated quartz cuvette attached to an external waterbath. Spectra were recorded in

100 mM Na^+ buffer (20 mM Na-Cacodylate, 20 mM Na-Acetate) at either pH 8.0, pH 6.75 or pH 4.5 at the appropriate temperatures determined by the thermal melting experiments.

Ethidium Bromide Fluorescence Measurements

Fluorescence measurements were carried out in 10 mm fluorescence cuvetts using an Aminco SPF 500 corrected spectra spectrofluorimeter equipped with an external heating bath. The sample was heated at approximately 2°C/min and the temperature was monitored with the help of a temperature probe in a reference cuvette. The oligonucleotide concentration was 1 absorbance unit/ml in 100 mM Na^+ buffer, the ethidium bromide concentration was 81 nM to give an approximate ratio of base-pairs : ethidium of 1:1. The excitation wavelength chosen for all measurements was 230 nm; emission was monitored at 608 nm to record the fluorescence intensity as a function of temperature.

Scanning Calorimetry

Calorimetric experiments were performed using a Differential Scanning Calorimeter DASM 4 (Teshmash Export, Moscow) (Privalov & Potekhin (1986)). The oligonucleotide ITS-$4G_1$ in 100 mM Na^+ buffer at pH 8.0, pH 6.75 or pH 4.5 was heated from 20°C to 80°C at 1°C/min and the excess heat capacity ΔC_p was recorded as a function of temperature. The area under the peak represents the energy involved in unfolding of the ordered structure. Repeat determinations of the area using the same sample were reproducible to within < 5%.

RESULTS

4.1 Formation of the Intramolecular Triple Helix

Consecutive Folding of the Oligonucleotides into Intramolecular Triple Helices

In order to investigate the impact of global and local composition on the stability of a triple helical DNA the six sequences **(Table V)** were designed and synthesized. They all follow the same design pattern to facilitate the folding into isosterical intramolecular triple helices. Only one parameter is changed at a time, either the global composition is kept constant and the local composition varied, or the local composition is kept constant as much as possible and the global composition is systematically varied. To this aim each oligonucleotide contains a ten base homopurine run at the 5' end followed by a block of four cytosines. Next comes a ten base homopyrimidine region which is Watson & Crick complementary to the homopurine run, followed by four thymine bases. (or in the case of the oligonucleotide ITS-ATT by four cytosines). Finally a ten base homopyrimidine region Hoogsteen complementary to the homopurine run is situated a the 3' end. At room temperature, pH 8.0 and Na^+ concentrations not higher than 100 mM the consecutive folding of the purine section onto the following pyrimidine section is observed, leaving the 3' pyrimidines as a dangling end and locking the four spacer cytosines into a hairpin loop. Lowering the pH and/or raising the ionic strength causes the second pyrimidine section to fold back into the major groove of the stem loop structure. This will lock the other four spacer pyrimidines into the triple helix loop **(Figure 4.1)**. In general only thymines are included in the triplex loop sequence, with the exception of the oligonucleotide ITS-ATT. To exclude the possibility of variable loopsize the triplex loop of ITS-ATT consist of four cytosines. Provided that each oligonucleotide folds itself into the same intramolecular triple helical conformation, it will be possible to compare the

effects of local and global composition on triple helix stability. It should not be overlooked that there is a possibility to form intermolecular complexes between these single strands. However the concentration required to favor these structures over the intramolecular complexes in general exceeds the conditions commonly used for spectroscopic measurements.

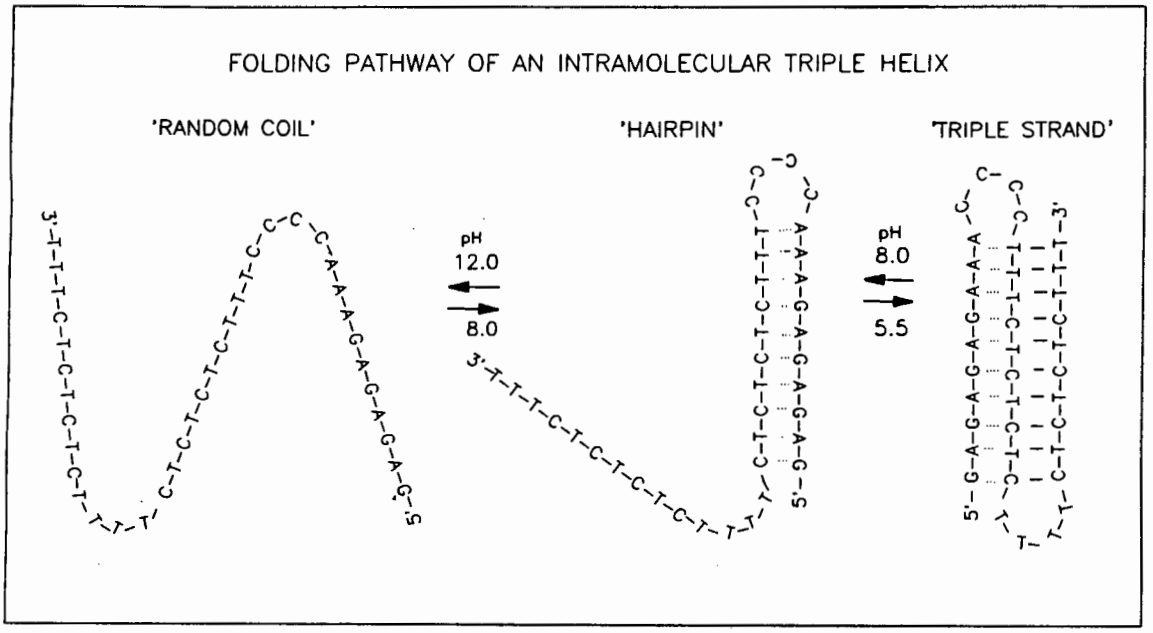


Figure 4.1: The proposed pH dependent folding pathway of an intramolecular triple helix

In the following the physical properties of the oligonucleotides is exemplified by the description of the typical behavior of a selected sequence in detail. In all other cases, where the global and local compositional variations lead to altered properties, this will be pointed out appropriately.

UV-Absorption of the Different Folding States

The folding of the oligonucleotides from a random coil first into a hairpin and then into an intramolecular triple helix is accompanied by a stepwise increase in the degree of order. Consequently each of the three conformations will have distinct extinction coefficients. It is thus possible to monitor the folding of the oligonucleotide by following the absorption as a function of the appropriate environmental conditions. Figure **4.2** shows how the formation of ordered secondary structures can be monitored by following the absorbance as a function of wavelength and pH (**Figure 4.2B**) or as a function of pH at a given wavelength (**Figure 4.2A**). At pH 12 and above the absorption at 260 nm is maximal, since the oligonucleotide is in the random coil state. Between pH 11.5 and pH 10.5 a hypochromic shift is observed. This corresponds to the cooperative formation of a hairpin with dangling 3' tail. Between pH 8.0 and pH 7.0 a second hypochromic shift occurs. Associated with this hypochromic shift is a slight red shift in the absorbance maximum from 264 nm at pH 8.0 to 268 nm at pH 7.0. This second hypochromic shift is due to the binding of the single stranded 3' overhang to the major groove of the hairpin helix, resulting in the formation of the intramolecular triple helix. All oligonucleotides listed in **Table V** can be shown to follow a corresponding two step folding process by recording the UV-absorbance as a function of the pH and/or ionic strength (in the case of ITS-ATT), but with a reduced redshift at low pH. It has been noted that protonation of cytosines is accompanied by a redshift in the absorption spectrum (Manzini et al (1990)).

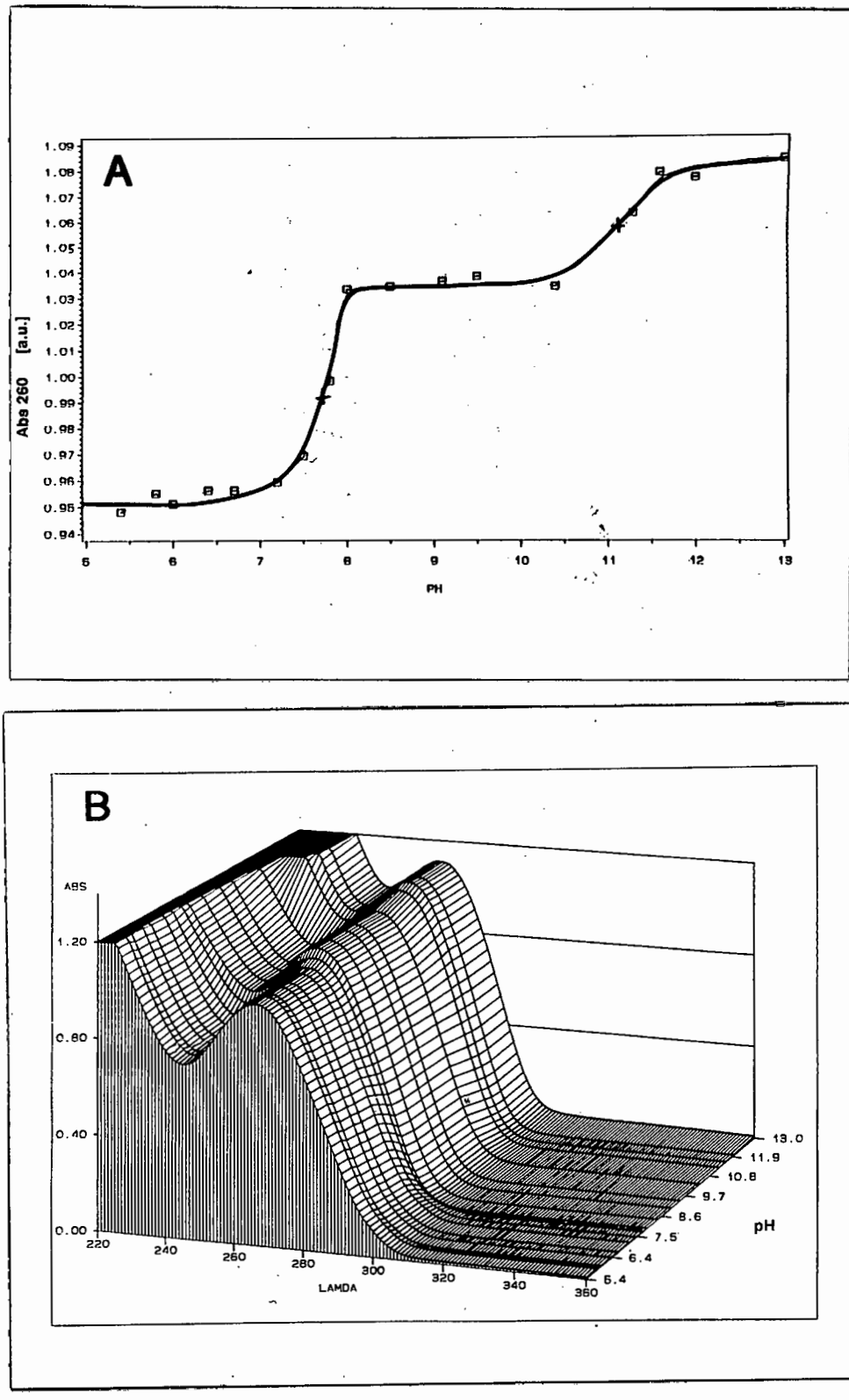
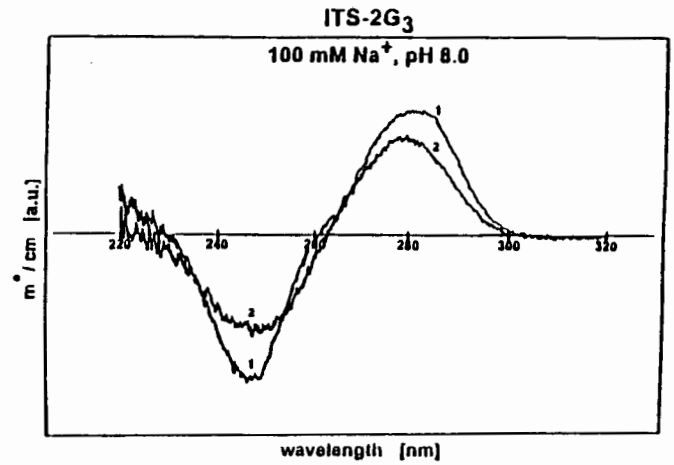


Figure 4.2: Change in UV spectrum of ITS-4G$_1$ as a function of pH.
(**A**) Monitored at 260 nm.
(**B**) Three dimensional representation of the UV-spectrum as a function of pH.

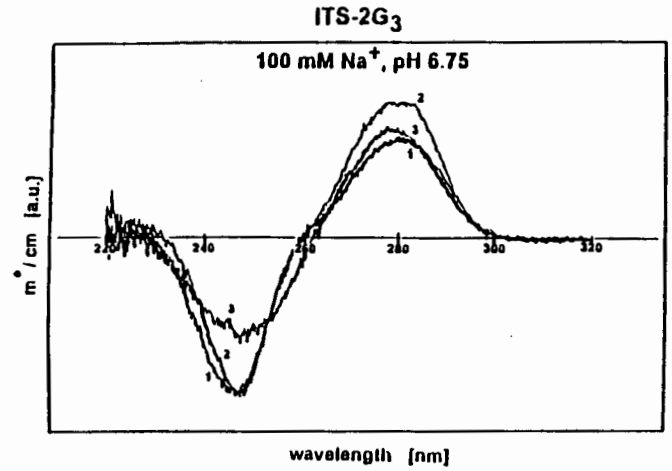
Probing of the Conformational State by CD-Spectroscopy

Because circular dichroism is sensitive to changes in the helical arrangement of DNA CD-Spectroscopy provides a useful tool for the investigation of the different conformations of the oligonucleotides. The CD-spectra of the oligonucleotide ITS-2G_3 in the region of 220 nm to 320 nm in different environmental conditions are shown in **Figure 4.3 A-C**. **Figure 4.3A** shows the CD-spectrum of the hairpin with single stranded dangling 3' tail in 100 mM Na^+, pH 8.0 and room temperature (curve 1). There is a positive band between 300 nm and 260 nm with a broad maximum around 280 nm. Between 260 nm and 230 nm the spectrum has a strong negative band with a maximum negative Cotton effect at about 248 nm. A small positive Cotton effect in the 260 nm region of the spectrum results in a shoulder. The CD-spectrum of the random coil state recorded at 70 °C shows a similar pattern, in that the ellipticity is positive between 300 nm and 260 nm and negative between 260 nm and 230 nm, however, the intensity of the negative and positive ellipticity is strongly reduced compared to the hairpin (curve 2). No evidence of the positive shoulder at 260 nm remains at 70°C. The CD-spectrum of the hairpin conformation with the exception of the positive shoulder at 260 nm is typical for a right handed B-DNA (Johnson (1990)). The major difference between the set of CD-Spectra of the different hairpin helices at pH 8.0 is the magnitude and sign of the ellipticity of the shoulder at around 260 nm. There is no evidence for this shoulder in the CD-Spectra of the oligonucleotides ITS-4G_1, ITS-3G_2, and ITS-3G_1, while for ITS-3G_0 it is weakly detectable. For ITS-2G_3 the shoulder becomes more pronounced, and finally for ITS-ATT this shoulder has changed into an independent peak. The appearance of the shoulder is positively correlated to the length of the A•T runs.

A



B



C

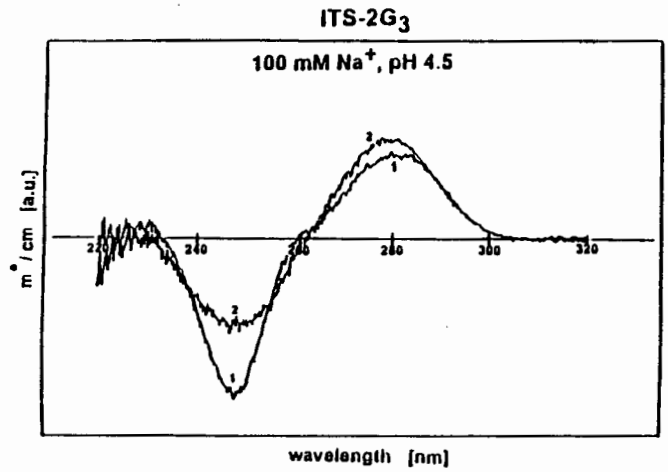


Figure 4.3: The effect of pH and temperature on the CD-spectrum of ITS-2G3
A) pH 8.0 at RT (1) and 70°C (2)
B) pH 6.75 at RT (1), 40 °C (2) and 70°C (3)
C) pH 4.5 at RT (1) and 70°C (2)

Its absence in any of the spectra at 70°C leads to the conclusion that it is due to structural features in the hairpin and not to special sequence effects. The comparison of the CD-spectrum of the hairpin conformation of ITS-ATT with that of the hairpin $A_{10}C_4T_{10}$ (van de Sande et al (1988) and own observations) indicates that the single stranded 3'dangling end contributes minimally to the ellipticity. It also confirms that the pH 8.0 conformation of ITS-ATT is indeed a hairpin.

The CD-spectrum of the triple helix conformation of ITS-$2G_3$ at a 100 mM Na^+, pH 6.75 and room temperature is shown in **Figure 4.3B**. The triple helical complex (curve 1) is characterized by a negative Cotton effect between 230 nm and 260 nm and a positive Cotton effect between 260 nm and 300 nm. The ellipticity changes sign at 260 nm. The negative Cotton effect has a minimum at 249 nm and the positive Cotton effect has a maximum at 280 nm. Compared to the hairpin (curve 2) the positive Cotton effect is reduced in intensity and is comparable to that of the random coil (curve 3). The negative band is considerably larger than that of the random coil and comparable to that of the hairpin, indicating that the triplex conformation is distinct from either random coil or hairpin. The shoulder observed in the CD-spectrum of the hairpin persists in the spectrum of the triplex. This indicates that the binding of the third strand does not grossly change the conformation of the hairpin. The reduced positive ellipticity of the triple helix makes the shoulder appear more pronounced. The CD-spectra of all other oligonucleotide triplexes at room temperature and pH 6.75 follow the same scheme. The only difference is the size of the shoulder at 260 nm which is correlated to the length of the A•T region of the hairpin component. There is a small but persistent red shift of the positive Cotton effect as the number of third strand cytosines increases, similar to the redshift in the UV-spectrum. The oligonucleotide ITS-$4G_1$ consequently shows the largest redshift in the region of positive ellipticity of all oligonucleotide triplexes. The CD-spectrum of ITS-

$4G_1$ corresponds closely to that observed for the poly d(T-C^+) • poly d(A-G) • poly d(T-C) triple helix (Antao et al (1988)).

At pH 4.5 (**Figure 4.3C**), where the intramolecular triple helix is more stable than at pH 6.75 there is an indication for a strong negative Cotton effect with a minimum at wavelength below 220 nm. Because of the increasing absorbance of the oxygen at wavelength of 220 nm and below the signal to noise ratio decreased rapidly, making it impossible to determine a reliable CD-signal in this wavelength region. A strong negative Cotton effect with a minimum at 215 nm has in the past been described as characteristic for triple helices (Johnson et al (1991), Steely et al (1986), Callahan et al (1991)). All other features of the CD-spectrum at pH 4.5 resemble the features of the CD-spectrum at pH 6.75 and room temperature as discussed above. (**Figure 4.3C**). It can be concluded that the oligonucleotide conformation at both pH values is the same.

Interaction of ITS-$4G_1$ with Ethidium Bromide

It is possible to discriminate between a double and a triple stranded conformation by its interaction with ethidium bromide. Ethidium bromide is shown to intercalate into double stranded DNA up to maximal binding capacity of 1 ethidium bromide molecule for every 2.5 base-pairs (Le Peq & Paoletti (1967)). Once intercalated the ethidium bromide fluoresces strongly in the 600 nm range. The interaction of ethidium bromide with triple helices is more complicated. While it has been reported that ethidium bromide intercalates, albeit with a reduced binding density, into the poly d(T) • poly d(A) • poly d(T) triple helix (Scaria & Shafer (1991)) no characteristic binding has been observed for the poly d(C^+T) • poly d(GA) • poly d(CT) triple helix (Lee et al (1979).

Electrostatic repulsion between the protonated third strand cytosines and the positively charged ethidium molecule have been cited as the reason for this difference (Sun et al (1991).

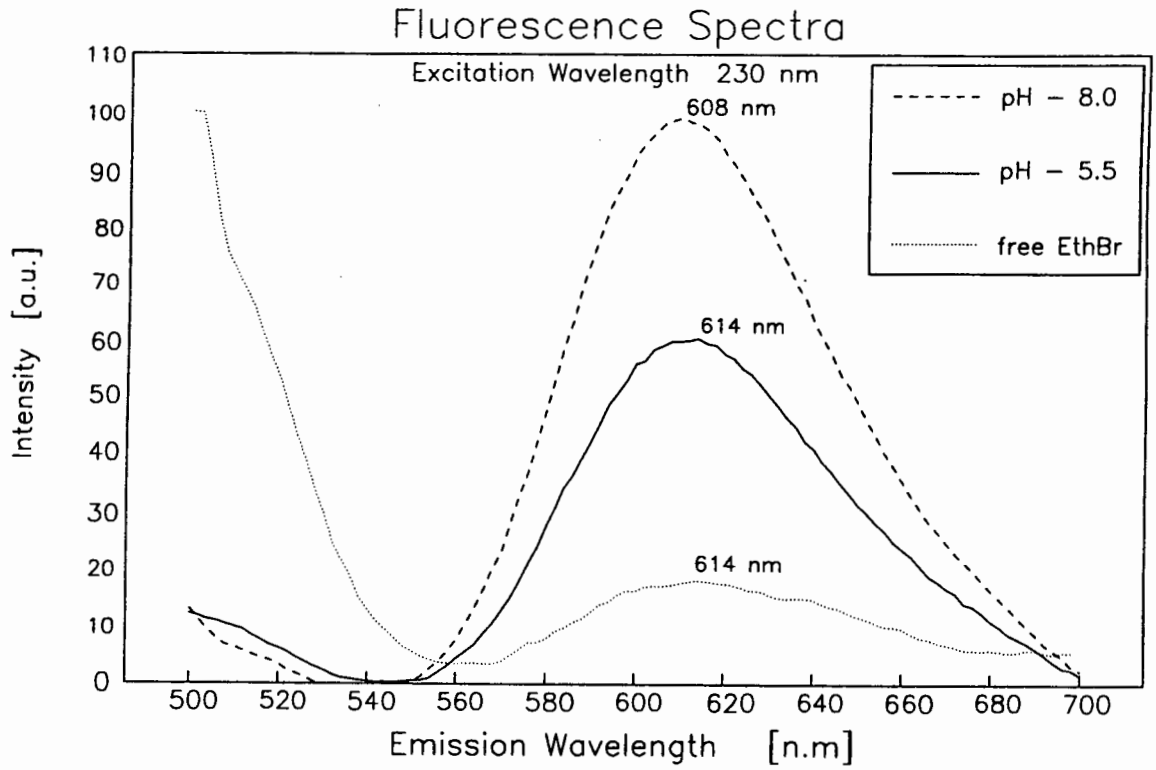


Figure 4.4: Fluorescence spectra of ethidium bromide binding to ITS-4G_1 as a function of pH

The interaction of ethidium bromide with the oligonucleotide ITS-4G_1 is shown in **Figure 4.4**. At pH 8.0 ethidium bromide exhibits all the features of intercalation into double stranded DNA. The fluorescence intensity increases strongly compared to that of free ethidium bromide, and the emission maximum is blue shifted from 614 nm to 608 nm. Reducing the pH from pH 8.0 to pH 5.5 leads to a decrease in fluorescence

intensity and a red shift in the emission maximum back to 614 nm indicating that ethidium bromide is most probably expelled back into the aqueous environment. The residual fluorescence enhancement at pH 5.5 may be due to the binding of ethidium bromide to the exterior of the triple helix and the bases in the loops. A drop in the pH for double stranded random sequence chicken DNA (146bp) does not induce a detectable change in fluorescence intensity or maximum.

Temperature dependent Fluorescence Intensity

Since the intercalation of ethidium bromide into a DNA helix is associated with an increase the fluorescence intensity, it is possible to follow the helix-coil transition by monitoring the fluorescence intensity as a function of temperature. The melting of genomic chicken DNA and of the oligonucleotides $d(A)_{10} \bullet d(T)_{10}$ and $d(T)_{10} \bullet d(A)_{10} \bullet d(T)_{10}$ has been monitored by observing the fluorescence intensity of ethidium bromide at 614 nm as a function of temperature (data not shown). Increasing the temperature leads to a gradual linear decrease in fluorescence, due to thermal quenching. In the temperature interval in which the DNA melts, a cooperative decrease in intensity is superimposed on the gradual fluorescence quenching. The midpoint of this cooperative change corresponds to the midpoint of a corresponding melting curve monitored by UV-spectroscopy. In the case of the oligonucleotides $d(A)_{10} \bullet d(T)_{10}$ and $d(T)_{10} \bullet d(A)_{10} \bullet d(T)_{10}$ the original fluorescence intensity is regained after cooling to room temperature, indicating that the intercalation is completely reversible. The melting experiments of the oligonucleotide triple helix $d(T)_{10} \bullet d(A)_{10} \bullet d(T)_{10}$ in 50 mM Mg^{2+} (Conditions as in Pilch et al (1990)) reveals two distinct cooperative melting transitions corresponding to the triple helix to duplex transition at low temperature and the double helix to coil transition at higher temperature. The results of fluorimetric experiment correspond to that of UV-melting in all details. The observation that a

cooperative decrease in fluorescence intensity accompanies the triple helix to duplex transition indicates that for this sequence ethidium bromide can intercalate into the triple helix. The oligonucleotide triple helix $d(T)_{10} \bullet d(A)_{10} \bullet d(T)_{10}$ behaves similar to the corresponding polynucleotide triple helix (Scaria & Shafer (1991)).

The melting of the oligonucleotide ITS-$4G_1$ at pH 8.0 in the presence of ethidium bromide monitored by fluorimetry reveals a single cooperative hypochromic effect. The melting temperature of this transition corresponds to that determined by UV-spectroscopy (shown later) under the same set of experimental conditions. The cooperative melting of the hairpin stem displaces the intercalated ethidium bromide molecules from the hydrophobic interior of the hairpin helix to the aqueous environment resulting in a reduction in the fluorescence intensity. At pH 6.8 the behavior of the fluorescence melting curve is more complicated. At the temperature of the triplex to hairpin transition a cooperative increase in the fluorescence intensity is observed superimposed on the gradual thermal quenching. As the third strand melts, sites on the hairpin stem become available for ethidium bromide intercalation, leading to an increase in the fluorescence intensity. At higher temperature a cooperative decrease in fluorescence intensity occurs corresponding to the melting of the hairpin stem. At pH 5.5 only the gradual decrease in fluorescence intensity due to thermal quenching is observed, while under the same conditions a UV-melting curve shows a single cooperative helix to coil transition. This one step melting corresponds to the intramolecular triple helix to the random coil transition. Since no corresponding change in fluorescence intensity can be detected it must be concluded that the ethidium bromide can not intercalate into the pH 5.5 conformation of ITS-$4G_1$.

UV-melting curves of ITS-$4G_1$ at pH 6.7 in the presence of an increasing ratio of ethidium bromide to nucleotide phosphates reveal that the ethidium bromide selectively

stabilizes the hairpin and destabilizes the intramolecular triple helix. This is in accordance with the theoretical assumption that a ligand selectively stabilizes the DNA conformation to which it binds preferentially and destabilizes the conformation to which it binds with a reduced binding affinity. Mergny et al (1991) observed a similar behavior for a triple helix formed from three independent oligonucleotides.

Enzymatic Probing with a Single Strand Specific Nuclease

The two oligonucleotide secondary structures (hairpin and intramolecular triple helix) are characterized by the difference in the length and position of single stranded regions. The hairpin conformation at pH 8.0 contains a single stranded extension of 14 bases at the 3'end and a partially single stranded four membered hairpin loop. In the intramolecular triple helix on the other hand only the four bases in the hairpin loop and those in the triplex loop are partially single stranded. It should be easily possible to distinguish between the two conformations by the pattern of degradation by single strand specific agents such as P_1-nuclease followed by denaturing gel electrophoresis of the radiolabled fragments. P_1-nuclease was chosen in preference to the more commonly used S_1-nuclease, because of the broader pH and salt tolerance of the enzyme. An additional advantage of using P_1-nuclease is that the essential cofactor Zn^{2+} is complexed to the enzyme and hence does not change the environmental conditions by additional Zn^{2+} ions in the reaction buffer (Blaho et al (1988)). Since Zn^{2+} can potentially interact with DNA at the bases and thus induce conformational changes (e.g. Bernues et al (1989)) it is essential to keep the concentration of free Zn^{2+} ions in solution to a minimum.

Figure 4.5 shows the result of the P_1-nuclease digestion of ITS-4G_1 at pH 8.0 and pH 5.5. At pH 8.0 the P_1-nuclease cuts predominantly the 3'tail of the oligonucleotide, consistent with the hairpin conformation of ITS-4G_1. No cutting within the hairpin stem

region and very little cutting within the hairpin loop is observed. At pH 5.5 a small degree of cutting in the triplex loop and two very strong cuts in the hairpin loop are observed. No other site shows any sign of nuclease digestion. This pattern of electrophoretic bands is consistent with the formation of an intramolecular triple helix, which has no single stranded regions except within the loops. It is somewhat surprising that the hairpin loop is not digested at pH 8.0, but shows two strong cutting sites at pH 5.5. The lack of cutting of P_1-nuclease in the hairpin loop at pH 8.0 suggests that the loop bases are not recognized as the usual single stranded substrate by the P_1-nuclease. Baumann et al (1986) have observed that in short hairpins with three to six adenines in the loop region both P_1 and S_1 nucleases show only a small preference for the loop. The bases in a hairpin loop can potentially form a partially stacked structure. A considerable degree of order has been reported for four membered hairpin loops previously. (Hare & Reid (1986), Boulard et al (1991), Blommers et al (1991), Erie et al (1993)).

The observation that the hairpin loop suddenly becomes accessible to nuclease digestion at pH 5.5, while it is completely inaccessible at pH 8.0, indicates that the loop is in a different conformational state. The change can potentially be brought about by the binding of the third strand to the major groove in the triple helix. The binding of the third strand causes a conformational change in the double helix host (Arnott et al (1976), Raghunathan et al (1993)). A conformational change in the hairpin stem can propagate into the hairpin loop - leading to the unstacking of the loop bases. These bases are then substrates for the P_1-nuclease. Alternatively it is reasonable to assume that the cytosines in the hairpin loop become partially protonated at pH 5.5. Model building shows that this can give rise to a two membered hairpin loop closed by an unusual $C \bullet C^+$ base-pair between cytosine 1 and cytosine 4. Due to sterical constraints the cytosine 2 & 3 in a two membered loop can not stack with themselves or the unusual $C \bullet C^+$ base-pair and will thus provide a substrate for the P_1-nuclease.

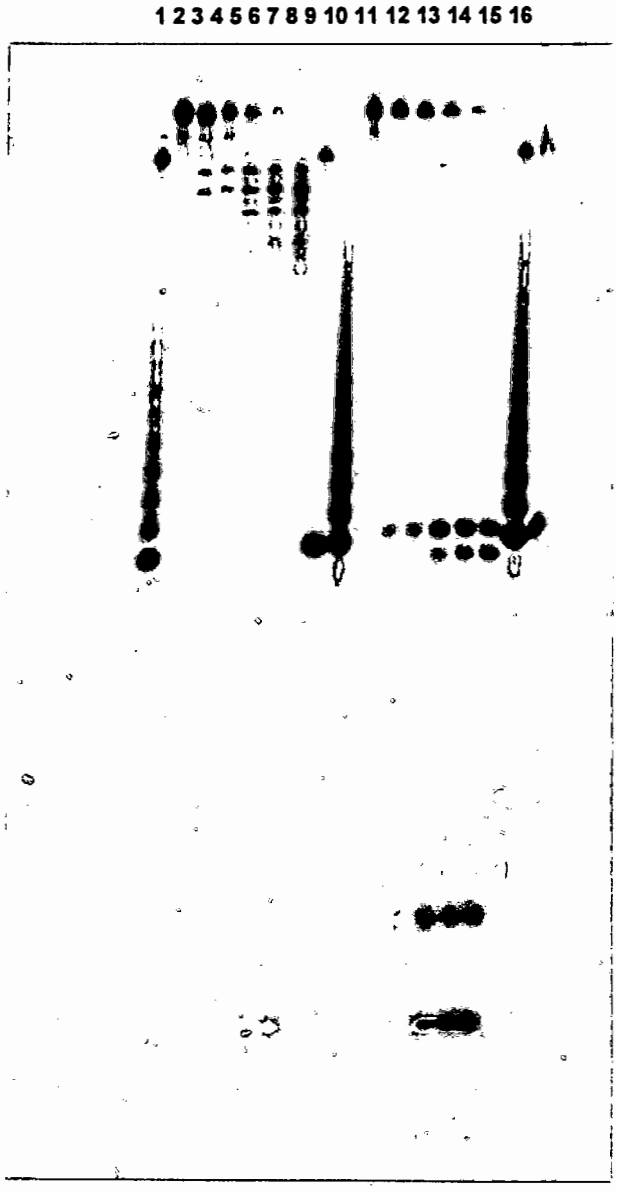


Figure 4.5: P_1-nuclease digestion of ITS-4G_1 at pH 8.0 and pH 5.5. Three units (pH 8.0) or 0.003 units (pH 5.5) of P_1-nuclease were added to (γ-^{32}P) labeled oligonucleotide. Samples were taken at intervals of 0 s, 60 s, 120 s, 300 s, 600 s and 900 s, and run on a 20% denaturing polyacrylamide gel. (**A**) Lanes 1,9 and 16 contain a standard ladder produced by the addition of random nucleotides to a T_{10} mer oligonucleotide with the help of terminal nucleotide transferase and an untreated 33 mer. Lane 8 is a standard lane of untreated 33 mer only. Lanes 2 to 7 represent the digestion at pH 8.0 at increasing time intervals, lanes 10 to 15 represent the digestion at pH 5.5 at increasing time intervals.

B

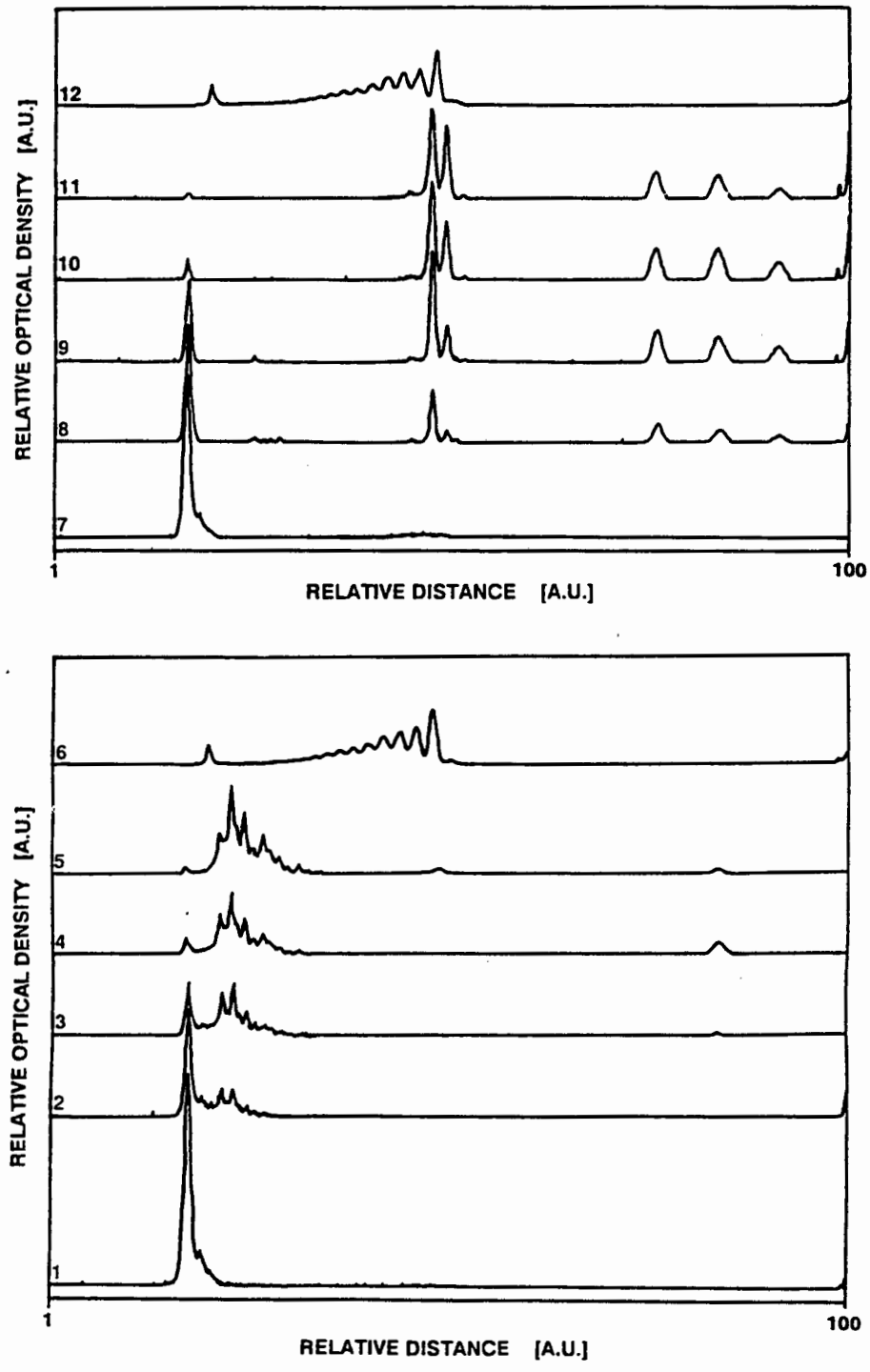


Figure 4.5: (**B**) Densitometric trace of selected lanes of the denaturing gel in (**A**). Lanes 1 to 5 digestion at pH 8.0, lanes 7 to 11 digestion at pH 5.5. Lanes scanned were for digestion times of 0s, 120 s, 300s, 600 s and 900 sec respectively. Lanes 6 and 12 are standard lanes.

This latter explanation is also consistent with the observation that the enzyme cleaves predominantly at two position in the loop. Hairpin loops of only two bases closed by an unusual base-pair have been observed before by NMR (e.g. Blommers et al (1989)). Supporting evidence for this latter explanation comes also from pH titration of the intramolecular triple helix.

P_1-nuclease digestion pattern of all other oligonucleotides is very similar to the digestion pattern of the corresponding reaction of ITS-$4G_1$, indicating the same formation of the hairpin with dangling 3' tail and the intramolecular triple helix.

A completely different behavior is observed for the oligonucleotide ITS-ATT. Digestion with P_1-nuclease results in nearly complete digestion. The reason for this anomalous behavior is simply that the A_{10}•T_{10} helix is unstable, so that any nick in the hairpin or triplex loop will lead to a denaturation of the complex. This results in the digestion of the single strands.

Another interesting observation based on the intensity of the bands in the denaturing gel is that the digestion of the 3' overhanging terminus at pH 8.0 by P_1-nuclease is not uniform throughout. **Figure 4.5** reveals a pattern of alternating strong and weak cutting sites. Since the single stranded tail of the oligonucleotide ITS-$4G_1$ consists predominantly of a sequence of alternating thymine and cytosine bases, it is obvious that the P_1-nuclease shows a sequence preference for one of these bases. A close inspection of all digestion patterns reveals that the strongest cutting frequency always corresponds to the position of the cytosines. It seems that P_1-nuclease has a sequence preference for cytosines at least under the conditions used here. From the resolution of the polyacrylamide gel it is not possible to decide whether P_1-nuclease cuts preferentially the $^{5'}C\text{-}T^{3'}$ or the $^{5'}T\text{-}C^{3'}$ step. Sequence specific cutting has previously been assumed for

S_1-nuclease by Pullyblank et al (1988) to explain the results of the S_1-nuclease digestion of homopurine • homopyrimidine runs in supercoiled plasmids. The observed sequence preference by P_1-nuclease raises some concern over analyzing nucleic acid conformations based on the reaction towards "single strand specific" reagents, if little or nothing is known about the reaction mechanism of these reagents.

All six Oligonucleotides form Intramolecular Triplexes

The pH induced change in absorbance observed by UV-spectroscopy shows the sequential formation of two ordered conformations in agreement with the proposed model, i.e. the formation of first the hairpin with an extended 3'end and second the intramolecular triple helix. CD-spectroscopy results have confirmed the presence of these two conformers. The digestion pattern of P_1-nuclease is in accordance with the hairpin with single stranded 3' tail at pH 8.0 and the intramolecular triple helix at pH 5.5. The interaction of ITS-$4G_1$ with ethidium bromide is also consistent with this interpretation. Comparing the set of UV-spectra, of CD-spectra and of P_1-nuclease digestion patterns of the six oligonucleotides listed in **Table V** indicates that all these oligonucleotides fold along the same pathway. Minor differences in CD-spectra and the P_1-nuclease pattern are a reflection of the local and global differences in sequence and hence small sequence dependent variations in conformation, rather than indications for the formation of alternative structures. Since all oligonucleotides fold into the same overall conformations it is possible to compare the thermodynamic data for the unfolding in order to elucidate the impact of local and global composition on triple helix stability.

4.2 Thermodynamics of the Intramolecular Triple Helix Formation

Thermal Denaturation Experiments

The thermal stability of oligonucleotide and polynucleotide secondary structures can be measured by following the change in absorption at 260 nm as a function of temperature. **Figure 4.6** shows the melting curves of the different conformations of ITS-$4G_1$ in 100 mM Na^+ at the appropriate pH. At pH 8.0 a single cooperative melting transition is observed. This transition is characterized by a transition temperature (T_m) of $T_m = 63 \pm 1°C$, a transition width of $\delta T = 12.8 \pm 0.8°C$ and a hyperchromic change of %h = 7.8%. It corresponds to the melting of the hairpin to the random coil. At pH 6.75 the change in absorbance with temperature reveals two different transitions. The upper transition corresponds to the pH 8.0 transition. It is reasonable to assume that it represents the melting of the hairpin to the random coil. The lower transition is characterized by a melting temperature $T_m = 41.0 \pm 1°C$, a transition width of $\delta T = 12.6 \pm 0.5°C$ and a hyperchromicity of similar magnitude as the hairpin to coil transition. This transition is assigned to the unfolding of the intramolecular triple helix to the hairpin. At pH 4.5 only one transition remains. This transition is characterized by a melting temperature of $T_m = 71 \pm 1.0°C$, a transition width of $\delta T = 8.0 \pm 0.7°C$ and a hyperchromicity of %h = 16%. It represents the melting of the intramolecular triple helix to the random coil. It is a matter of choice of the appropriate environmental conditions to show the same conformational changes (hairpin $\leftrightarrow$ coil, triplex $\leftrightarrow$ hairpin $\leftrightarrow$ coil and triplex $\leftrightarrow$ coil) for all other oligonucleotides listed in **Table V**. The results are listed in It seems appropriate to impose an arbitrarily determined confidence interval of $T_m \pm 1.0°C$ on the experimental data.

TABLE VI (A):

THERMODYNAMIC PARAMETERS OF ITS-ATT IN 100 mM NaCl BUFFER:

	pH 8.0	pH 6.75	
		1.st **	2.nd
Tm [°C]	49.3 ± 1.0***	28.3 ± 1.0	49.8 ± 1.0
δT [°C]	12.6 ± 0.6	14.6 ± 0.4	13.2 ± 1.2
ΔH_{vH} [kcal/M]*	67.4 ± 4.7	(53.2 ± 3.0)	67.5 ± 4.2
ΔS_{vH} [cal/M*K]	207.2 ± 12.1	(171.0 ± 6.6)	206.8 ± 9.6

---

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

TABLE VI (B):

THERMODYNAMIC PARAMETERS OF ITS-$2G_3$ IN 100 mM NaCl BUFFER:

	pH 8.0	pH 6.75		pH 4.6
		1.st **	2.nd	
Tm [°C]	53.8 ± 1.0***	32.1 ± 1.0	54.2 ± 1.0	56.0 ± 1.0
δT [°C]	12.5 ± 0.4	16.6 ± 0.9	12.9 ± 0.4	10.3 ± 0.9
ΔH_{vH} [kcal/M] *	71.5 ± 3.4	(47.7 ± 2.9)	69.5 ± 3.7	90.5 ± 9.4
ΔS_{vH} [cal/M*K]	214.0 ± 8.0	(151.8 ± 8.3)	207.6 ± 8.5	266.8 ± 25.3

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

TABLE VI (C):

THERMODYNAMIC PARAMETERS OF ITS-3G_0 IN 100 mM NaCl BUFFER:

	pH 8.0	pH 6.75		pH 4.6
		1.st **	2.nd	
Tm [°C]	57.8 ± 1.0***	31.9 ± 1.0	58.3 ± 1.0	62.0 ± 1.0
δT [°C]	12.7 ± 0.4	9.6 ± 0.5	11.4 ± 0.3	8.4 ± 0.4
ΔH_{vH} [kcal/M]*	68.8 ± 1.8	(77.5 ± 3.4)	77.3 ± 2.1	106.6 ± 4.8
ΔS_{vH} [cal/M*K]	207.9 ± 5.6	(254.2 ± 11.4)	233.2 ± 6.5	318.3 ± 14.5

---

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

TABLE VI (D):

THERMODYNAMIC PARAMETERS OF ITS-3G_1 IN 100 mM NaCl BUFFER:

	pH 8.0	pH 6.75		pH 4.6
		1.st **	2.nd	
Tm [°C]	57.7 ± 1.0***	36.5 ± 1.0	60.5 ± 1.0	66.7 ± 1.0
δT [°C]	11.9 ± 0.4	11.3 ± 0.9	11.4 ± 0.5	7.6 ± 0.4
ΔH_{vH} [kcal/M]*	75.9 ± 4.7	(70.3 ± 3.7)	80.2 ± 2.7	127.2 ± 8.2
ΔS_{vH} [cal/M*K]	225.9 ± 8.8	(223.2 ± 12.5)	235.8 ± 5.65	364.9 ± 19.9

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

TABLE VI (E):

THERMODYNAMIC PARAMETERS OF ITS-3G_2 IN 100 mM NaCl BUFFER:

	pH 8.0	pH 6.75		pH 4.6
		1.st **	2.nd	
Tm [°C]	56.8 ± 1.0***	37.0 ± 1.0	58.9 ± 1.0	65.0 ± 1.0
δT [°C]	13.3 ± 0.7	15.3± 1.4	12.8 ± 0.6	N.D.****
ΔH_{vH} [kcal/M]*	67.8 ± 1.9	(52.3 ± 3.8)	71.5 ± 2.2	N.D.
ΔS_{vH} [cal/M*K]	202.0 ± 7.8	(166.3 ± 14.0)	212.0 ± 7.4	N.D.

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

**** N.D. - not determined

TABLE VI (F):

THERMODYNAMIC PARAMETERS OF JV-ITS IN 100 mM NaCl BUFFER

	pH 8.0	pH 6.75		pH 4.6
		1.st **	2.nd	
Tm [°C]	63.0 ± 1.0***	41.0 ± 1.0	63.0 ± 1.0	71.0 ± 1.0
δT [°C]	12.8 ± 0.8	12.6 ± 0.5	12.8 ± 0.8	8.0 ± 0.7
ΔH_{vH} [kcal/M]*	71.7 ± 4.2	(58.7 ± 4.2)	71.7 ± 4.0	119.4 ± 11.9
ΔS_{vH} [cal/M*K]	212.2 ± 12.8	(191.7 ± 12.8)	212.2 ± 12.8	346.6 ± 34.2

---

* Triplex to Hairpin transition (pH 6.75) does not proceed according to a two state process.

** 1 kcal = 4.18 kJ

*** ±1°C Error estimate for T_m to exclude the effect of changes in buffer with time. The calculated error was ±0.2°C - the same as the sampling density

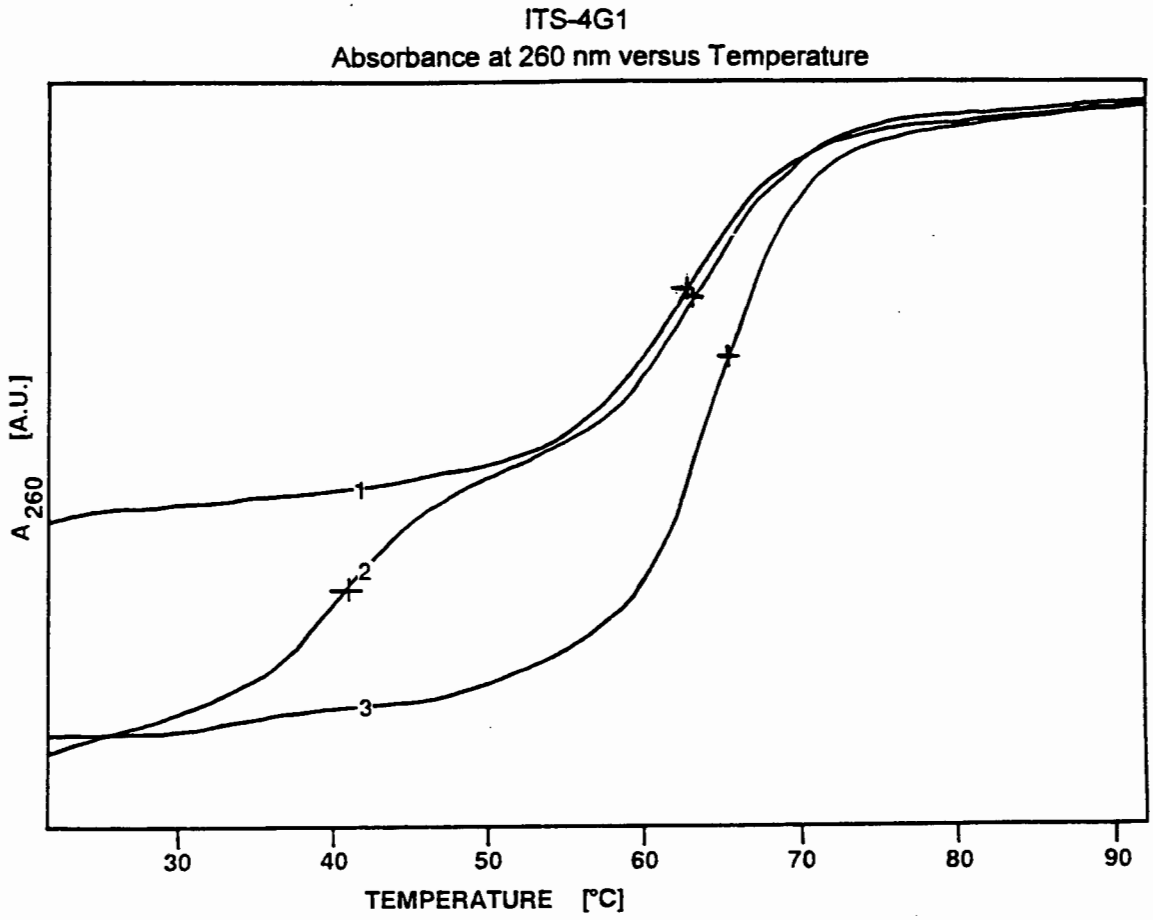


Figure 4.6: Thermally induced helix to coil transitions of ITS-4G_1: UV melting curves at 260 nm at pH 8.0 (1), pH 6.75 (2) and pH 5.5 (3);

Since T_m can be measured very accurately and changes in T_m are readily detectable, it is tempting to put special significance on small T_m changes. By imposing an arbitrary confidence interval marginal effects are excluded and only relatively large changes are considered significant. For example the T_m values of the hairpin to coil transition of ITS-3G_0, ITS-3G_1 and ITS-3G_2 are considered to be identical, even though they vary slightly.

Unfolding of intramolecular secondary structures should be independent of nucleotide concentration. To verify this effect of oligonucleotide concentration on the melting temperature of the three conformational transitions of ITS-4G_1 and ITS-3G_0 was

determined. A 10 fold increase in concentration does not result in any change of the T_m of the three conformational transitions. As expected there is no evidence for the competing intermolecular reaction at oligonucleotide concentrations used for spectroscopic measurements (UV- & CD-spectroscopy). The effect of oligonucleotide concentration on the melting temperature of the other oligonucleotides was not determined separately. Since all these oligonucleotides show the same physical characteristics and behave similarly in every other aspect, it was assumed that no significant contributions from competing intermolecular complexes occur.

To determine whether the occurrence of intermediate states in the unfolding of the intramolecular triple helix conformation was kinetically inhibited, the effect of the heating rate on the melting temperature of ITS-$4G_1$ at pH 6.75 (100 mM Na^+) was determined. Increasing the heating rate 10 fold from about 0.2 °C/min to 2 °C/min does not effect the melting temperature or the shape of the melting curve of the triple helix to hairpin and hairpin to coil transition. In either case the thermal behavior of the system was controlled by the thermal inertia of the equipment than by kinetic effects of the system.

The Effect of the Environmental Conditions on the T_m

From the theoretical considerations of triple helix stability outlined in chapter II it is clear that the local and global compositional effects modulate the impact of environmental changes on the thermal stability of the triple helix. The environmental conditions most likely to influence triple helix stability are the gross ionic strength and the pH of the solution. **Figure 4.7** shows the effect of the ionic strength at pH 8.0 on the hairpin to coil transition of the oligonucleotide ITS-$3G_0$ observed by the change in UV-absorption at 260 nm.

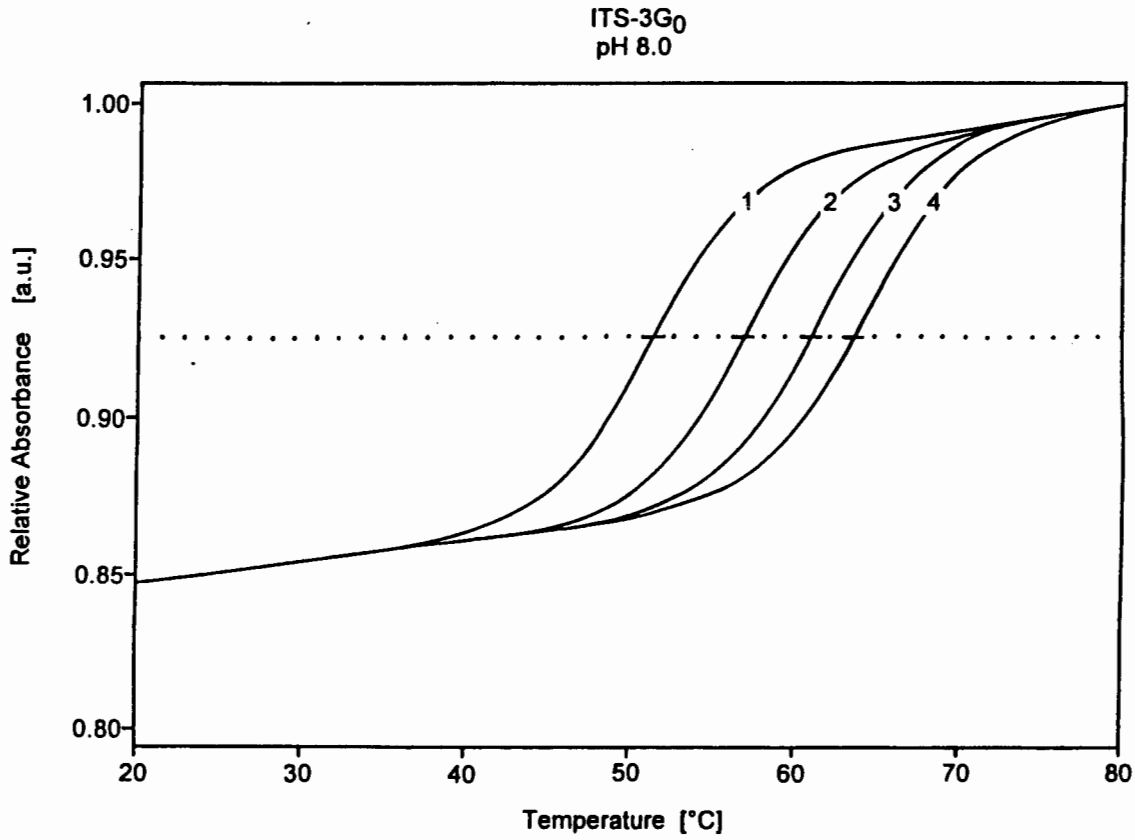


Figure 4.7: The effect of ionic strength on the melting temperature of the hairpin to coil transition of ITS-$3G_0$ at pH 8.0:
1 - 50 mM Na^+; **2** - 100 mM Na^+; **3** - 200 mM Na^+; **4** - 300 mM Na^+

The effect of the ionic strength on the triple helix to hairpin and on hairpin to coil transition at pH 6.75 is shown in **Figure 4.8**. **Figure 4.9** shows the effect of the ionic strength on the intramolecular triple helix to coil transition at pH 4.5. The results are summarized in a phasediagram of T_m versus log [Na^+] in **Figure 4.10**. Phasediagrams of this kind have been obtained for all oligonucleotides (**Appendix A-C**).

From the linear region of the plots (T_m versus log [Na^+]) the slope ($d(T_m)/d(\log [Na^+])$) can be obtained by linear least square curve fitting. For the hairpin to coil transitions (curve 1 & 2) the plot is linear up to at least 700 mM Na^+. A further increase of the counterion concentration results in a nonlinear negative deviation. This is observed for the helix to coil transition of polynucleotides as well. It is assumed that the nonlinear behavior is due to effects of high ion concentration on the water structure.

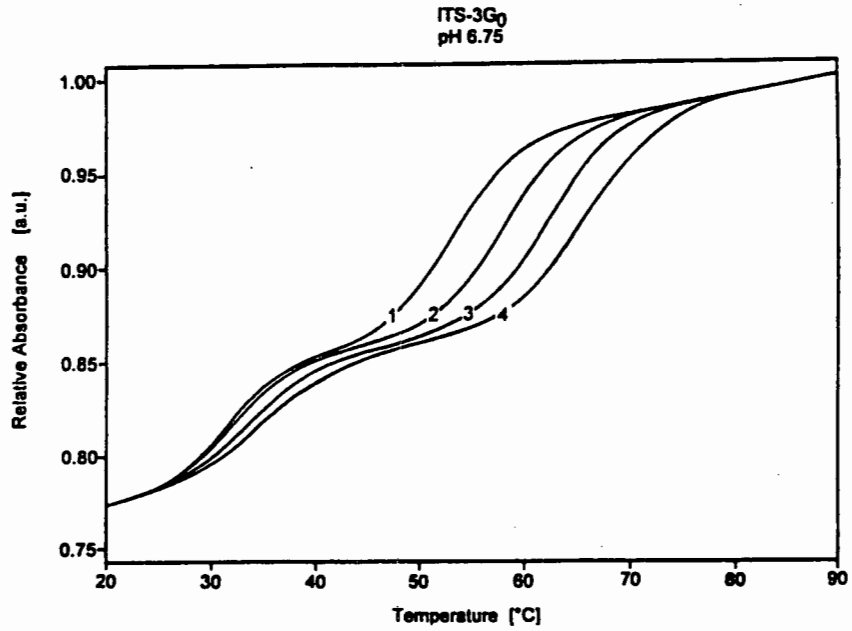


Figure 4.8: The effect of ionic strength on the melting temperature of the triple helix to hairpin and the hairpin to coil transition of ITS-$3G_0$ at pH 6.75:
1 - 50 mM Na^+; **2** - 100 mM Na^+; **3** - 200 mM Na^+; **4** - 300 mM Na^+

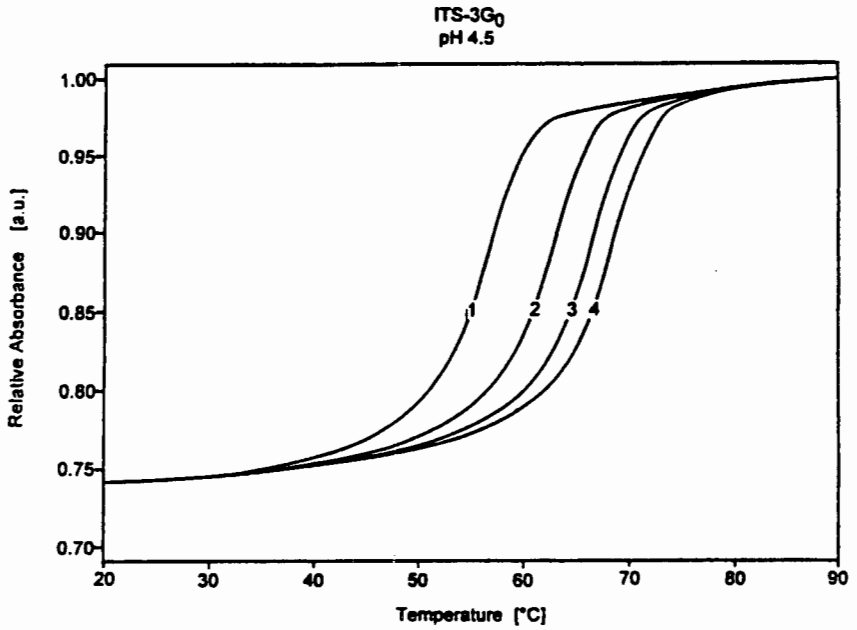


Figure 4.9: The effect of ionic strength on the melting temperature of the triple helix to coil transition of ITS-$3G_0$ at pH 4.5:
1 - 50 mM Na^+; **2** - 100 mM Na^+; **3** - 200 mM Na^+; **4** - 300 mM Na^+

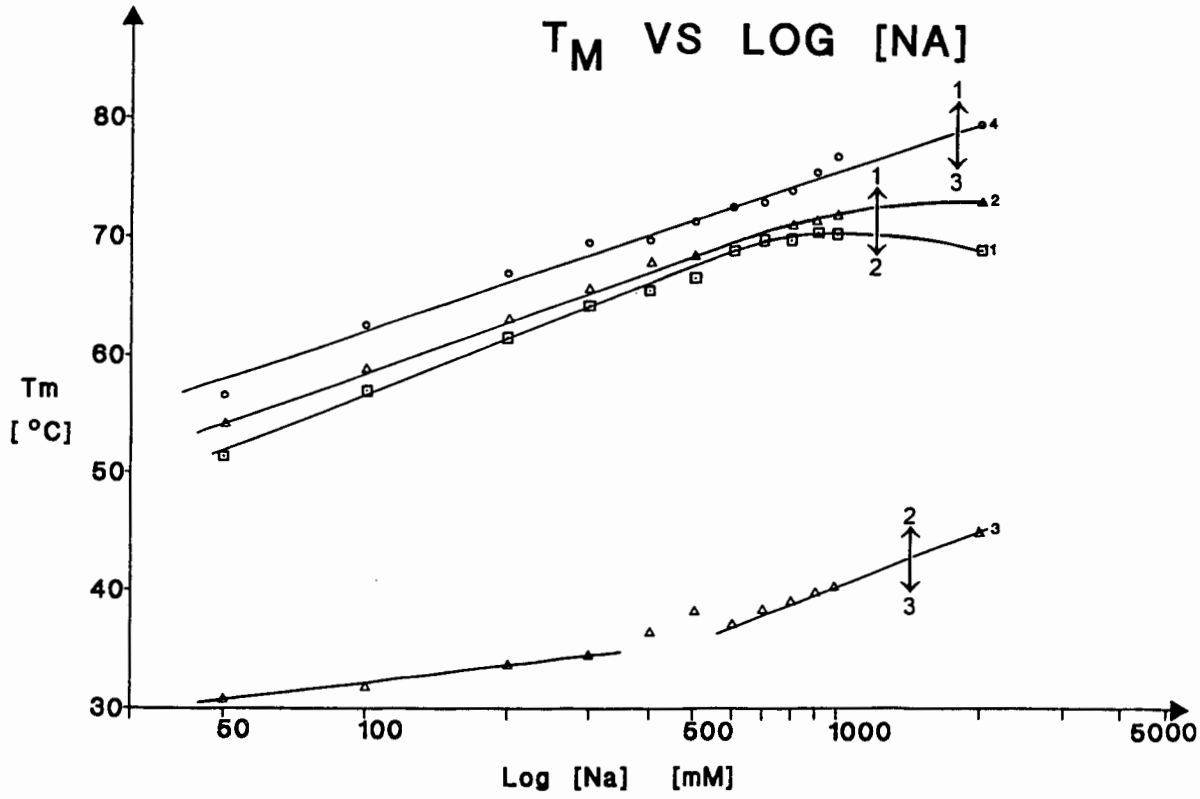


Figure 4.10 Phasediagram of ITS-3G_0: transition temperature (T_m) versus ionic strength for the various conformational transitions of ITS-3G_0:
(1) the hairpin to coil transition at pH 8.0
(2) the hairpin to coil transition at pH 6.75
(3) the triple helix to hairpin transition at pH 6.75 and
(4) the triple helix to coil transition at pH 4.5

For the triple helix to coil transition (curve 4) the linear region of a plot extends to a concentration of at least 1 M Na^+. It is possible that the linearity at high Na^+ concentrations results from compensation effects. For the triple helix to hairpin transition at pH 6.75 (curve 3) a drastic change in slope occurs at around 500 mM Na^+. Below 200 mM Na^+ the slope is three fold less than the slope at 1 M Na^+. This can only be understood, if one assumes that the extend of screening changes drastically over the concentration range. This will be discussed later. The $d(T_m)/d(\log [Na^+])$ values for the different oligonucleotides are listed in **Table VIII** (hairpin to coil transition), **Table IX** (triple helix to hairpin transition) and **Table X** (triple helix to coil transition).

To investigate the effect of pH on the stability of the different conformations the T_m was determined as a function of pH at a fixed ionic strength of 150 mM Na^+. The results for the oligonucleotide ITS-$3G_0$ are compiled in a phasediagram in **Figure 4.11**. Phasediagrams of this kind have been obtained for all oligonucleotides (**Appendix D**). A linear least square fitting procedure allows the determination of the linear equation describing the relationship between the T_m and the solution pH ($d(T_m)/d(pH)$) for the various conformers. Similar phasediagrams have been obtained for all other oligonucleotides studied here. The T_m of the hairpin to coil transition is only marginally dependent on pH.

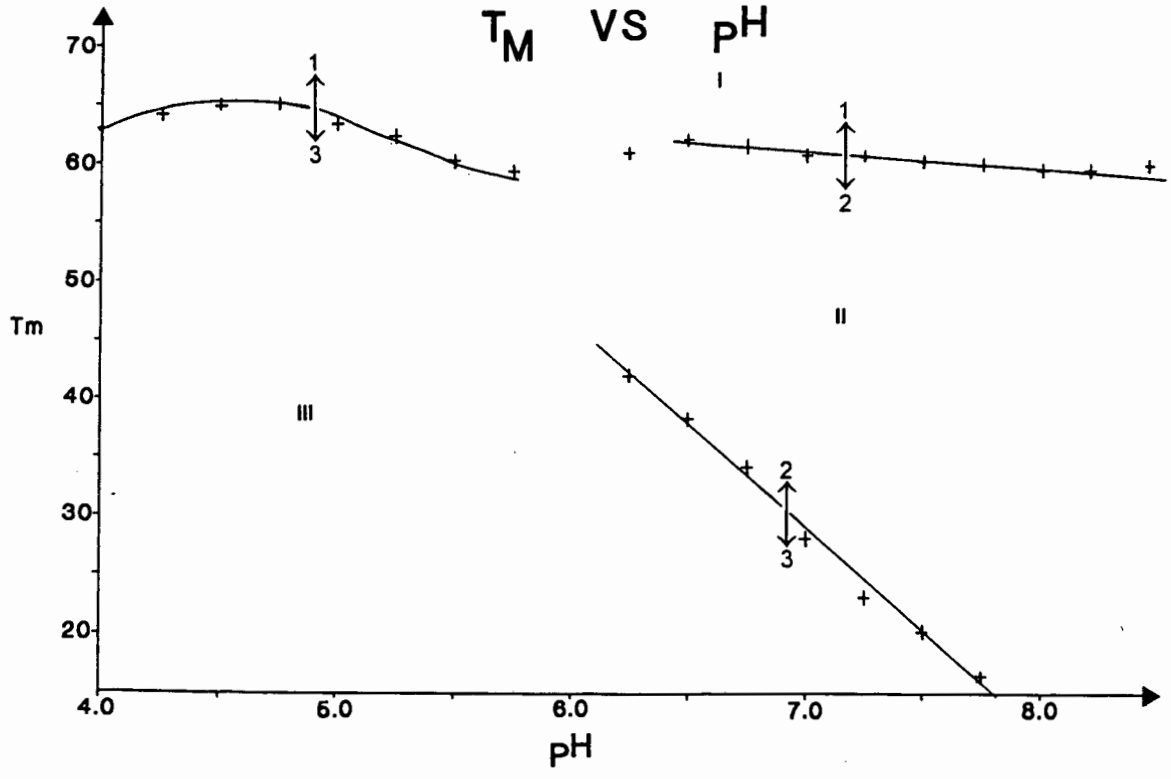


Figure 4.11 Phasediagram of ITS-$3G_0$: transition temperature (T_m) versus pH. Roman numericals denoted the region of stability for the various conformations of ITS-$3G_0$:
(**I**) random coil
(**II**) hairpin with 3'dangling end
(**III**) intramolecular triple helix

This corresponds to the behavior observed for double stranded polynucleotides and genomic DNA (Record (1967)). The T_m of the intramolecular triple helix to hairpin transition on the other hand increases strongly with a decrease in pH. The T_m of the intramolecular triple helix to coil transition also changes by the change in pH. But unlike the case of the intramolecular triple helix to hairpin transition the T_m for the intramolecular triple helix to coil transition does not change linearly with a change in pH. The T_m increases almost linearly with decreasing pH up to pH 5.0. As the pH approaches the pka of free cytosine, the T_m increase levels of, reaching a maximum at pH 4.5. With further lowering of the pH the T_m becomes progressively smaller. The $d(T_m)/d(pH)$ values for the different conformational transitions of the different oligonucleotides are listed in **Table VIII** (hairpin to coil transition), **Table IX** (triple helix to hairpin transition) and **Table X** (triple helix to coil transition).

The Energetics of Secondary Structure Formation

The van't Hoff enthalpy for a temperature dependent cooperative equilibrium reaction can be calculated form the UV-melting curves according to the equation:

$$\Delta H_{vH} = 2(n+1) * R * T_m^2 * (\delta\theta/\delta T)_{Tm} \qquad (4.1)$$

where R is the universal gas constant, T_m the melting temperature (in Kelvin) and $(\delta\theta/\delta T)_{Tm}$ the slope of the melting curve at T_m. **n** refers to the number of single strands that form the helical complex (Marky & Breslauer (1987)). For an intramolecular reaction **n** = 1 and for an intermolecular reaction that leads to a double helix **n** = 2. Since the hairpin and the triple helical state are intramolecular reactions involving one and the same strand, **n** for these reactions has to be **n** = 1. The van't Hoff equation is valid only for true two state processes. The formation of partially melted intermediate states makes the application of equation (**4.1**) invalid, and hence leads to incorrect ΔH_{vH} values. In

case these conformational states are not significantly populated, application of equation (**4.1**) leads to a reasonable approximation of the ΔH_{vH} values. The occurrence of partially melted intermediate states can be deduced from a comparison of the calorimetric ΔH_{cal} value and the model dependent ΔH_{vH} value for a particular transition.

Hairpin Melting

The denaturation enthalpy of the different conformations of the oligonucleotide ITS-4G$_1$ have been determined by calorimetry (ΔH_{cal}) and van't Hoff analysis of UV-melting curves (ΔH_{vH}). The calorimetric enthalpy of the hairpin to coil transition of ITS-4G$_1$ at pH 8.0 (ΔH_{cal} = 69.9 ± 3.5kcal/M (1cal = 4.184J)) is found to corresponds within the margins of the experimental error to the van't Hoff enthalpy (ΔH_{vH} = 71.7 ± 4.0kcal/M). From the ratio of ΔH_{cal} to ΔH_{vH} it can be concluded that the melting of the hairpin to the coil state can be described as a two state process. For comparison it is possible to calculate the denaturation enthalpy from incremental contributions of the individual base-pairs using a data set derived from the melting of a variety of model polynucleotides (Klump (1988)). It is necessary to take end effects into account if the increment method is applied to oligonucleotides. Summation of the enthalpy of each of the ten base-pairs of the hairpin stem yields a computational enthalpy of ΔH_{comp} = 76kcal/M (see **Table VII**). The close agreement between the computational enthalpy and the experimental enthalpy considering the unknown contributions of the hairpin loop and 3' single stranded tail is gratifying. Due to instrument failure it was not possible to determine the calorimetric denaturation enthalpy ΔH_{cal} for any of the other oligonucleotides. However the close agreement of the thermal behavior of the hairpin to coil transition ITS-4G$_1$ with that of the other oligonucleotides justifies the comparison of the ΔH_{vH} values with the ΔH_{vH} value and the ΔH_{cal} value of ITS-4G$_1$. The thermodynamic data are compiled in **Table VI A-F**.

TABLE VII: Nearest Neighbour Interactions in the Hairpin Stem

Nearest Neighbour Δ H [kcal/M]	A - A 8.08	A - G 8.61	G - A 8.61	G - G 9.09	END (X-Y)/2	LOOP (A-C)/2	ΔH_{comp}
ITS-ATT	8 64.64	- -----	- -----	- -----	A - A 4.04	1 4.22	72.90
ITS-$2G_3$	6 48.48	1 8.61	1 8.61	- -----	G - A 4.31	1 4.22	74.23
ITS-$3G_0$	5 40.40	1 8.61	1 8.61	1 9.09	G - A 4.31	1 4.22	75.24
ITS-$3G_1$	4 32.32	2 17.22	2 17.22	- -----	G - A 4.31	1 4.22	75.29
ITS-$3G_2$	4 32.32	2 17.22	2 17.22	- -----	G - A 4.31	1 4.22	75.29
ITS-$4G_1$	2 16.16	3 25.83	3 25.83	- -----	G - A 4.31	1 4.22	76.35

The good agreement between the experimental van't Hoff enthalpy values is mirrored by the corresponding close agreement of the computational enthalpy values. Differences in the sequence reflected in differences of the computational enthalpy values result in a maximal difference ($\Delta\Delta H_{comp}$) not exceeding 4 kcal/M. This is well within the experimental error for the melting of a 10-mer and hence no large differences in ΔH_{vH} for the different oligonucleotide hairpins can be expected.

Triple Helix Melting

The calorimetric enthalpy for the intramolecular triple helix to coil transition of ITS-4G_1 at pH 4.5 is $\Delta H_{cal} = 121.9 \pm 6.1$kcal/M which compares well with the corresponding van't Hoff enthalpy $\Delta H_{vH} = 119.4 \pm 11.9$kcal/M. A tentative computational enthalpy for this transition has been calculated from the ΔH_{comp} of the hairpin and an enthalpy value of $\Delta H_{cal} = 4$kcal/M $_{\text{third strand pyrimidine base}}$ determined for the poly r(U) • poly r(A) • poly r(U) by Neumann & Ackermann (1969). This computational enthalpy comes to $\Delta H_{comp} = 116$kcal/M. The computational enthalpy for the triple helix formation, although it comes persuasively close to the experimental values, has to be regarded as tentative. Due to lack of calorimetric data it is based on the melting enthalpy of a ribonucleic acid triple helix and on a small set of calorimetric data. Never the less the close agreement of ΔH_{cal}, ΔH_{vH} and ΔH_{comp} provide some confidence for this approach. It also indicates that the triple helix to coil transition can be described as a two state process. The apparent one step melting of the intramolecular triple helix to the random coil can be considered as the superposition of the melting of the triple helix to the hairpin and that of the hairpin to the random coil. Considering that the melting temperature of the intramolecular triple helix at pH 4.5 is considerably higher than that of the (integrated) hairpin at pH 8.0 it is obvious that on melting of the third strand the resulting hairpin becomes very unstable and melts immediately. It was not possible to

determine whether the differences in ΔH_{vH} observed for the intramolecular triple helix to coil transitions of the other oligonucleotides is reflected in differences in enthalpy of the third strand melting or whether it is due to a non-two-state melting process (**Table VI A-F**). There is a noticeable difference between ΔH_{cal} and ΔH_{vH} of third strand binding in the literature. Plum et al. (1990) have determined an enthalpy of melting for an oligonucleotide triple helix of as little as ΔH_{cal} = 2 kcal per mole of third strand base by microcalorimetry. Introducing this value into the computation of ΔH_{comp} results in a ΔH_{comp} = 96kcal/M for the complete intramolecular triple helix to coil transition. This value is very close to the experimental van't Hoff enthalpy values for some of these oligonucleotides.

Third Strand Melting

For the intramolecular triple helix to hairpin transition of ITS-$4G_1$ the situation is more complicated due to the relatively large difference between the ΔH_{cal} value and the ΔH_{vH} value. This difference indicates that the reaction deviates considerably from a two-state process. The usual explanation for the observation that the ΔH_{cal} value is smaller than the corresponding ΔH_{vH} value is that either an aggregation or a competing bimolecular process occurs. Even though both possibilities are theoretically possible very little experimental evidence for either exists. In that respect it is noteworthy that Plum et al (1990) observed a similar discrepancy between the ΔH_{cal} value and the ΔH_{vH} value for the intermolecular triple helix to duplex transition. Interestingly the recent investigation of the single stranded $d(A\text{-}G)_{10}$ selfstructure at low pH, thought to be formed by ionic interactions between protonated adenines and the negatively charged phosphate backbone, detect a similar discrepancy between the ΔH_{cal} value and the ΔH_{vH} value (Dolinnaya et al. (1993)). From these observations it is possible to conclude that the discrepancies between ΔH_{cal} and ΔH_{vH} may originate from ionic interaction

of protonated bases and phosphate groups absent in conventional DNA structures. The intramolecular triple helix to hairpin transition of the various oligonucleotides show large local sequence specific differences in ΔH_{vH} (**Table VI C-E**). Despite these differences in ΔH_{vH} the corresponding differences in T_m are small and the T_m varies in the opposite direction. This can only be understood, if we assume that there is a compensating entropy difference. In the absence of any obvious entropy differences, we must conclude that the observed discrepancies result from ionic interaction of the protonated cytosines. Because of this it is not possible to determine reliable thermodynamic data for the triple helix to hairpin transition based solely on spectroscopic measurements.

Entropy

Tinoco (DeVoe & Tinoco (1962)) has outlined a procedure to evaluate the upper limit of the transition entropy per nucleotide based on the Bolzmann formalism that relates the probability of accessible states to the entropy of the system. It has been shown experimentally that there is no sequence dependency of the entropy change per nucleotide, supporting the assumption that the entropy change can be attributed to the conformational entropy of the sugar phosphate backbones. The entropy of the conformational transition can be evaluated from the enthalpy of this transition and its melting temperature by equation (**4.2**):

$$\Delta S_x = \frac{\Delta H_x}{T_m} \qquad \textbf{(4.2)}$$

The ΔS_{cal} and ΔS_{vH} values calculated for the hairpin to coil transition of ITS-4G$_1$ (210.0 cal/M K and 212.2 cal/M K respectively) are typical for the denaturation of a double helix of 10 bases in length. The ratio of the entropy of the hairpin to coil

transition to the entropy of the intramolecular triple helix to coil transition is 2/3, reflecting the difference in the number of backbone strands. Accordingly the ratio of the calorimetric entropy of the triple helix to hairpin transition at pH 6.7 to that of the hairpin to coil transition is 1/2. This again reflects the number of backbone strands partaking in the conformational transition. Because of the overestimation of the van't Hoff enthalpy for the triple helix to hairpin transition the van't Hoff entropy also overestimates the corresponding calorimetric entropy. The van't Hoff entropy values of the different conformational transitions of the other oligonucleotides follow a similar trend (**Table VI A-F**).

Summary

1) Thermal denaturation experiments have shown that each 38 mer oligonucleotide can undergo three different conformational transition, depending on the environmental conditions.

2) These different conformational transitions have been identified as the hairpin ↔ coil transition, the intramolecular triple helix ↔ hairpin transition and the intramolecular triple helix ↔ coil transition.

3) The thermal stability of each of these transitions as well as the environmental conditions at which each of the conformations occur depends strongly on the local and global composition.

4) The comparison of the impact of the environment (pH & ionic strength) on the stability of each of the three conformers for the different 38 mer oligonucleotides can help to asses the local and global contributions of base composition on the stability of the particular conformation. Especially the number and position of protons in the third strand is obviously the determining factor for the variations. The effect of T_m, ionic strength and pH on each of the three conformational transition will be discussed in detail below.

DISCUSSION

THE HAIRPIN TO COIL TRANSITION

5.1.1) The Effect of Local and Global Composition on the Melting Temperature

The T_m Changes Linearly with Base Composition

Table VIII lists the melting temperature (T_m) of the hairpin to coil transition at pH 8.0 and pH 6.75 respectively (100 mM Na^+) of the six oligonucleotides. The T_m increases linearly as a function of the percentage of GC base-pairs (%GC) in the hairpin stem **(Figure 5.1.1)**, with the exception of the T_m of ITS-ATT. Mamur and Doty have shown this linear relationship as early as 1962 for genomic DNA. The contribution to the linear increase originates both from stacking interactions and from the number of hydrogen bonds between AT and GC base-pairs. (e.g. Klump (1988), Breslauer (1986)). As expected local variations in sequence at constant composition (%GC base-pairs) do not significantly effect the T_m of the hairpin to coil transition. In general local variations in sequence at constant composition only change the nearest neighbor stacking interactions. But, due to sequence constraints in the 38mer oligonucleotides, only minor differences in the nearest neighbor interactions of ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ can occur (**Table VIII**). The linear dependence of T_m on %GC base-pairs can be described by the equation:

$$T_m = 0.46*(\%GC) + 44.01°C \quad \text{at pH } 8.0 \qquad (5.1.1)$$

and

$$T_m = 0.44*(\%GC) + 45.50°C \quad \text{at pH } 6.75 \qquad (5.1.2)$$

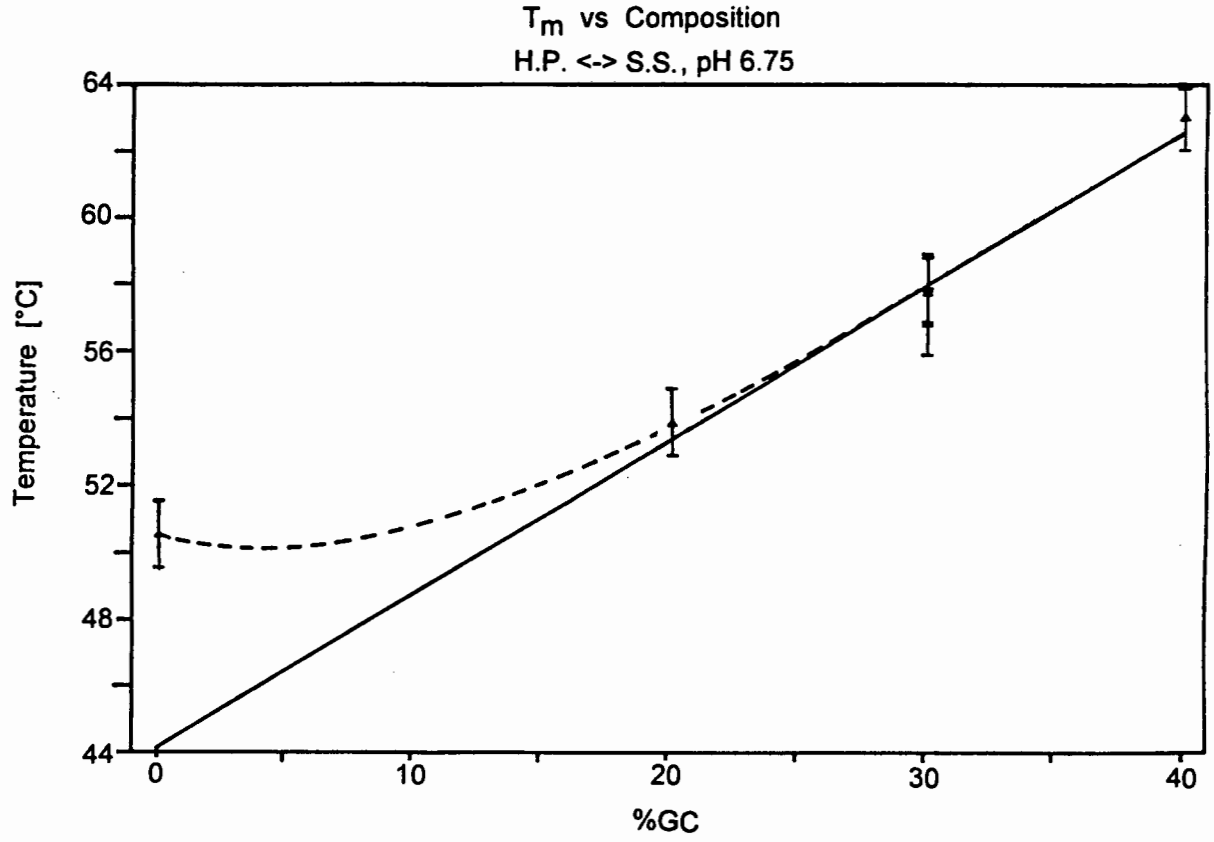


Figure 5.1.1 The effect of base composition on the melting temperature of the hairpin to coil transition at 100 mM Na^+ pH 6.75

where the numerical values for the slope and intercept are obtained by linear least square fitting of the T_m values of ITS-2G_3, ITS-3G_2 and ITS-4G_1 to the %GC base-pairs. These equations compare well with the linear equation derived by Benight and coworkers (Doktycz et al (1992)) for a series of oligonucleotide dumbbells with different GC content:

$$T_m = 0.44*(\%GC) + 66.52°C \quad \text{in 115 mM } Na^+, \text{ pH 6.8} \qquad (5.1.3)$$

The intercept part of the linear equations reflects differences in salt concentration and end effects such as hairpin loops or tailing ends, while the slope should only depend on the %GC content.

TABLE VIII : The Hairpin to Coil Transition

	T_m pH 8.0 [°C]	T_m pH 6.75 [°C]	ΔH_{vH} pH 8.0 [kcal/M]	$\delta T_m/\delta \log[Na^+]$ pH 8.0 [°C]	$\delta T_m/\delta \log[Na^+]$ pH 6.75 [°C]	$\delta T_m/\delta$(pH) 150 mM Na^+ [°C]
ITS-ATT pH <6.0						-2.42 ± 0.37
ITS-ATT pH >6.0	50.5 ± 1.0	49.8 ± 1.0	67.4 ± 4.7	16.39 ± 0.56	17.57 ± 0.56	-0.34 ± 0.11
ITS-$2G_3$	53.8 ± 1.0	54.2 ± 1.0	71.5 ± 3.4	14.69 ± 0.64	16.22 ± 0.57	-0.69 ± 0.19
ITS-$3G_0$	57.8 ± 0.2	58.3 ± 0.2	68.8 ± 1.8	15.32 ± 0.47	14.06 ± 0.67	-1.28 ± 0.19
ITS-$3G_1$	57.7 ± 1.0	59.5 ± 1.0	75.9 ± 4.7	12.93 ± 0.52	14.07 ± 0.19	-0.94 ± 0.12
ITS-$3G_2$	56.8 ± 1.0	58.9 ± 1.0	67.8 ± 1.9	13.26 ± 0.71	13.94 ± 0.74	-1.47 ± 0.14
ITS-$4G_1$	63.0 ± 1.0	63.0 ± 1.0	71.7 ± 4.2	12.00 ± 0.71	12.24 ± 0.71	-1.34 ± 0.12

Oligonucleotides contain a limited number of base-pairs resulting in a significant nonrandom distribution of nearest neighbor base stacking interactions. This may influences both the intercept and the slope of the above equation. Perhaps this is the reason for the small difference in slope obtained for the oligonucleotide hairpin to coil transition compared to that derived for the melting temperatures of genomic DNA (Marmur & Doty (1962). Mamur & Doty (1962) derived the equation:

$$T_m = 0.41 * (\%GC) + 69.3°C \qquad \textbf{(5.1.4)}$$

to describe the dependence of T_m on the %GC base-pairs. The numerical results of the equations **(5.1.1)** and **(5.1.2)** are supported by the results of Benight et al (1992) and Mamur & Doty (1962), even though they were derived from relatively small database only.

Homo A • Homo T Regions form an Altered DNA Structure - A-Tract DNA

What is the cause then for the higher than expected T_m observed for the oligonucleotide ITS-ATT? Applying equations **(5.1.1)** or **(5.1.2)** to the T_m values of ITS-ATT, containing only AT base-pairs, at pH 8.0 or pH 6.75 results in a apparent GC content of 10% (pH 6.75) and 14% (pH 8.0) respectively. A similar increase in stability has been observed for poly d(A) • poly d(T) previously (Klump (1988), Klump (1985)). Fitting the T_m values of poly d(A) • poly d(T) into a linear equation results in an apparent GC content of about 20 %. Since this is patently not true, the increased T_m value for both poly d(A) • poly d(T) and ITS-ATT must be an inherent property of the secondary structure of the homo d(A) • homo d(T) sequence. Based on the x-ray crystal structure analysis of oligonucleotides containing short homo d(A) • homo d(T) inserts Nelson et al (1987) and Coll et al (1987) have proposed the formation of a bifurcated hydrogen-bond system in the major groove of the helix between the adenine N6H group of one AT

base-pair and the thymine O4 atom of an adjacent AT base-pair. This predicted bifurcated hydrogen bond system is made possible by a high degree of propeller twist for the AT base-pair maximizing the stacking interaction between adjacent adenine bases. A high degree of propeller twist between the adenine bases is also found in fiber diffraction studies of poly d(A) • poly d(T) giving rise to a structure called B_h DNA (Lipanov et al (1990)). Both the increased stacking interactions of the adenine bases and the presence of the additional bifurcated hydrogen bond system can contribute to the increased stability of both poly d(A) • poly d(T) and ITS-ATT.

The Effect of A-Tract DNA on the T_m

Introducing GC base-pairs into an AT run causes an interruption of the bifurcated hydrogen bond system, since the GC base-pairs do not contain the appropriate hydrogen-bond donor and acceptor groups in the major groove. Also, since GC base-pairs are unable to exhibit the same degree of propeller twist, a reduction in the stacking interactions will occur at the site of the insertion of a GC base-pair (Nelson et al (1987)). Thus insertion of GC base-pairs should reduce the stability of the homo d(A) • homo d(T) tract. At the same time introducing GC base-pairs into an oligonucleotide will increase the melting temperature because of the increased number of conventional (as opposed to bifurcated) hydrogen-bonds. It is therefor expected that the T_m of the hairpins will represent a trade off between these two properties. At low %GC content the T_m will be dominated by the excess stability of the A-tract while at higher %GC content the conventional hydrogen-bonds will be mainly responsible for the T_m. The stippled line in **Figure 5.1.1** indicates the deviation from linearity of the T_m at decreasing %GC content. In principle it should be possible to more accurately describe the change in melting temperature with the number of GC base-pairs in the hairpins by the following equation:

$$T_m = [SLOPE * (\%GC) + CONSTANT] + [f(A \bullet T)] \qquad (5.1.5)$$

where the function f(A•T) describes the increased melting temperature caused by the homo d(A) • homo d(T) tract. It can be expected that f(A•T) is not a trivial function of the number of AT base-pairs in a row, since a minimum length of consecutive homo d(A) • homo d(T) base-pairs will be required to nucleate the A-tract conformation. NMR studies by Nadeau & Crothers (1989) seem to indicate that a minimum length of 7 consecutive AT base-pairs are required before the central base-pair will show the same imino proton resonance as adenine in poly d(A) • poly d(T). A insertion of a GC base-pair into an AT run will always create a junction between a B-DNA like conformation and an A-track conformation. It is likely that this junction region depends on the orientation of the inserted interrupting base-pair (e.g. Nelson et al (1987) assigned a 3' GC base-pair the status of an "honorary AT base-pair", since it seemed to interrupted the A-tract structure less than a 5'CG base-pair). At present there is insufficient experimental data to allow a quantification of the function f(A•T). In any case A-tract formation and the effect of interruptions by a variable number of GC base-pairs on the A-trackt will provide the first evidence for long range effects beyond the nearest neighbor stacking interactions in DNA. To a first approximation one can assume that the junction consists of three AT base-pairs on either side of the insert. It can be estimated that the term f(A•T) in equation (**5.1.5**) will be small or zero for all oligonucleotides except ITS-ATT. f(A•T) will reflect only the deviation of the T_m from the T_m expected for an oligonucleotide of AT base-pairs in random orientation. Hence a linear function of T_m versus %GC as given in equation (**5.1.1**) or (**5.1.2**) should adequately describe the T_m's of all oligonucleotide hairpins except ITS-ATT. As shown in Figure **5.1.1** this in fact the case.

CD-Spectroscopy Supports A-Tract Formation

Direct experimental support for the occurrence of A-tract regions (B_h-DNA) comes from the CD spectra of the different oligonucleotide hairpins. The shoulder at 260 nm in the more AT rich oligonucleotides is a characteristic feature of the CD-spectrum of the homopolymer poly d(A) • poly d(T) (e.g. Shan et al (1990)) and of the $d(A)_{10}$•$d(C)_4$•$d(T)_{10}$ hairpin (van de Sande et al. (1988)). It has not been observed for any other polynucleotide sequence, not even for the alternating copolymer poly d(AT) • poly d(AT) (Shan et al (1990)). It is therefor reasonable to assume that this characteristic CD-spectral band is due to the slightly altered helical structure and base tilt of the B_h-DNA conformation. It is interesting to note that the CD-spectrum of single stranded poly d(A) (Johnson (1990)) exhibits similar features, indicating that the base tilt in the stacked single stranded poly d(A) may be similar to that of the double stranded polymer. The disappearance of the shoulder in the more GC rich oligonucleotides ITS-$3G_1$, ITS-$3G_2$ and ITS-$4G_1$ indicates a corresponding disappearance of the B_h type character in these oligonucleotides.

5.1.2) The Effect of Counterion Concentration on the Melting Temperature

The $d(T_m)/d(\log[Na^+])$ Changes Linearly with Base Composition

Table VIII lists the change in T_m as a function of the logarithm of the Na^+ concentration ($d(T_m)/d(\log[Na^+])$) for the six oligonucleotides under investigation. The $d(T_m)/d(\log[Na^+])$ values vary according to the base composition of the hairpins stem. In **Figure 5.1.2** the $d(T_m)/d(\log[Na^+])$ is shown as a function of the overall composition (%GC) of the hairpin stem.

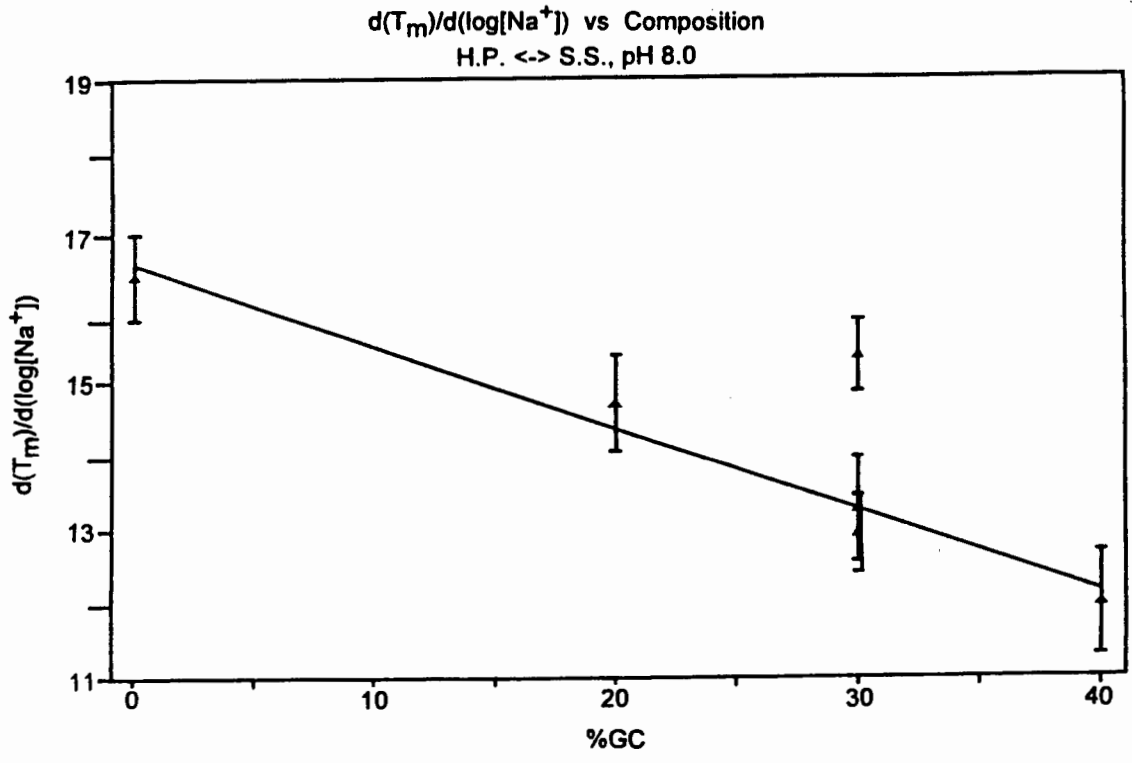


Figure 5.1.2 The effect of base composition on the degree of counterion release ($d(T_m)/d(\log[Na^+])$) of the hairpin to coil transition at pH 8.0

It is intriguing to note that the $d(T_m)/d(\log[Na^+])$ decreases linearly with an increase in %GC. This is somewhat unexpected, considering that the amount of counterions released on melting should only depended on the difference in charge density between the helix and the coil states, which as a first approximation should be independent of sequence or base composition (Manning (1978)). However the counterion condensation theory was developed for ideal conditions of polyions of infinite length and point charges only. In order to asses the importance of this linear change in $d(T_m)/d(\log[Na^+])$ with hairpin composition, it is necessary to discuss the effect of oligoions with complex structure (hairpin stem and loop as well as single stranded 3'overhanging terminus) on the experimental $d(T_m)/d(\log[Na^+])$ values.

The Impact of End Effects and Hairpin Loops on the $d(T_m)/d(\log[Na^+])$

Even though the oligonucleotides under investigation here are of length **N** = 38 nucleotides, only 24 nucleotides are involved in the formation of the hairpin helices. Thus the hairpin helices should just fall in the range were a linear behavior of $d(T_m)/d(\log[Na^+])$ with oligonucleotide length is expected (Olmsted et al (1989)). It is assumed that the end effects in all oligonucleotides will contribute equally to the observed $d(T_m)/d(\log[Na^+])$ value of the hairpin to coil transition. Similarly it is assumed that the loops contribute equally to the $d(T_m)/d(\log[Na^+])$ value, since in all oligonucleotides the hairpin loop is formed by four cytosine bases and closed by an AT base-pair.

The Impact of the Single Stranded Overhanging 3'Terminus on the $d(T_m)/d(\log[Na^+])$

The contribution of the dangling 3' single stranded tail to the ion condensation of the hairpin helix is more difficult to estimate. It is very likely that the exact nature of the ion condensation onto this tail is a function of the distance from the terminal base-pair of the hairpin helix. The overall effect of the single stranded tail will be to reduce the end effect parameter introduced by Record and Lohman (1978), effectively mimicking a hairpin with a longer stem region. This will be expressed in additional nucleotides included into the helix stem (apparent length $\mathbf{N_{app}}$). Experimental confirmation of this assumption comes from the comparison of the $d(T_m)/d(\log[Na^+])$ value of the hairpin $A_{10}C_4T_{10}$ (HP-AT, data not shown) with that of the oligonucleotide ITS-ATT. For the hairpin HP-AT the $d(T_m)/d(\log[Na^+]) = 13.89 \pm 0.36°C$, which is smaller than the $d(T_m)/d(\log[Na^+]) = 16.39 \pm 0.56°C$ of ITS-ATT (pH 8.0).

Equation **(2.10b)** as shown above allows a semiempirical calculation of the reduction in slope caused by the end effects in oligoions. The $d(T_m)/d(\log[Na^+])$ value of poly d(A) • poly d(T) was calculated to be $S_\infty = 18.79 \pm 0.33°C$, based on a data set compiled by Guschlbauer (1990). The calculated S_N for the 24mer hairpin (HP-AT) is $S_{Ncalc} = 11.84°C$. This is 2°C smaller than the experimentally observed S_N value. The differences between the calculated and experimental S_N value must reflect either differences in the magnitude of the end effect parameter **r * (1-f)** or differences due to the sequence in the hairpin loop. The end effect parameter **r * (1-f)** determined by Record & Lohman (1978) is based on the behavior of $(AT)_N$ hairpins, which may differ from that of other sequences. Differences in loop conformation may also contribute to differences in the overall charge density and therefor to the end effect parameter.

Assuming a constant overestimation of $d(T_m)/d(\log[Na^+])$ for all homo d(A) • homo d(T) hairpins and correcting the S_{Nexp} value of ITS-ATT appropriately allows the calculation of an apparent length N_{app} of the oligonucleotide ITS-ATT. N_{app} calculated thus is $N_{app} = 32.89 \approx 33$ nucleotides. Thus the contributions of the 14 nucleotide long single stranded 3'dangeling tail to the electrostatic behavior of the hairpin will be the same as those expected for a hairpin with a stem which contains 4 to 5 extra base-pairs. It is encouraging that the calculated number of base-pairs N_{app} is reasonable. It is reasonable to assume that this consideration holds for all other oligonucleotide hairpins investigated. It is however not possible to validate this assumption, due to the lack of $d(T_m)/d(\log[Na^+])$ data of corresponding polynucleotides.

The $d(T_m)/d(\log[Na^+])$ is Proportional to the Manning Parameter $\eta/2$

Since end effects must be equal any differences in the $d(T_m)/d(\log[Na^+])$ values of the different oligonucleotides must be caused by effects other than the oligoion character of the hairpins. According to equation **(2.5c)** the $d(T_m)/d(\log[Na^+])$ value is proportional to the Manning parameter $\eta/2$, with the proportionality constant given by the term $\mathbf{(2.3*R*T_m^2/\Delta Hobs)}$. Both the T_m and the ΔH_{obs} increase with increasing GC content. Hence it is conceivable that the differences in slope are due to composition and sequence dependent changes in the term $\mathbf{(2.3*R*T^{m_2}/\Delta Hobs)}$. The change of T_m with base composition has already been discussed in the previous chapter. Since the ΔH_{obs} must be model independent, only the calorimetrically determined ΔH_{cal} can be substituted for ΔH_{obs} in equation **(2.5c)**. However since the ΔH_{cal} and ΔH_{vH} of ITS-$4G_1$ are in close agreement with each other and since all other sequences obey the same folding pattern as ITS-$4G_1$, it is acceptable to substitute the ΔH_{vH} value for ΔH_{cal} in all cases. Because the sequence dependent differences in ΔH_{vH} do not exceed the experimental error it is more appropriate to substitute ΔH_{comp} **(Table VII)** into

equation **(2.5c)**. The sequence dependent change in ΔH_{comp} and the experimentally determined T_m^2 compensate each other, such that the term $(2.3*R*T_m^2/\Delta Hobs)$ is constant, i.e. independent of the GC content or sequence. Record et al (1978) came to the same conclusion from surveying the literature on the calorimetry of nucleic acids of natural sources. They found that the term $(2.3*R*T^{m_2}/\Delta Hobs)$ is independent of temperature, base composition, chain length (for chains longer than the cooperative melting unit) and ionic strength. This can only be, if the entropy of the denaturation is practically sequence independent. Klump (1988) showed that the entropy of melting changes marginally with the melting temperature, however the change is small enough that it will not effect the validity of the above statement. Consequently any change in slope must be directly proportional to composition dependent changes in the Manning parameter η.

The Charge Density of the Helix Depends on the Sequence

Sequence dependent changes in the Manning parameter η must reflect sequence dependent changes in the linear charge density **b** of either the hairpin ($\mathbf{b}_h$) or the coil ($\mathbf{b}_c$) state. Currently no independent determination of either $\mathbf{b}_h$ or $\mathbf{b}_c$ for the hairpin helices is available and it is therefor impossible to quantify this difference precisely. Conventionally it is assumed that the axial charge density of the helical state $\mathbf{b}_h$ as a first approximation can be described by the average helix parameters of a B-DNA (i.e. 2 negative charges for every 3.4 Å), while the axial charge density of the coil state $\mathbf{b}_c$ is assumed to vary. Small composition dependent differences in the Manning parameter η of genomic DNA have indeed been observed (Owen et al (1969), Frank-Kamenetskii (1971)) and have been assigned to some residual degree of stacking in the single stranded coil state. (Blake & Haydock (1979)). Base stacking interactions in single stranded polynucleotides are temperature dependent (Brahms (1964), Fresco & Klemperer

(1959)). If one assumes that residual stacking interactions are responsible for differences in $\mathbf{b_c}$, then the hairpin that denatures at the lowest temperature should retain the largest degree of single stranded stacking and hence show the least change in axial charge distance reflected in the smallest smallest $d(T_m)/d(\log[Na^+])$. The opposite is observed experimentally. It seems therefor more appropriate to assume that the axial charge distance of all coil states is identical at elevated temperatures. Consequently differences in the Manning parameter η have to be due either to differences in the helical axial charge density $\mathbf{b_h}$, or to some other reason.

The Effect of Hydration on the $d(T_m)/d(\log[Na^+])$

Implicit in the derivation of the polyelectrolyte theory is the assumption that the change in hydration on denaturation of the DNA is sequence, pH and temperature independent and therefor need not be accounted for. However it is now known that the degree of solvation of A•T base-pairs is different from that of G•C base-pairs (Schneider et al (1992)). It can be assumed that the degree of hydration increases on denaturation. It is likely that the number of water molecules surrounding a single stranded adenine base is very similar to the number of water molecules surrounding a guanine base, due to their similar shape and hydrophobicity. It can also be assumed that the degree of solvation of the pyrimidine bases are nearly identical in the coil state. Hence it is likely that there are composition dependent differences in the change of hydration on denaturation, which can give rise to the composition dependent changes in $d(T_m)/d(\log[Na^+])$.

Composition Dependent Changes in $d(T_m)/d(\log[Na^+])$ are due to Differential Hydration and Change in Helical Charge Density

The observed linear dependence of $d(T_m)/d(\log[Na^+])$ on the percentage of GC base-pairs for genomic DNA has been fitted to the empirical equation (Frank-Kamenetskii (1970)):

$$d(T_m)/d(\log[Na^+]) = -0.070 * (\%GC) + 18.3°C \qquad \textbf{(5.1.6a)}$$

Similarly the salt dependence of the T_m of different melting regions of a lambda DNA have been fitted to the empirical equation (Blake and Haydock (1979)):

$$d(T_m)/d(\log[Na^+]) = -0.067 * (\%GC) + 19.96°C \qquad \textbf{(5.1.6b)}$$

It is noticeable that the observed slopes for genomic DNA and for the Lambda DNA are very similar. In contrast to this the slope of the linear equation modeling the change of $d(T_m)/d(\log[Na^+])$ with %GC base-pairs for the hairpin to coil transition shown in **Figure 5.1.2** is much more pronounced. Linear least square fitting of the change in $d(T_m)/d(\log[Na^+])$ with %GC results in equation (**5.1.7a**) :

$$d(T_m)/d(\log[Na^+]) = -0.11 * (\%GC) + 16.55°C \qquad \textbf{(5.1.7a)}$$

at pH 8.0 and equation (**5.1.7b**):

$$d(T_m)/d(\log[Na^+]) = -0.13 * (\%GC) + 18.00°C \qquad \textbf{(5.1.7b)}$$

at pH 6.75.

It is assumed that the change in $d(T_m)/d(\log[Na^+])$ observed by Frank-Kamenetskii (1970) and Blake & Haydock (1979) is mainly due to differences in the hydration of G•C and A•T base-pairs. The much steeper slope observed for the oligonucleotide hairpin to coil transitions may reflect contributions from end effects, change in hydration,

and sequence dependent variations in secondary structure. Since the oligonucleotides consist of only a fraction of all the statistically possible sequences of the appropriate length and composition, composition dependent hydration effects will be more pronounced.

As was shown by x-ray crystallography the strong propeller twist and the bifurcated hydrogen bond system of A•T base-pairs in a homo d(A) • homo d(T) helix cause the helical rise of B_h-DNA to decrease compared to the standard 3.4Å per base-pair for an average B-DNA helix (Nelson et al (1987) , Coll et al (1987)). Homo d(A) • homo d(T) tracts are thus characterized by a reduced spacing between adjacent charges and hence a lower $\mathbf{b}_h$ value. G•C base-pairs, which interrupt the B_h-DNA structure cause a change in the helical rise and the axial charge density at these sites. Consequently the average linear charge density parameter $\mathbf{b}_{h,av}$ will increase with the introduction of GC base-pairs into the oligonucleotides. Consequently it is expected that the charge density of the hairpin helix of ITS-ATT and the change upon melting reflected in the $d(T_m)/d(\log[Na^+])$ value will be the largest for all the oligonucleotides while the corresponding value will be smallest for ITS-$4G_1$. This is indeed observed experimentally. It is likely that all the effects stated above will act in concert. The observed composition dependent differences in $d(T_m)/d(\log[Na^+])$ may thus reflect subtle structural differences in the backbone conformation as well as differences in the hydration of the different oligonucleotide hairpins.

5.1.3) The Effect of pH on the Melting Temperature

The $d(T_m)/d(pH)$ Changes Marginally with Base Composition

Table VIII lists the change in melting temperature T_m with change in pH of the solution ($d(T_m)/d(pH)$) for the hairpin to coil transition. The hairpin to coil transition for the different oligonucleotides can only be monitored at pH values larger than pH 6.0, because at lower pH values the triple helix becomes the dominant structure. The exemption to this is the oligonucleotide ITS-ATT, for which the hairpin to coil transition can be detected throughout the pH range investigated. In the pH range between pH 6.0 and pH 9.0 the T_m of the hairpin to coil transition is only marginally changed by a change in pH - hence the slope $d(T_m)/d(pH)$ is close to zero. This behavior of the hairpins is the same as the pH dependent melting behavior of genomic DNA sequences, such as calf thymus DNA (Record (1967)) in this pH range. The lack of change of the T_m with pH in the pH range between pH 6 and pH 9 is not very surprising, since the pka values for the protonation and deprotonation of any of the bases are more than 1 pH unit outside this pH range (Sänger (1984)). The bases in either the hairpin or the coil state will not be protonated, and hence no differential interaction of protons with either the helix or coil is observed.

The small differences observed for the $d(T_m)/d(pH)$ values are probably due to small changes in the ionic strength of the solution brought about by adjusting the pH. There is a gradual change in $d(T_m)/d(pH)$ between the different oligonucleotides from $d(T_m)/d(pH) = -0.34°C$ for ITS-ATT to $d(T_m)/d(pH) = -1.34°C$ for ITS-$4G_1$. The reasons for this apparent trend in $d(T_m)/d(pH)$ with GC content are unknown. The effects observed are indeed very small compared with those observed for the triple helix.

It is not possible to find statistical support at any reasonable confidence level for a slight sequence dependent $d(T_m)/d(pH)$ and hence these differences may be irrelevant.

Formation of a $C \bullet C^+$ Base-pair in the Hairpin Loop at Low pH

Only for the oligonucleotide ITS-ATT is it possible to detect a clear pH dependent change in T_m at low pH values. While in the pH range between pH 9.0 and pH 6.0 the $d(T_m)/d(pH)$ is close to zero as expected, a sudden change in the slope at pH values below pH 6.0 occurs. The $d(T_m)/d(pH)$ value changes to $d(T_m)/d(pH) = -2.42 \pm 0.37°C$. When the pH approaches pH 4.5 the $d(T_m)/d(pH)$ approaches zero again. This observation indicates that this effect must be due to the protonation of cytosines, since only cytosine mono nucleotides have a pka value between pka = 4.2 and pka = 4.5. The only cytosines involved in the hairpin structure of ITS-ATT are those found in the hairpin loop. A $C \bullet C^+$ base-pair can potentially form between cytosines 1 and 4 of the hairpin loop, resulting in the protonation of one cytosine at a pH values above the pka. The magnitude of the $d(T_m)/d(pH)$ value at pH values below pH 6.0 is consistent with the protonation of only one cytosine. The formation of the unusual $C \bullet C^+$ base-pair could provide sufficient free energy to overcome the inherently positive free energy of protonation of a cytosine base 1.5 to 1.8 pH units above its pka value. Corroborating evidence for the formation of the $C \bullet C^+$ base-pair in the hairpin loop comes from the peculiar pattern of P_1-nuclease digestion at pH 5.5 of all oligonucleotides. As was discussed in the result section previously, the observed P_1-nuclease digestion pattern at pH 5.5 can be explained by the formation of a possible $C \bullet C^+$ base-pair between cytosines 1 and 4 of the hairpin loop. At pH 8.0 , where no cytosine protonation occurs, the four cytosines form a partially stacked hairpin loop not recognized by P_1-nuclease as a substrate.

THE INTRAMOLECULAR TRIPLE HELIX TO HAIRPIN TRANSITION

5.2.1) The Effect of Local and Global Composition on the Melting Temperature

The T_m Depends on the Local and Global Composition

At pH 6.75 the UV melting curves show two independent conformational transitions. The transition with the lower melting temperature has been assigned to the triple helix to hairpin transition, while the upper transition represents the melting of the hairpin to the coil state and has already been discussed extensively in the preceding chapter. In this chapter the emphasis is on the intramolecular triple helix to hairpin transition. **Table IX** lists the thermodynamic data for this transition. The change in the T_m as a function of the global/local composition of the third strand (%cytosines (%C)) is shown in **Figure 5.2.1**.It is obvious that the T_m is not only dependent on the %C in the third strand, but also dependent on the exact sequence (c.f. the difference in T_m between ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$). To facilitate the interpretation of the data the change of T_m is discussed is terms of global and local composition *vide infra.*

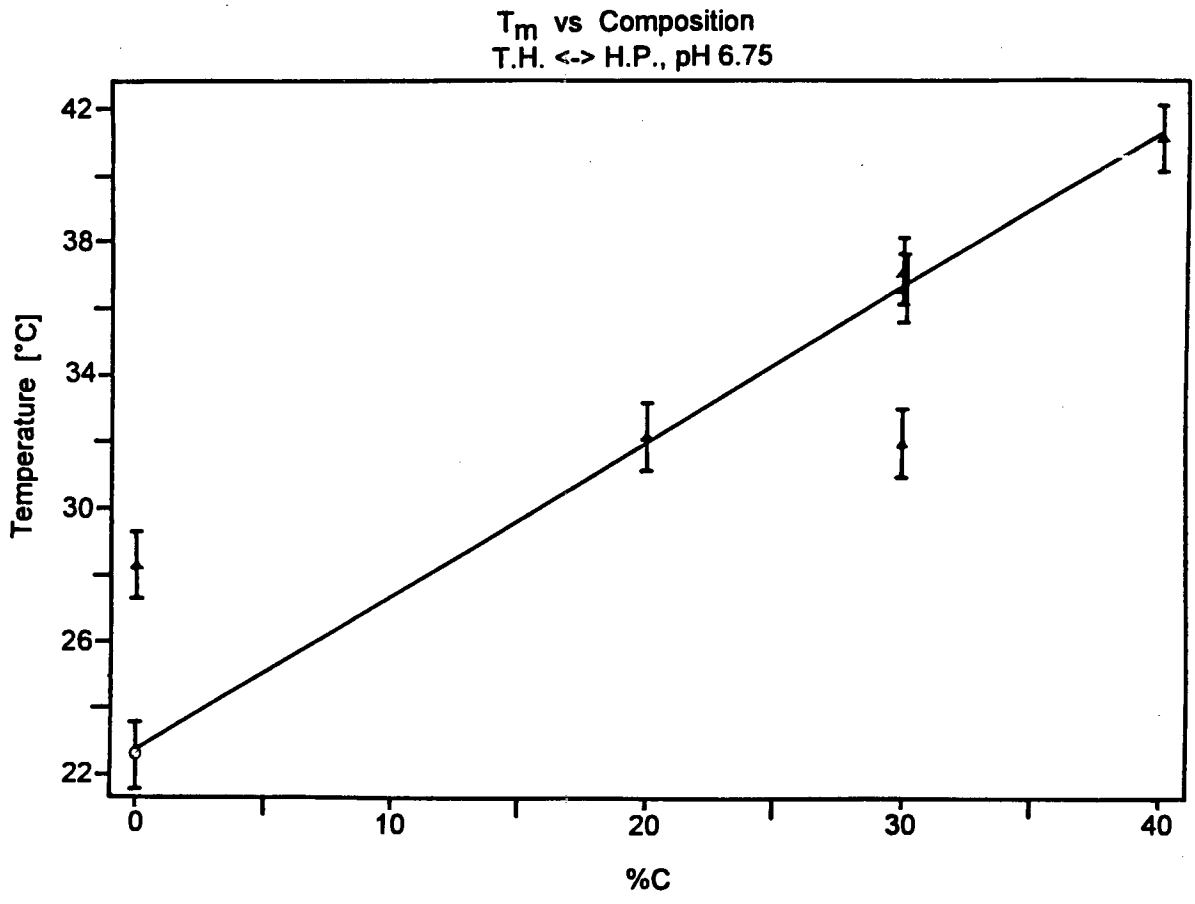


Figure 5.2.1 The effect of sequence composition and base-composition on the melting temperature of the triple helix to hairpin transition at 100 mM Na^+ pH 6.75

TABLE IX: The intramolecular triple helix to hairpin transition

	T_m pH 6.75 TS -> HP [°C]	T_m pH 6.75 HP -> C [°C]	ΔH_{vH} TS -> HP [kcal/M]	$\delta T_m/\delta \log[Na^+]$ TS -> HP < 200 mM Na^+ [°C]	$\delta T_m/\delta \log[Na^+]$ TS -> HP > 400 mM Na^+ [°C]	$\delta T_m/\delta(pH)$ TS -> HP [°C]
ITS-ATT pH 8.0				33.80 ± 0.50		-1.47 ± 0.30
ITS-ATT pH 6.75	28.3 ± 1.0	49.8 ± 1.0	53.2 ± 3.0	29.60 ± 1.49		-6.50 ± 0.25
ITS-$2G_3$	32.1 ± 1.0	54.2 ± 1.0	47.7 ± 2.9	11.46 ± 1.25	20.66 ± 1.38	-14.1 ± 0.91
ITS-$3G_0$	31.9 ± 0.2	58.3 ± 0.2	77.5 ± 3.4	4.49 ± 0.16	15.50 ± 1.03	-17.3 ± 0.73
ITS-$3G_1$	36.5 ± 1.0	59.5 ± 1.0	70.3 ± 3.7	7.60 ± 0.56	17.75 ± 0.83	-17.69± 0.31
ITS-$3G_2$	37.0 ± 1.0	58.9 ± 1.0	52.3 ± 3.8	8.04 ± 0.68	14.43 ± 2.40	-20.7 ± 0.59
ITS-$4G_1$	41.0 ± 1.0	63.0 ± 1.0	58.7 ± 4.2	2.42 ± 0.25		-26.0 ± 1.15

5.2.1.1) **The Impact of Global Base Composition on the T_m**

Sequence Effects are due to the Relative Position of Cytosines in the Third Strand

In order to determine the impact of the global composition on the T_m of the triple helix to coil transition it is essential to consider only oligonucleotides for which local sequence specific effects can be excluded. Since a change in base composition of the intramolecular triple helix comes about by the replacement of T•A•T triplets in a T•A•T triple helix by C^+•G•C triplets, local and global composition specific effects will be assigned to the introduction of cytosines only. If the cytosine bases are sufficiently spaced by intervening thymine bases, interactions between neighboring cytosines can be regarded as negligible and only global compositional effects will be observed. Comparing the T_m's of ITS-3G_0, ITS-3G_1 and ITS-3G_2 it seems that there is a strong local sequence specific effect, which is most pronounced, if adjacent cytosines are not separated by any intervening thymines. A minimum of one thymine must separate consecutive cytosines in order to minimize local sequence specific effects. Better still is the separation by two or more thymines. Consequently the discussion of global compositional effects on the T_m shall be confined to the oligonucleotides ITS-ATT, ITS-2G_3 and ITS-3G_2, were at least two thymines separate consecutive cytosines. The oligonucleotide ITS-4G_1 has been included in the discussion in order to span a larger range of %C, even though in this oligonucleotide the cytosines are only separated by a single thymine.

The T_m Changes Linearly with Overall Base Composition

Figure 5.2.2 shows the change in T_m with change in %C for the intramolecular triple helix to hairpin transition of the four oligonucleotides. The T_m changes as a perfectly linear function of the percent third strand cytosines (%C), provided that the T_m of the triplex to hairpin transition of ITS-ATT at pH 8.0 is plotted rather than that at pH 6.75.

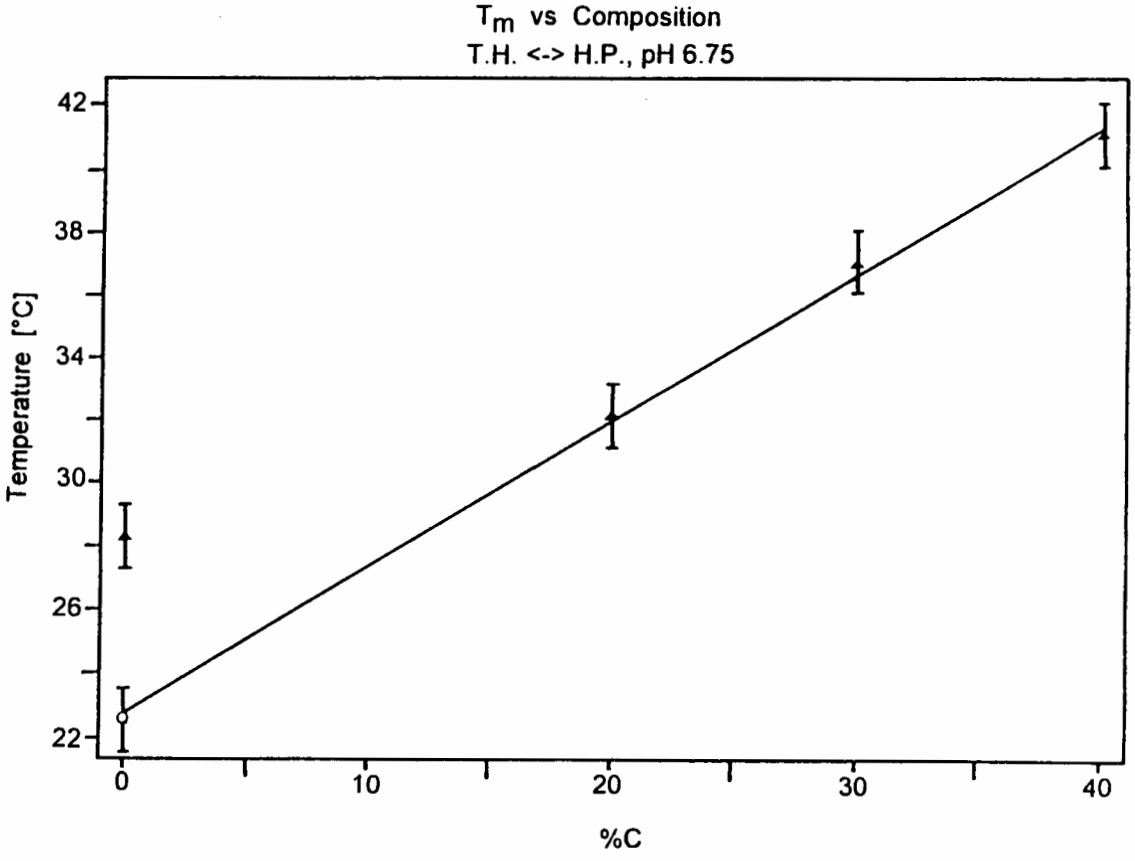


Figure 5.2.2 The effect of global base composition on the melting temperature of the Triple helix to hairpin transition at 100 mM Na^+ pH 6.75

Due to sequence constraints imposed by the homo d(A) • homo d(T) • homo d(T) stem region, ITS-ATT is the only oligonucleotide containing a four membered cytosine (C_4) loop connecting the hairpin to the third strand. In all other oligonucleotides the triple helix loop is made up from four thymine bases. As will be discussed below, one of the four cytosines in this loop can become protonated below pH 7.0 resulting in the formation of a $C \bullet C^+$ base-pair. In the previous chapter it has already been shown that the C_4 hairpin loop can become partially protonated at pH values below pH 6.0 and that this leads to the formation of the unusual $C \bullet C^+$ base-pair between cytosines 1 and 4 of the loop. There is no reason that the same partial protonation of the loop cytosines and the formation of the unusual $C \bullet C^+$ base-pair will not also occur in the intramolecular triple helix at somewhat higher pH values. The formation of the unusual $C \bullet C^+$ base-pair will result in an pH dependent increase in the T_m of the intramolecular triple helix to hairpin transition, even though there is no cytosine included in the third strand.

The change in T_m with %cytosines at pH 6.75 (100 mM Na^+) can be fitted by linear least square fitting to the linear equation:

$$T_m = 0.46 * (\%C) + 22.73\ ^\circ C \qquad \textbf{(5.2.1)}$$

No comparable data set is available in the current literature.

In analogy to the hairpin to coil transition it can be expected that the intercept in equation **(5.2.1)** reflects the influence of environmental conditions, configurational entropy and end effects, and that the slope is only dependent on the base composition. It is somewhat surprising that this slope is of the same order of magnitude as the slope for the hairpin to coil transition.

The Protonation of Third Strand Cytosines Causes the Increase in T_m

What is the underlying physical principle for the increase in T_m with increased %C and why is the slope so similar to that of the hairpin to coil transition? For the hairpin to coil transition as shown above the difference in T_m with base composition is primarily due to the difference in the number of hydrogen bonds of AT and GC base-pairs. The difference in stacking interactions between different nearest neighbor base-pairs can be considered to be small (Benight (1992)). The models of Hoogsteen base-pairs predict the same number of hydrogen bonds of thymines and cytosines **(Figure 1.1 & 1.2)**, provided the cytosines are protonated. Hence differences in hydrogen bonding cannot be the reason for the composition dependent change in T_m. Little is known about the base-stacking enthalpy of third strand bases. But it is unlikely that differences in stacking interactions can account for the observed large differences in T_m. Is the assumption that the Hoogsteen cytosines are protonated correct? Since the triple helix to hairpin transition is studied at pH 6.75, more than two pH units above the pka of free cytosines, the third strand cytosines can potentially exist in either the protonated (2 Hoogsteen hydrogen-bonds) or unprotonated (1 Hoogsteen hydrogen-bond) form. If the cytosines are unprotonated the number of hydrogen-bonds are reduced. This should result in triple helices of reduced thermal stability. Since the exact opposite is observed, it can be assumed that the third strand cytosines are protonated. The positive free energy of protonation at more than two pH units above the pka of free cytosine has to be compensated for by the free energy of third strand binding. This free energy could conceivably be provided by the second hydrogen bond formed on protonation. If however the free energy gained on formation of the additional hydrogen bond was solely responsible for the protonation of the triple helix, the overall free energy of triple helices with third strand cytosines would be less than that of comparable triple helices consisting of T•A•T base triplets only. Since the T_m is directly proportional to the overall free

energy of the triple helix, the T_m of cytosine rich triple helices should again be lower than that of thymine rich triple helices - again the opposite of the observed behavior.

The protonation of cytosines will not only lead to the formation of the second hydrogen bond in C^+•G•C base triplets, but also introduce a positive charge at the cytosine bases. The introduction of a positive charge in the center of a strongly negatively charged triple helix will reduce the electrostatic repulsions of the three phosphate backbones and hence decrease the repulsive electrostatic free energy component. The compensation between the positive free energy of protonation at pH values above the pka of free cytosines and the negative free energy of formation of additional hydrogen bonds associated with the negative free energy due to reduced electrostatic repulsions between the backbones will be responsible for the overall increase in T_m observed in triple helices with higher cytosine content. The result that the slope in equation **(5.2.1)** is almost identical to the corresponding slope of the hairpin to coil transition hence must be due to the fortuitous compensation of the different forces responsible for triple helix stability. It is obvious that at different conditions of pH and ionic strength the numerical results for equation **(5.2.1)** will be different. This can indeed be observed for the oligonucleotides investigated here. Figure **5.2.3** shows the change in slope of equation **(5.2.1)** with change in pH extrapolated to low pH. While the slope of a change of T_m with %C for the hairpin to coil transition is relatively independent of the environmental conditions over a wide range of these conditions, the slope of the triple helix to hairpin transition will depend strongly on the environmental conditions of pH and ionic strength.

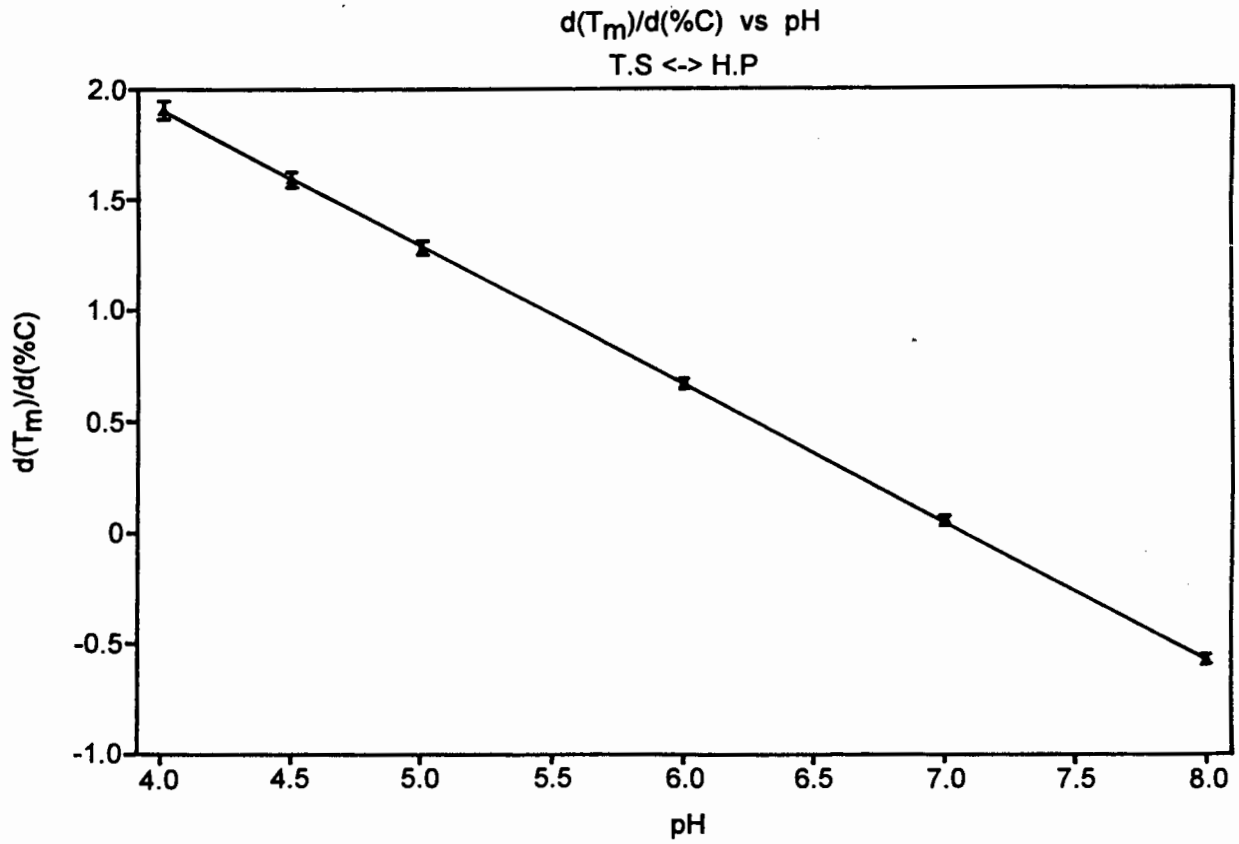


Figure 5.2.3 Plot of the slope of equation **(5.2.1)** $\{d(T_m)/d(\%C)\}$ versus pH at 150 mM Na^+

5.2.1.2) **The Impact of Local Composition on the T_m**

Local Composition Effects Arise from the Interaction of Neighboring Cytosines

The three oligonucleotides ITS-3G_0, ITS-3G_1 and ITS-3G_2 have the same overall base composition but contain different sequences. The sequences differ by the increasing separation of the central two cytosines due to the insertion of thymine residues. From the melting temperatures tabulated in **Table IX** it can be seen that the T_m is a function of the number of thymines inserted between adjacent cytosines. The T_m of ITS-3G_0 is smallest that of ITS-3G_2 is largest and the T_m of ITS-3G_1 is only marginally smaller than that of ITS-3G_2. It is clear from this that the T_m is not a linear function of the number of intervening thymines (**Figure 5.2.4**).

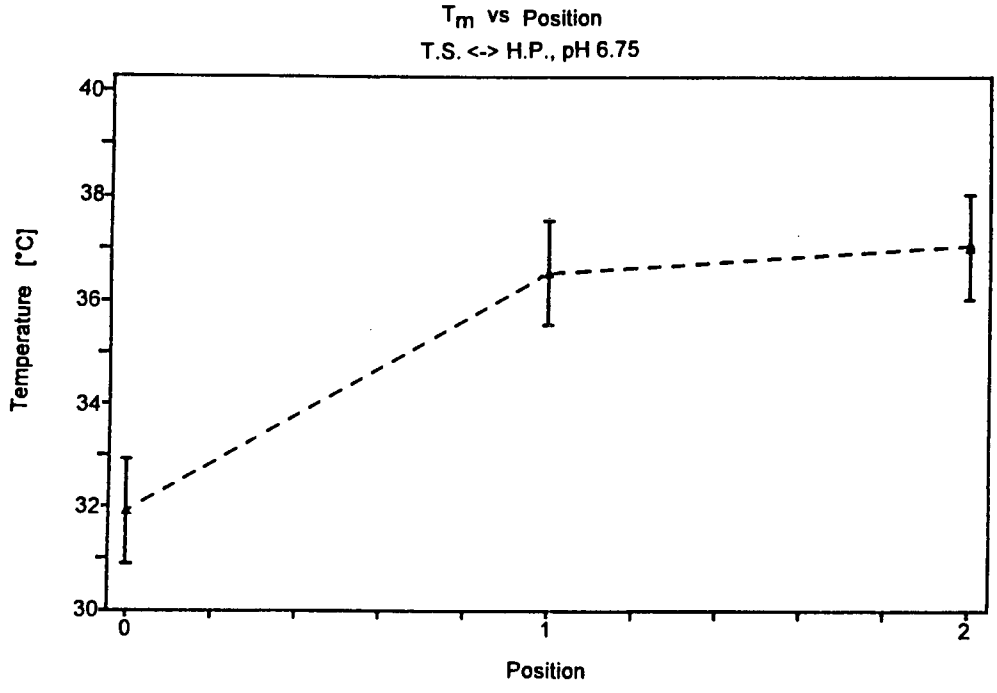


Figure 5.2.4 The effect of sequence composition on the melting temperature of the triple helix to hairpin transition at 100 mM Na^+ pH 6.75

It is likely that the same forces which are responsible for the global effects are also responsible for the local sequence specific change in T_m. The binding of a proton and the associated formation of a second hydrogen bond in a third strand cytosine will only exhibit its influence on the formation of the C^+•G•C base-triplet and should have no effect on the neighboring base triplets. Associated with the protonation is the placing of a positive charge at the cytosines. This charge is responsible for the attractive electrostatic contribution to the overall stability of the triple helix. It is therefor suggested that these electrostatic interactions are responsible for local sequence specific variations in T_m.

The Positive Charges Cause Repulsion between Adjacent Protonated Cytosines

On the one hand the positively charged cytosines cause a decrease in the electrostatic repulsion of the negatively charged triple helix backbone, but on the other hand the positive charges will cause an increased electrostatic repulsion of adjacent third strand cytosines. In other words they act attractive on the phosphates and repulsive on the neighboring cytosines. The latter effect will certainly decrease with increasing distance. The repulsion of the positive charges of adjacent cytosines can be considered to increase the amount of work needed (≈free energy) for a proton to approach a cytosine. This should decrease (more acidic) the pka_{app} for the second cytosine compared to the pka_{app} of the first cytosine. As a consequence this results in a decrease of the overall free energy of triple helix formation compared to triple helices of identical base composition but well separated cytosines and therefor a lower T_m is expected. This is exactly what the experiment shows. The positive charges are partially compensated for by the negative charges located on the phosphate backbone in the plane of the triplet and to a lesser extend along the helix axis. Insertion of thymines between neighboring cytosines brings about a displacement of the positive charges around the helical cylinder and along its axis. It can be expected that the electrostatic repulsion between these two cytosine bases

will diminish very rapidly with increasing number of intervening thymine bases. This is reflected in the observed increase in T_m with increasing number of intervening thymine bases.

In view of this it is interesting to compare the sequences ITS-$3G_0$ and ITS-$2G_3$ which have an almost identical T_m at pH 6.75 (100 mM Na^+). The anti-cooperative effect of the electrostatic repulsion between adjacent cytosines at pH 6.75 can abolish the attractive electrostatic effect of the third cytosine. As is the case for the effects of global composition, this should change with changing solution conditions. Thus at lower pH the unfavorable free energy of protonation of the central cytosines of ITS-$3G_0$ is expected to diminish and hence the overall free energy and T_m for the formation of triple helices with adjacent cytosines will become more favorable than the triple helix with fewer cytosines. This is indeed observed experimentally.

5.2.2) The Effect of Counterion Concentration on the Melting Temperature

5.2.2.1) The Effect of Global Composition on the $d(T_m)/d(\log[Na^+])$

The $d(T_m)/d(\log[Na^+])$ Changes with Ionic Strength

Table IX lists the slope of the change in T_m with change in the logarithm of counterion concentration $\{d(T_m)/d(\log[Na^+])\}$ for the intramolecular triple helix to hairpin transition at pH 6.75. For some oligonucleotides the plot of T_m versus log $[Na^+]$ shows different slopes at ionic strength of less than 200 mM Na^+ and at ionic strength of more than 400 mM Na^+ (see **Figure 4.10** & **Appendix B**). The $d(T_m)/d(\log[Na^+])$ values for conditions of low and high ionic strength are given in **Table IX**. The change in the slope for oligonucleotides ITS-$2G_3$, ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ is unexpected. According to the theory it is indicative for the presence of a low salt and a high salt conformation with different linear charge densities. A similar behavior for an intermolecular triple helix to duplex transition has been observed by Plum et al (1990), who noted that their triple helix showed a very low slope $d(T_m)/d(\log[Na^+])$ at ionic strength of less than 200 mM Na^+, but did show an ionic strength dependent increase in hyperchromicity. Once the triple helix was completely formed above 200 mM Na^+ they observed a linear change in T_m with log $[Na^+]$. The slope at high ionic strength for the intermolecular triple helix is of the same order of magnitude as those observed for the intramolecular triple helix to hairpin transition at high ionic strength.

Triple Helices have Reduced End Effects Compared to Double Helices

A comparison of the $d(T_m)/d(\log[Na^+])$ value of ITS-ATT ($d(T_m)/d(\log[Na^+]) = 33.8 \pm 0.5°C$) with that of the corresponding polymer triple helix poly d(T) • poly d(A) • poly d(T) ($d(T_m)/d(\log[Na^+]) = 36°C$) (Guschlbauer (1990)) reveals that for this oligonucleotide end effects influence the degree of ion condensation only marginally. As was shown before the difference in $d(T_m)/d(\log[Na^+])$ between the hairpin conformation and the corresponding poly d(A) • poly d(T) double helix is much larger. At present it is difficult to discern whether the higher charge density in the triple helix reduces the end effect parameter, or whether this is due to fortuitous compensation caused by the loop regions. An intramolecular triple helix formed from 12 T•A•T base-triplets linked by uncharged hexaethylene glycol loops has a $d(T_m)/d(\log[Na^+])$ value of ($d(T_m)/d(\log[Na^+] = 36.2 \pm 0.5°C$) (Durand et al (1992)) close to the polymer value, indicating that the higher charge density in triple helices may predominantly be responsible for reduced end effects. Because of the potential applications of oligonucleotide triple helices, it will be useful to determine the effect of triple helix size in more detail in the future.

The T_m of Intermolecular Triple Helices varies with Composition and Ionic Strength

Figure 5.2.5 shows an overlay of a plot of T_m versus log $[Na^+]$ derived from the experimentally determined $d(T_m)/d(\log[Na^+])$ values for the triple helix to hairpin transition of the oligonucleotides ITS-ATT, ITS-$2G_3$, ITS-$3G_2$ and ITS-$4G_1$. The overall cytosine content in the third strand changes from 0% to 40%.

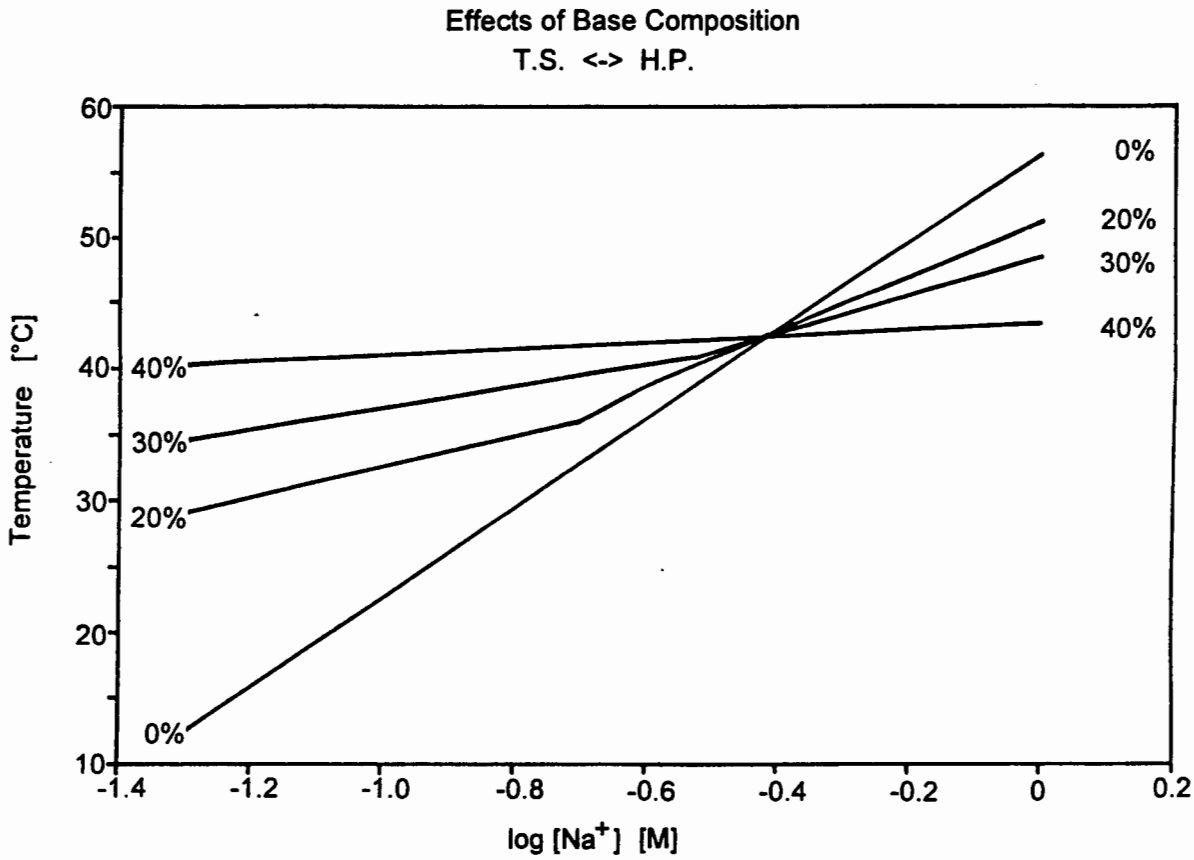


Figure 5.2.5 The effect of ionic strength on the T_m of the triple helix to hairpin transition of oligonucleotides with different global composition

At 0 %C and at 40 %C the $d(T_m)/d(\log[Na^+])$ is a straight line over the whole range of ionic strength. The slope of the cytosine rich species is much smaller than that of the intramolecular triple helix ITS-ATT. At intermediate cytosine content (ITS-$2G_3$ and ITS-$3G_2$) there is a gradual change in slope starting at about 200 mM Na^+. All four curves intersect at 400 mM Na^+, i.e at this ionic strength the triple helix to hairpin transition is independent of the global composition. The gradual increase in slope with decreasing cytosine content is a reflection of the different forces stabilizing T•A•T and C^+•G•C triplets. High ionic strength favors T•A•T triplets because of increased shielding of the negatively charged backbones, while it disfavors C^+•G•C triplets because an increase in salt concentration causes a decrease the pka of the third strand

cytosines. An increase in ionic strength also leads to repulsion between the positive charge on the protonated cytosine and the cations condensed onto the exterior of the DNA helix. Hence high ionic strength will favor T•A•T triplets and disfavor C^+•G•C triplets. It is conceivable that the difference in the forces stabilizing the two different types of base-triplets is responsible for the observed change in $d(T_m)/d(\log[Na^+])$ with ionic strength.

Application of Polyelectrolyte Theory to the Triple Helix to Hairpin Transition

Equation **(2.7)** predicts a linear decrease in the degree of counterion condensation with a linear increase in the number of protonated cytosines in the third strand. This can only be the case if we assume that the term $(2.3*R*T_m^2/\Delta H_{obs})$ of equation **(2.7)** is independent from the local and global composition effects. Previously it has been shown (chapter 5.2.1) that the T_m of the triple helix to hairpin transition increases linearly with increasing number of third strand cytosines. Hence the T_m must give rise to a corresponding increases in ΔH_{obs}. But ΔH_{obs} can not be determined because of the lack of calorimetric data and because of the uncertainty regarding the validity of the determination of the van't Hoff enthalpy (Plum et al(1990), Völker et al. (1993)). It is however possible to attempt to estimate the change in ΔH_{obs} and any possible error introduced by assuming that the term $(2.3*R*T_m^2/\Delta H_{obs})$ is independent of global and local composition.

Composition Dependent Changes in T_m are Predominantly Enthalpic in Origin

Any increase in T_m corresponds to a proportional increase in the free energy ΔG for the transition. Since the free energy is related to the enthalpy of the transition by equation **(5.2.2)**:

$$\Delta G = \Delta H - T * \Delta S \qquad (5.2.2)$$

any increase in ΔG must be due to either a corresponding increase ΔH at constant ΔS (1), a increase of ΔS at constant ΔH (2), or a combination of both (3). For cases (1) and (2) it is possible to calculate the expected change in equation **(2.7)**, while in case (3) the possibility of large and compensating enthalpy/entropy changes make this more difficult. In conventional double stranded DNA any increase in free energy is almost exclusively due to an increase in enthalpy (Klump (1988)), since the entropy change is due to the gain in the degrees of freedom of the phosphate backbone. It is not unreasonable to expect that the same will hold true for the triple helix to hairpin transition as well even if the release of protons is not taken into account (i.e. case 1), resulting in the term $(2.3*R*T_m^2/\Delta H_{obs})$ to be independent of sequence. If it is assumed that the increase in free energy is entirely due to an increase in entropy (i.e. case 2), then the numerical values of the term $(2.3*R*T_m^2/\Delta H_{obs})$ will differ by less than 10 % for the oligonucleotides with the largest difference in T_m. Since it is highly unlikely that any change in free energy for triple helical DNA is only due to entropic effects, it can be expected that changes in enthalpy and entropy are responsible for any change in free energy (i.e. case 3). In the absence of any obvious reasons for potential large and compensating composition dependent enthalpy and entropy changes it can be assumed that most of the free energy change is due to enthalpic contributions. Hence it is likely that only small variations (less than 10%) in the term $(2.3*R*T_m^2/\Delta H_{obs})$ do occur and consequently any error introduced by assuming a constant term in equation **(2.7)** will be small.

The $d(T_m)/d(\log[Na^+]$ is not Proportional to the Global Composition

An inspection of the plot of $d(T_m)/d(\log[Na^+])$ versus %C in the third strand (**Figure 5.2.6**) reveals that even though the $d(T_m)/d(\log[Na^+])$ values at high and low ionic strength decrease with increasing cytosine content, this decrease at best only approaches a linear behavior.

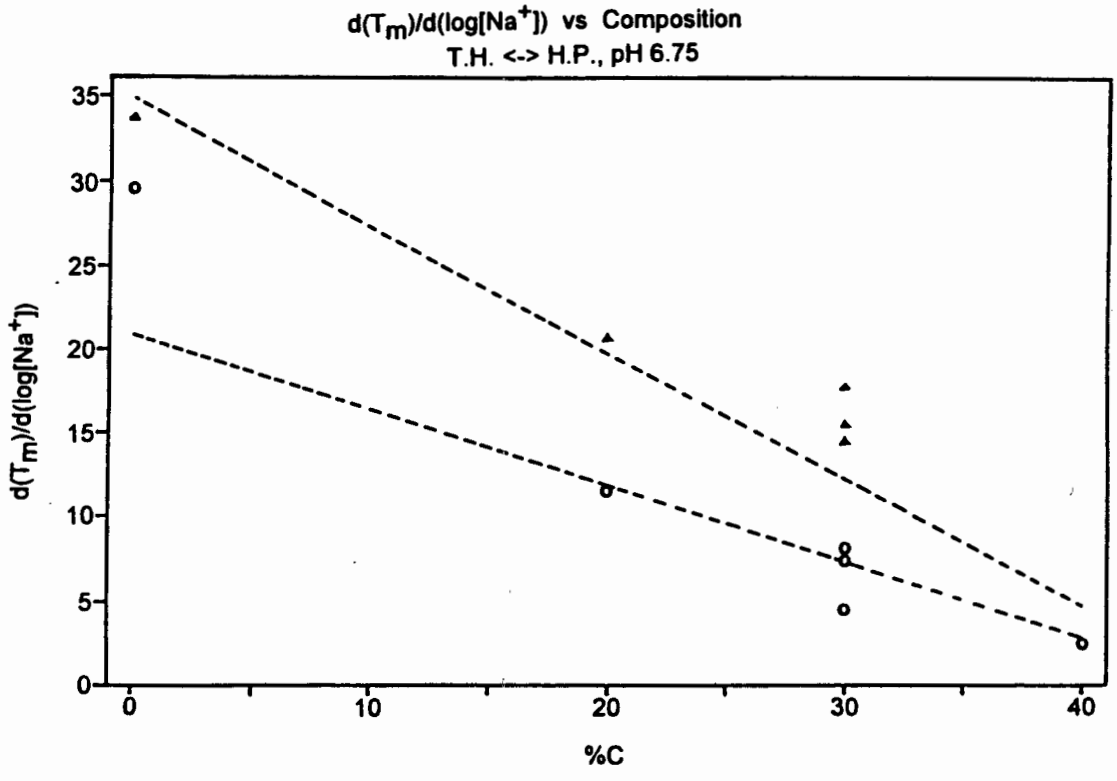


Figure 5.2.6 The effect of base-composition on the degree of counterion release ($d(T_m)/d(\log[Na^+])$) of the triple helix to hairpin transition at pH 6.75

It is impossible to fit a linear equation to the data points with any reasonable degree of confidence. The most obvious reason coming to mind is that the Manning parameter η

depends on the difference in counterion condensation between the triple helix and the hairpin with single stranded 3' extension. Since as was shown previously the amount of counterions condensed onto the hairpin depends strongly on the sequence (chapter 5.1.2), composition dependent ion condensation onto the hairpin helix and single stranded tail may contribute to the nonlinear behavior of the $d(T_m)/d(\log[Na^+])$ with composition. The most likely source for the nonlinear behavior however is to be found in the fundamental assumption of equation **(2.7)** that a proton can be considered to simply reduce the overall charge density. Since the protons are site-bound in the interior of the triple helix it is very likely that protonation events cannot be treated by the conventional counterion condensation theory. It is currently not possible to evaluate the effect of proton site-binding on the experimental $d(T_m)/d(\log[Na^+])$ value in more detail.

5.2.2.2) **The Effect of Local Composition on the $d(T_m)/d(\log[Na^+])$**

The effect of local composition on the change in T_m with log $[Na^+]$ can be determined by comparing the $d(T_m)/d(\log[Na^+])$ for the oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ (**Table IX**). All three oligonucleotides contain the same number of cytosines in the third strand but have different local composition. For each of the oligonucleotides the slope $d(T_m)/d(\log[Na^+])$ increases at higher ionic strength. Any differences in the $d(T_m)/d(\log[Na^+])$ should only be due to differences in the degree of ion condensation of the hairpin with 3' single stranded tail, the intramolecular triple helical state, or the term $\mathbf{(2.3*R*T_m^2/\Delta H_{obs})}$ of equation **(2.7)**. Alternatively differences in $d(T_m)/d(\log[Na^+])$ reflect the effect of site-binding of the protons to the interior of the triple helix.

Local Changes in T_m are Predominantly Enthalpic in Origin

As discussed above differences in the term $\mathbf{(2.3*R*T_m^2/\Delta H_{obs})}$ are small or insignificant and the term can be take as constant for the purpose of this discussion. The observed differences in $d(T_m)/d(\log[Na^+])$ value of the hairpin to coil transition of ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ are small (Chapter 5.1.2), indicating that the degree of counterion condensation to the hairpin state is very similar. Hence to a first approximation differences in $d(T_m)/d(\log[Na^+])$ with local composition can be considered to be due to differences in the degree of ion condensation onto the triple helix caused by site-binding of the protons.

Protonation of Adjacent Cytosines is Disfavored

The strongly reduced $d(T_m)/d(\log[Na^+])$ value of ITS-$3G_0$ compared to the other oligonucleotides must be a reflection of the direct neighborhood of the central two protonated cytosines. It is possible that as the ionic strength is increased, the difference in the pka_{app} of the central two cytosines also increases. Hence any increase in stability due to increased shielding of the phosphate charges will be offset by the increased free energy cost of protonating the second cytosine. For the oligonucleotides ITS-$3G_1$ this effect becomes less noticeable and vanishes completely for ITS-$3G_2$, where the cytosines are separated by intervening thymines.

At high ionic strength, where the electrostatic contributions of the negatively charged phosphate backbones of T•A•T triplets predominates, the $d(T_m)/d(\log[Na^+])$ value of all three oligonucleotides approach a common value. Hence at high ionic strength the effect of protonation of the cytosines becomes less important for the stability of the triple helix. Because of the relatively large uncertainty of the $d(T_m)/d(\log[Na^+])$ values at high ionic strength and since the Manning formalism strictly does not apply under these conditions it is difficult to judge the observed differences in the $d(T_m)/d(\log[Na^+])$ value.

5.2.3) The Effect of pH on the Melting Temperature

5.2.3.1) The Effect of Global Composition on $d(T_m)/d(pH)$

The T_m of the Triple Helix to Hairpin Transition increases with pH

Table IX lists the change in T_m with change in solution pH $\{d(T_m)/d(pH)\}$ for the intramolecular triple helix to hairpin transition. Because of the equilibrium between the protons in solution and the protons bound to the Hoogsteen cytosines the experimentally observed strong change in T_m with pH is expected. Equation **(2.14)** predicts that the $d(T_m)/d(pH)$ value is proportional to the difference in the number of protons bound to the triple helix and to the hairpin. The intramolecular triple helix to hairpin transition is observed between pH 8.0 and approximately pH 6.0 - more than 1.5 pH units above the pka of free cytosines. The cytosines of the third strand will be unprotonated in the coil state (i.e $\mathbf{k_c} = 0$). In this pH range the hairpin does not take up any protons and therefor the $d(T_m)/d(pH)$ value will be proportional to the number of protons bound to the intramolecular triple helix. In order to determine the number of protons bound to the triple helix, it is necessary to evaluate the term $\mathbf{(2.303*R*T_m^2/\Delta H_{obs})}$ in equation **(2.14)**. Due to the lack of calorimetric data this is not easily achieved. However, using the same argument presented previously for the dependence of the $d(T_m)/d(\log[Na^+])$ on the amount of counterions released on melting (Chapter 5.2.1.1), it can be assumed that the variations in T_m are mostly enthalpic in origin. The composition and sequence dependent variations in the numerical value of $\mathbf{(2.303*R*T_m^2/\Delta H_{obs})}$ are expected to be small and taken as constant in equation **(2.14)**. Hence the $d(T_m)/d(pH)$ value will change proportionally to the number of protons released from the triple helical state upon melting.

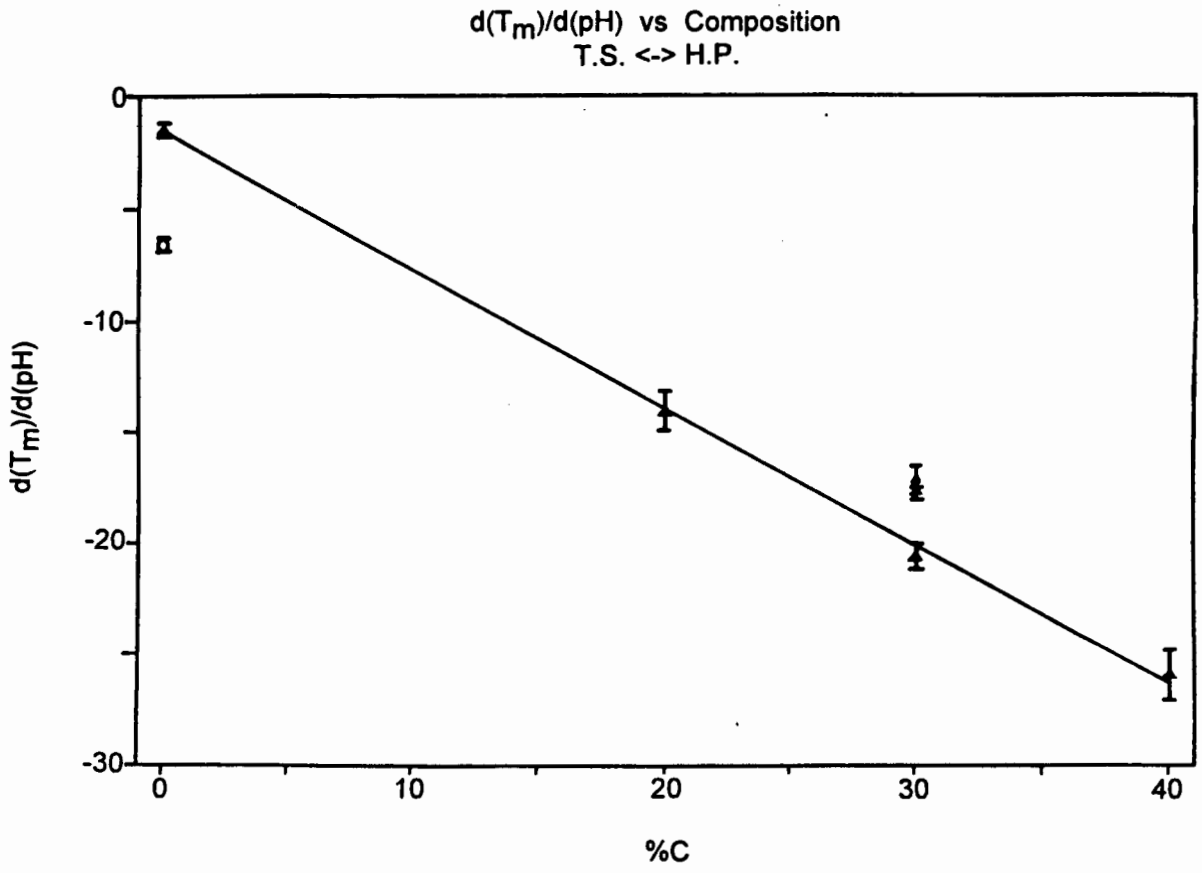


Figure 5.2.7 The effect of base composition on the change in T_m with pH ($d(T_m)/d(pH)$) of the triple helix to hairpin transition at pH 6.75

The $d(T_m)/d(pH)$ Changes Linearly with Global Composition

Figure 5.2.7 shows a plot of $d(T_m)/d(pH)$ versus the percentage of third strand cytosines (%C) for all the oligonucleotides with the Hoogsteen cytosines sufficiently well separated to exclude sequence specific effects. The relationship between %Hoogsteen cytosines and the $d(T_m)/d(pH)$ value can be fitted by linear least square fitting to the equation:

$$d(T_m)/d(pH) = -0.60 * (\%C) - 2.38 \qquad (5.2.3)$$

provided that the $d(T_m)/d(pH)$ value of ITS-ATT at pH above pH 7.0 is used. The linear least square fit gives an excellent statistical fit (significant at 0.1% level).

The Effect of pH on the T_m of the Intramolecular Triple Helix ITS-ATT

As was already discussed previously the oligonucleotide ITS-ATT cannot directly be compared to the other oligonucleotides, because of the presence of cytosines in both loops. A plot of T_m versus pH for the intramolecular triple strand to hairpin transition shows the presence of three distinct regions (Figure **5.2.8**). Between pH 8.5 and pH 7.0 the T_m changes barely with a change in pH giving a slope $d(T_m)/d(pH) = -1.47°C$. This is of the same order of magnitude as the $d(T_m)/d(pH)$ for the pH independent hairpin to coil transitions. Hence the intramolecular triple helix to coil transition for ITS-ATT above pH 7.0 is independent of the pH. Between pH 7.0 and pH 5.0 the T_m suddenly changes with pH giving rise to a linear dependence of T_m on the pH with a slope of $d(T_m)/d(pH) = -6.5°C$. Reducing the pH further the T_m gradually becomes less dependent on the pH of the solution. It is interesting that according to equation **(5.2.3)** a slope of $d(T_m)/d(pH) \approx 6°C$ would correspond to the intramolecular triple helix to hairpin transition of an oligonucleotide containing a single protonated cytosine residue. Also a gradual decrease in $d(T_m)/d(pH)$ close to pH 4.5 is expected. At pH values close to the pka of free cytosine the difference in the number of protons between the helix and coil state $\Delta\mathbf{k}$ is expected to approach zero as an increasing fraction of the single stranded cytosines will remain protonated. Previously it has been argued, based on the P_1-nuclease digestion pattern that a single cytosine residue of the C_4 hairpin loop can be protonated. A P_1-nuclease digestion pattern for the triple helix loop of ITS-ATT cannot be observed, due to the lack of stability of the nicked A•T helix.

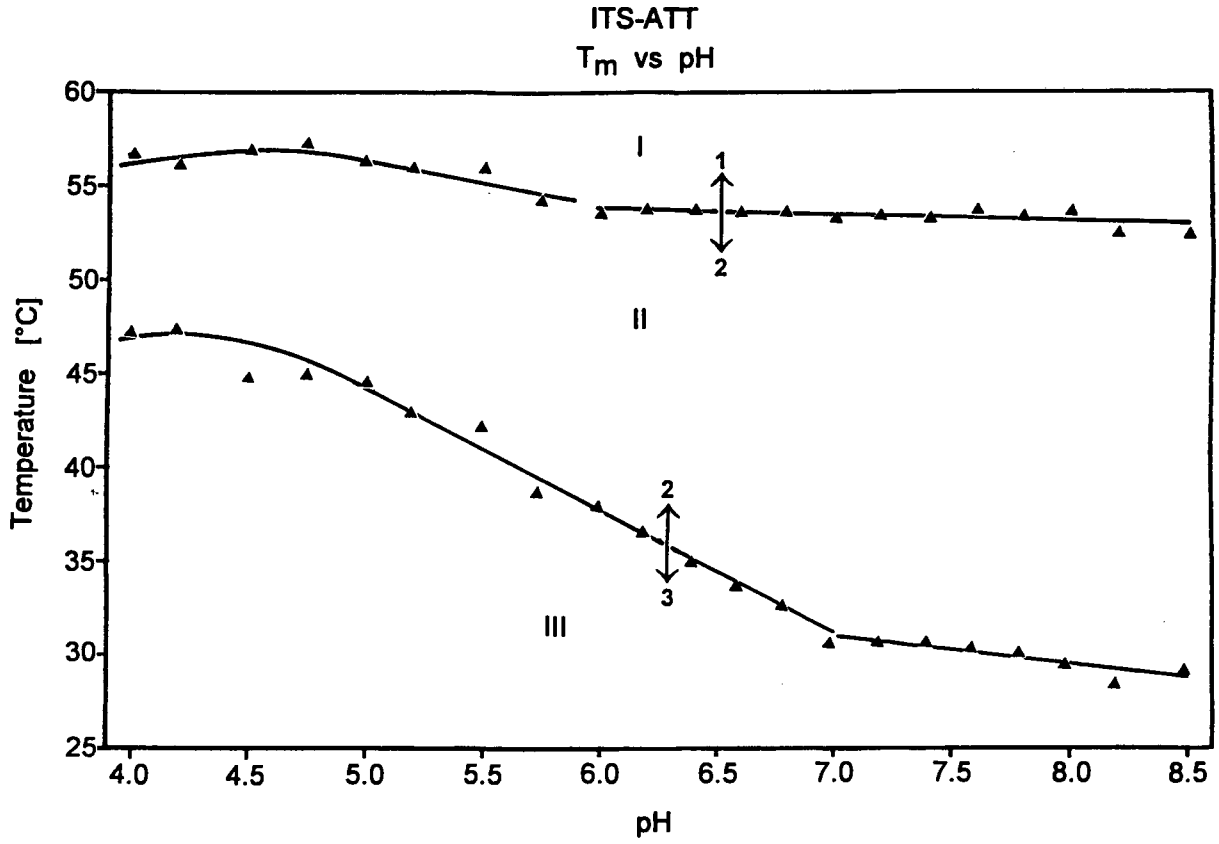


Figure 5.2.8 Phasediagram: T_m versus pH for ITS-ATT (150 mM Na^+)

There is however no apparent reason, why the C_4 triple helix loop of ITS-ATT is not also able to form the same unusual C^+•C base-pair. Backing for this approach comes from the close agreement between the T_m and $d(T_m)/d(pH)$ values of ITS-ATT at pH 8.0 with those expected from a linear extrapolation from the values of the other oligonucleotides.

All Third Strand Cytosines have the same pka Value

The linear change in $d(T_m)/d(pH)$ with the %Hoogsteen cytosines implies that the number of protons released on melting is proportional to the number of cytosine bases.

Identical pka_{app} values for the third strand cytosines can only be expected, if the protonation of one cytosine does not effect the protonation of any other cytosine. Hence the conditions for the study of global compositional effects, namely that the cytosines are sufficiently well separated to exclude local sequence specific effects is fulfilled by these oligonucleotides.

It needs to be pointed out however that the value for the slope of –0.6°C in equation (**5.2.3**) is characteristic only for the triple helix to hairpin transition. Protonated double helical structures with a different backbone charge density will show a different dependence of $d(T_m)/d(pH)$ on the percentage of protonated residues. It is therefor not surprising that the hairpin to coil transition of ITS-ATT shows a reduced $d(T_m)/d(pH)$ value at pH's below pH 6.0 compared to the triple helix to hairpin transition, even though the same number of protons are involved. Because of the linear relationship of $d(T_m)/d(pH)$ with %C it is possible to find a pH value at which the T_m's of any two intramolecular triple helix to hairpin transitions of oligonucleotides with different cytosine content are identical. **Figure 5.2.9** shows a plot of T_m versus pH for the triple helix to hairpin transition for four different oligonucleotides. Each straight line is derived from the experimental $d(T_m)/d(pH)$ value. It is interesting that they all intersect at pH 7.1 (150 mM Na^+), regardless of the base composition, indicating that under those conditions the triple helices are equally stable. At pH 7.1 and 150 mM Na^+ the free binding energy of a Hoogsteen cytosine in a C^+•G•C triplet is identical to that of a Hoogsteen thymine in a T•A•T triplet, despite the different forces contributing to the stability of the base triplets.

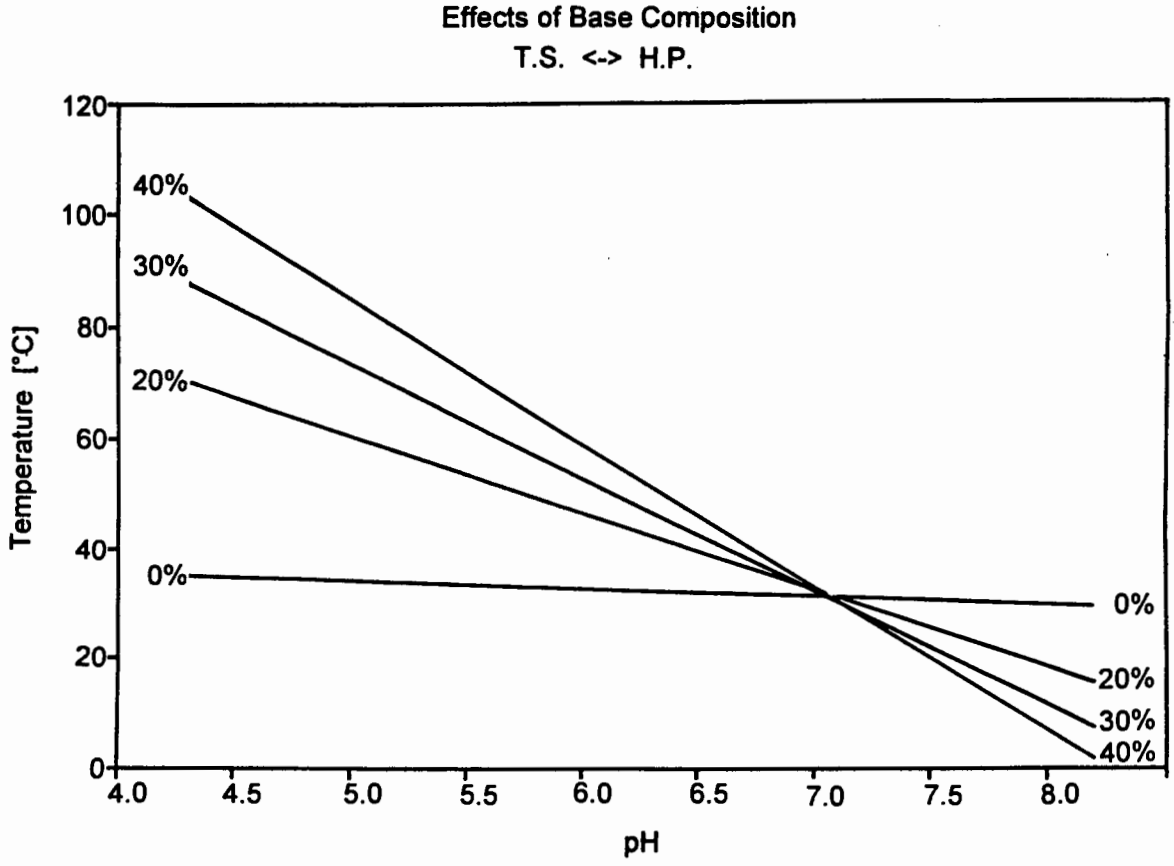


Figure 5.2.9 The effect of pH on the T_m of the triple helix to hairpin transition of oligonucleotides with different global composition.

At pH values above pH 7.1 in 150 mM Na^+ the cytosine binding free energy is lower than the thymine binding free energy, while at pH values below pH 7.1 (in 150 mM Na+) the cytosine binding free energy is higher than that of thymine. Hence the stability of the intramolecular triple helices is a linear function of the overall cytosine content and the environmental conditions (mainly pH).

5.2.3.2) The Effect of Local Composition on the $d(T_m)/d(pH)$

Local Composition Effects are due to Interactions between Adjacent Cytosines

Since base compositional effects are due to the replacement of T•A•T triplets by C^+•G•C triplets in a perfect T•A•T triple helix, any local sequence specific effects must be due to the interactions of neighboring C^+•G•C base-triplets. The oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ have the same global composition but different local composition, and can be used to elucidate sequence specific effects. In conventional double stranded DNA sequence specific effects arise from the hydrogen-bonding and the stacking interactions of nearest neighbor base-pairs. In contrast it has been shown above that the electrostatic forces dominate the binding of the third strand. Since electrostatic forces decay proportional to the square of the distance separating the charges, while the van der Waal's stacking interactions decrease as a function of the sixth power with distance, it is feasible that electrostatic effects reach beyond the nearest neighbors and are detectable in triple helices. In **Table IX** the $d(T_m)/d(pH)$ values of the oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ are listed.

Local Composition Effects Result in a Reduced pH Dependence of the T_m

It is somewhat surprising that the $d(T_m)/d(pH)$ values for both ITS-$3G_0$ and ITS-$3G_1$ are similar to each other and smaller than that of ITS-$3G_2$. According to equation **(2.14)** the $d(T_m)/d(pH)$ value corresponds to the number of protons release on melting ($\Delta\mathbf{k}$). Since only the cytosines are protonated and the three oligonucleotides contain the same number of Hoogsteen cytosines, it is expected that the $\Delta\mathbf{k}$ value (i.e. the number of protons released on melting) is identical independent of sequence. Differences in $d(T_m)/d(pH)$

can conceivably only be due to differences in the term **($2.303*R*T_m^2/\Delta H_{obs}$)** in equation **(2.14)**. In the previous discussion, it was assumed that the increase in T_m observed with change in base composition is essentially enthalpic in origin and hence any changes in the term **($2.303*R*T_m^2/\Delta H_{obs}$)** are expected to be small. Intuitively it can be assumed that this is also true for the oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$. If this is the case, then what is the cause for the difference in $d(T_m)/d(pH)$ between these three different oligonucleotides?

Substituting the $d(T_m)/d(pH)$ values of ITS-$3G_0$ and ITS-$3G_1$ into the equation **(5.2.3)** results in an apparent content of $\approx$ 25% protonated cytosines in the third strand instead of the 30% expected from the global cytosine content. The difference can only be explained if one assumes that the oligonucleotides are not fully protonated due to the repulsion between the adjacent positive charges. The slope of ITS-$3G_2$ appears normal, i.e. as expected from the equation **(5.2.3)**. This means that spacing of adjacent cytosines by two intervening thymines is sufficient to cancel the electrostatic repulsions between adjacent protonated cytosines.

THE INTRAMOLECULAR TRIPLE HELIX TO COIL TRANSITION

5.3.1 The Effect of Local and Global Composition on the Melting Temperature

5.3.1.1) The Effect of Global Composition on the T_m

One Step Triple Helix to Coil Transition

At pH values below pH 6.0 (100 mM Na^+) the triple helix to hairpin transition merges with the hairpin to coil transition, resulting in a single cooperative intramolecular triple helix to coil transition. The sole exception to this is the oligonucleotide ITS-ATT, where the triple helix to hairpin and the hairpin to coil transitions remain separated throughout the pH range investigated. Since the intramolecular triple helix of ITS-ATT does not involve any protonated base-pairs or triplets (with the possible exemption of the $C \bullet C^+$ base-pairs in the loop regions), this difference is expected. **Table X** lists the T_m values of the different oligonucleotide triple helix to coil transitions in 100 mM Na^+ at pH 5.5 and at pH 4.5.

The T_m is Proportional to the Global Composition

The T_m of the triple helix to coil transition can be presented as a linear function of the base composition $\%C^+GC$ (**Figure 5.3.1**). Because of the absence of a single step triple helix to coil transition for ITS-ATT, only the oligonucleotides ITS-$2G_3$, ITS-$3G_2$ and ITS-$4G_1$ are considered for the discussion of global composition effects on the T_m of this transition.

TABLE X : The intramolecular triple helix to coil transition

	T_m pH 5.5 [°C]	T_m pH 4.5 [°C]	ΔH_{vH} pH 4.5 [kcal/M]	$\delta T_m/\delta \log [Na^+]$ pH 5.5 [°C]	$\delta T_m/\delta \log [Na^+]$ pH 4.5 [°C]	$\delta T_m/\delta (pH)$ [°C]
ITS-ATT				19.22 ± 0.43	21.81 ± 0.45	
ITS-$2G_3$	53.3 ± 1.0	56.0 ± 1.0	90.5± 9.4	13.80 ± 0.38	17.74 ± 0.37	- 5.17± 1.14
ITS-$3G_0$	57.0 ± 0.2	62.0 ± 0.3	106.6± 4.8	12.16 ± 0.25	13.75 ± 0.50	- 5.46± 0.41
ITS-$3G_1$	60.5 ± 1.0	66.7 ± 1.0	127.2± 8.2	10.62 ± 0.24	14.72 ± 0.20	- 8.1 ± 0.36
ITS-$3G_2$	62.7 ± 1.0	65.0 ± 1.0	N.D.	5.42 ± 0.17	15.39 ± 0.30	- 8.2 ± 0.53
ITS-$4G_1$	65.0 ± 1.0	71.0 ± 1.0	119.4±11.9	8.00 ± 0.60	12.65 ± 0.27	-10.1 ± 0.85

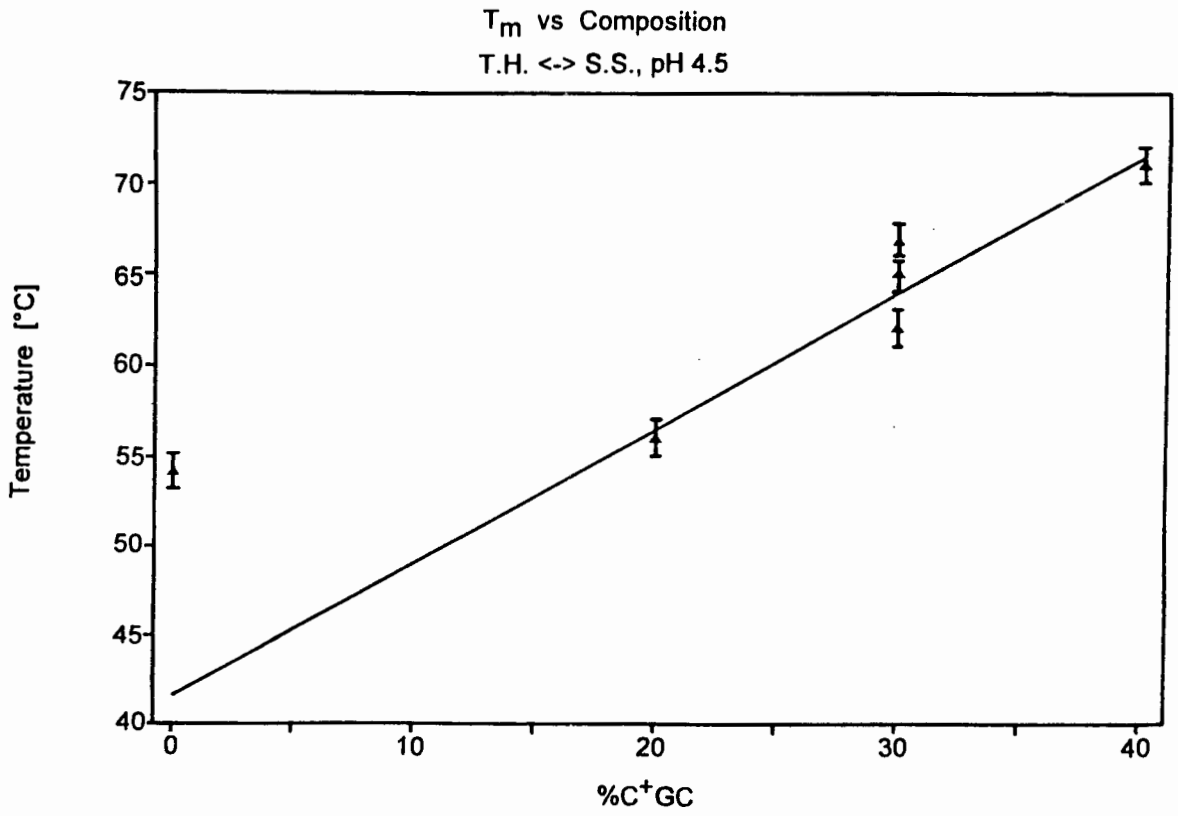


Figure 5.3.1 The effect of base composition on the melting temperature of the triple helix to random coil transition at pH 4.5 (100 mM Na^+).

Despite the scarcity and lack of range of the data points, the change of T_m with base composition at pH 4.5 can be described by a linear equation with a good statistical fit. The fit of a linear equation to the T_m values at pH 5.5 is less good, probably because at this pH the triple helix to coil transition may still result in the formation of a considerable proportion of hairpin intermediates. The formation of hairpin intermediates is also indicated by the observation that the van't Hoff enthalpy at pH 5.5 is much smaller than the van't Hoff enthalpy at pH 4.5, implying a multi-state melting process. If the fraction of hairpin intermediates formed during melting changes with global composition, then it can be expected that the transition width and T_m will also show some variation at pH 5.5. Never the less a linear equation does seem to give in a reasonable approximation of

the experimental data at pH 5.5. The linear least square fitting of the T_m values at pH 4.5 results in equation (**5.3.1**):

$$T_m = 0.75 * (\%C^+GC) + 41.5 \qquad (5.3.1)$$

The intercept of equation (**5.3.1**) depends on oligonucleotide end effects and environmental conditions, while the slope should only depend on the global composition. The straight line described by equation (**5.3.1**) has been extrapolated to 0 $\%C^+GC$ in **Figure 5.3.1**. The intercept indicates a melting temperature of 41.5°C for the triple helix to coil transition of an intramolecular complex without cytosines in the Hoogsteen position. This is considerably lower than the observed melting temperature of the hairpin to coil transition for ITS-ATT (T_m = 54.2°C). It is therefor not surprising that the one step triple helix to coil transition of ITS-ATT cannot be observed at pH 4.5.

The Slope of the Triple Helix to Coil Transition can not be derived from the slopes of the Component Transitions

A comparison of the slope given by equation (**5.3.1**) with that of the triple helix to hairpin transition (equation (**5.2.1**)) or that of the hairpin to coil transition (equation (**5.1.1**)) shows that the T_m of the intramolecular triple helix to coil transition is nearly twice as dependent on the global base composition, than either of the two component transitions at pH 6.75. What can be the reason for this behavior? Conceptionally the one step triple helix to coil transition can be considered to arise from the complete overlap of the two component transitions: the triple helix to hairpin transition and the hairpin to coil transition. Both component transitions are strongly dependent on the global composition (see chapter 5.1 and 5.2). For the hairpin to coil transition this is determined by the difference in the number of hydrogen bonds between G•C and A•T base-pairs, while for

the triple helix to hairpin transition the protonation of the cytosines is responsible for the change in T_m with base composition. The thermal stability of the intramolecular triple helix will depend on the protonated cytosines in the third strand component and on the extra hydrogen bond of the G•C base-pair in the hairpin component. An additional unwanted level of complexity will arise from the partial protonation of the C_4 loop of the hairpin component of the intramolecular triple helix at pH values below pH 6.0. As mentioned above the formation of the unusual $C•C^+$ base-pair in the hairpin loop increases the stability oft the hairpin component below pH 6.0 compared to the hairpin to coil transition at pH values above pH 6.0. But since all oligonucleotides contain the same C_4 loop, the slope of T_m versus global base composition at pH 4.5 is not effected. It may however contribute to the slope of T_m versus base composition at pH 5.5 by selectively stabilizing the hairpin component, and thus increasing the fraction of hairpin intermediate formed during the melting process.

Since the global composition effects the T_m of the triple helix to hairpin transition in a pH dependent way, while the T_m of the hairpin to coil transition is only marginally pH dependent, the gross compositional effects can not be just additive. Accordingly the slope of equation (**5.3.1**) cannot be derived by the simple addition of slopes of equations (**5.2.1**) and (**5.1.1**) at pH 4.5. The reason for this may be found in the observation that the intramolecular triple helix at pH 4.5 has a T_m that is higher than the extrapolated T_m expected for the hairpin to coil transition, but lower than the extrapolated T_m expected for the triple helix to hairpin transition. This can be explained, if one assumes that the third strand component stabilizes the hairpin component in the complex relative to the free hairpin. The physical properties of the intramolecular triple helix therefor appear to be determined by the physical properties of the individual structural components plus an additional function accounting for the interaction of these structural components. It is at present not possible to derive a numerical description for the putative additional

interactions, since the differential protonation of cytosines in the triple strand, hairpin and single strand conformations at pH 4.5 make it impossible to extrapolated the T_m values. Similarly at pH 5.5 the presence of a considerable fraction of hairpin intermediates during the thermal denaturation does not allow the calculation of a mean T_m from the contributions of the individual structural components.

5.3.1.2) **The Effect of Local Composition on the T_m**

Adjacent Cytosines Destabilize the Triple Helix

A comparison of the T_m 's of the oligonucleotide ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ at pH 4.5 reveals (**Table X**) that small local composition dependent variations in the thermal stability exist. As is the case for the triple helix to hairpin transition, the T_m of ITS-$3G_0$ is lower than that of the other oligonucleotide sequence variants. This underlines the importance of the electrostatic repulsion between adjacent protonated cytosines for the overall stability of the intramolecular triple helix. The decreasing difference in T_m compared to the triple helix to hairpin transition may be due to the lower pH, which will mask any changes in pka_{app} for the adjacent cytosines.

ITS-$3G_2$ shows a Small Fraction of an Altered Conformation at pH 4.5

The T_m of the oligonucleotide ITS-$3G_1$ is marginally higher than that of the oligonucleotide ITS-$3G_2$ at pH 4.5. In the pH range between pH 6 and pH 4.5 usually the opposite is observed. The differences are in fact very small. To give an explanation for this observation one can assume that this may not represent a sequence specific effect, but rather may be caused by the presence of a small percentage of an altered protonated conformation for ITS-$3G_2$. The melting curves of ITS-$3G_2$ show the presence of a small additional pre-transition at pH 4.5 not present in any of the other oligonucleotides. But similar additional acid induced pre-transition can be observed in the melting curves of all other oligonucleotides at pH values well below pH 4.5. The formation of a small fraction of a protonated alternate structure in ITS-$3G_2$ at pH 4.5 may be responsible for the somewhat smaller than expected T_m at pH 4.5. The exact nature of this additional

conformation however has not been determined. It can be assumed that this structure does involve protonated cytosines and possibly protonated adenine residues. The formation of a small percentage of an altered conformation of ITS-$3G_2$ does not change the general conclusions regarding the triple helix to coil transition at pH 4.5. In any case the T_m of the main transition of ITS-$3G_2$ falls on the line describing the change in T_m with global composition and is generally in good agreement with the expected behavior of an intramolecular triple helix with 30% cytosines.

Unlike for the triple helix to hairpin transition the T_m of the triple helix to coil transition of the oligonucleotide ITS-$3G_0$ is considerably higher than that of ITS-$2G_3$. As stated above (chapter 5.2.2.2) a decrease in the pka_{app} of a second cytosine due to the presence of a protonated cytosine in the immediate neighborhood is responsible for the decreased thermal stability of the triple helix of ITS-$3G_0$ at intermediate pH values. It can be expected that this effect will diminish with decreasing pH. Hence at pH 4.5 the differences in pka_{app} of the neighboring cytosines will influence the T_m of the triple helix to coil transition only marginally resulting in a much higher T_m for ITS-$3G_0$ than that for ITS-$2G_3$.

5.3.2) The Effect of Counterion Concentration on the Melting Temperature

5.3.2.1) The Effect of Global Composition on the $d(T_m)/d(\log[Na^+])$

The T_m of the Triple Helix to Coil Transition is Proportional to the Ionic Strength

The $d(T_m)/d(\log[Na^+])$ value of the triple helix to coil transition is tabulated in **Table X**. Unlike for the triple helix to hairpin transition, the change in T_m with a ten fold increase in the counterion concentration for the triple helix to coil transition remains linear between 50 mM Na^+ and 1000 mM Na^+. No positive or negative deviation from the linear relationship can be detected for any of the oligonucleotides. Negative deviations from the extrapolated linear relationship occur only at counterion concentrations higher than 1M Na^+. It is likely that the high charge density of the intramolecular triple helix causes the effects of nonideal solution conditions to become noticeable only at much higher ionic strength than for the duplex to coil transitions.

The Triple Helix to Coil Transition for ITS-ATT Occurs only at High Ionic Strength

Table X also lists the $d(T_m)/d(\log[Na^+])$ value for the intramolecular triple helix to coil transition of the oligonucleotide ITS-ATT. This transition is not observed at low counterion concentrations, but does occur at high counterion concentration at all pH values, since the T•A•T triple helix does not require the protonation of a base. **(Figure 5.3.2)**.

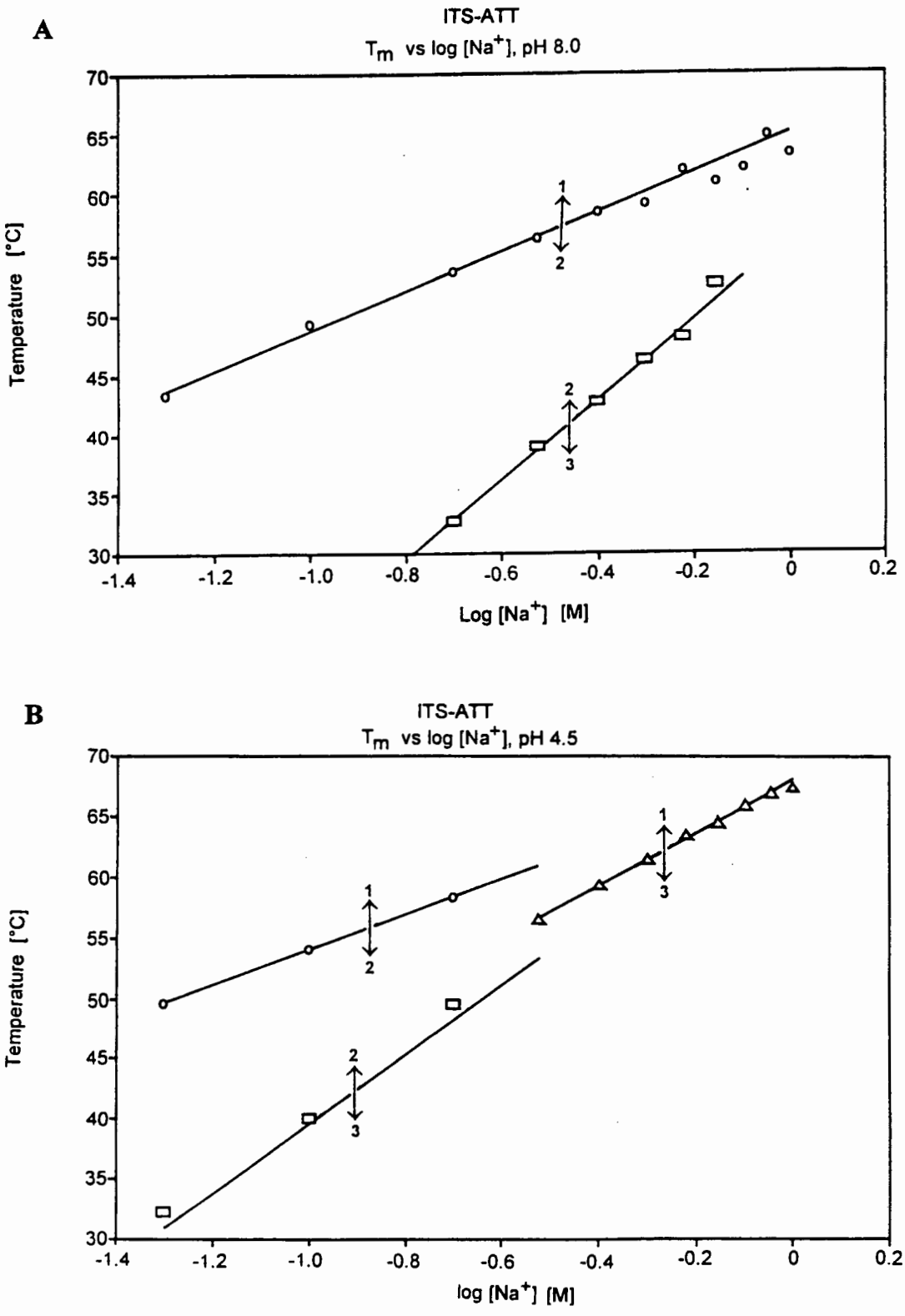


Figure 5.3.2 Phasediagram of T_m versus ionic strength of the oligonucleotide ITS-ATT at pH 8.0 (A) and pH 4.5 (B).

In the following, whenever the $d(T_m)/d(\log[Na^+])$ value for the triple helix to coil transition of ITS-ATT is referred to, it has to be remembered that this value is experimentally only accessible at high ionic strength, while the $d(T_m)/d(\log[Na^+])$ value for the triplex to coil transition of all other oligonucleotides can be measured at any ionic strength at pH 4.5. Extrapolating the melting temperature of the potential triple helix to coil transition of ITS-ATT in 100 mM Na^+ from the $d(T_m)/d(\log[Na^+])$ value results in a T_m of $T_m = 46.3°C$, which is considerably lower than the T_m of hairpin to coil transition in these conditions ($T_m = 54.2°C$). Hence the triple helix to coil transition of ITS-ATT can not occur in 100 mM Na^+ at pH 4.5 as the hairpin is thermodynamically more stable than the triple helix.

On Denaturation no Protons are Lost or Gained

Because the intramolecular triple helix contains protonated cytosine residues equation (**2.7**) predicts that the $d(T_m)/d(\log[Na^+])$ value of the triple helix to coil transition is proportional to the amount of protons and counterions **($k+\eta/2$)** released on melting. At pH 4.5 half of the cytosines in the coil state are protonated. In the intramolecular triple helix half of the cytosines are protonated as well. At least one of the cytosines in the hairpin loop (and the triple helix loop in the case of ITS-ATT) is also protonated at pH values below pH 6.0 in order to form the unusual $C \bullet C^+$ base-pair. Since the remaining two loop cytosine residues are expected to be exposed to the solvent, it is very likely that one of these cytosines will also be protonated. Hence it can be concluded that half of the cytosines in the intramolecular triple helix conformation are protonated at pH 4.5. Denaturation of the triple helix will likely result in a redistribution of the protons, but this will not be associated with a net loss or gain of protons. Consequently equation (**2.7**) predicts that the $d(T_m)/d(\log[Na^+])$ value is directly proportional to the amount of

counterions released on melting $\eta/2$, which in turn is proportional to the difference in linear charge density of the two conformations.

The $d(T_m)/d(\log[Na^+])$ is Proportional to the Global Charge Density of the Triple Helix

To calculate the amount of counterions released on denaturation it is necessary to evaluate the term $(2.3*R*T_m^2/\Delta H_{obs})$ in equation **(2.7)**. The lack of calorimetric enthalpy data makes the evaluation of this term impossible at present. It can however be argued that since the triple helix to coil transition is made up of the component triple helix to hairpin and hairpin to coil transitions, for which the term $(2.3*R*T_m^2/\Delta H_{obs})$ has been found to be either constant (hairpin $\leftrightarrow$ coil) or nearly constant (triple helix $\leftrightarrow$ hairpin), the term $(2.3*R*T_m^2/\Delta H_{obs})$ for the triple helix to coil transition as a first approximation can be treated as independent of the global composition. This assumption needs to be confirmed by a calorimetric determination of the enthalpy for the triple helix to coil transition. Having established that $(2.3*R*T_m^2/\Delta H_{obs})$ is independent of local and global composition, the term $d(T_m)/d(\log[Na^+])$ can be used to determine the difference in charge density between the triple helix and the coil state. Because of the nearly uniform distribution of the protonated cytosines in the oligonucleotides ITS-ATT, ITS-$2G_3$, ITS-$3G_2$ and ITS-$4G_1$, it can be assumed that local variations in charge density do not play a major role in the global composition dependent change in $d(T_m)/d(\log[Na^+])$. It is therefor expected that only the overall charge density will determine the magnitude of the $d(T_m)/d(\log[Na^+])$ value. Since the overall charge density is modulated by the number of protonated cytosines in the helix and the coil state a linear decrease in $d(T_m)/d(\log[Na^+])$ values with the increase of $\%C^+GC$ is expected. **Figure 5.3.3** shows convincingly that the $d(T_m)/d(\log[Na^+])$ values decrease linearly with the global composition at pH 4.5.

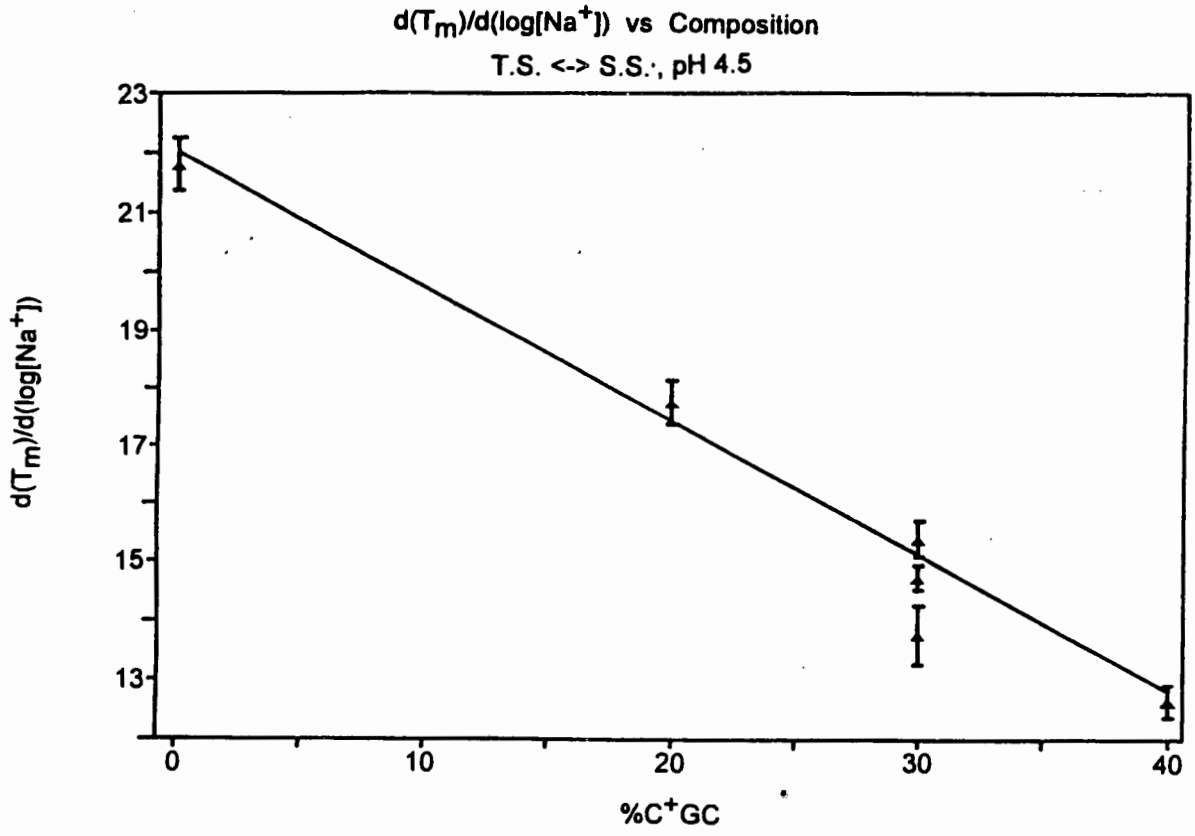


Figure 5.3.3 The effect of base-composition on the degree of counterion release ($d(T_m)/d(\log[Na^+])$) of the triple helix to random coil transition at pH 4.5

A similar behavior can be observed at pH 5.5. Linear least square curve fitting of the $d(T_m)/d(\log[Na^+])$ values with the $\%C^+GC$ at pH 5.5 and pH 4.5 results in equations (**5.3.2**) at pH 5.5:

$$d(T_m)/d(\log[Na^+]) = -0.28 * (\%C^+GC) + 19.26 \qquad (5.3.2)$$

and equation (**5.3.3**) at pH 4.5:

$$d(T_m)/d(\log[Na^+]) = -0.23 * (\%C^+GC) + 21.98 \qquad (5.3.3)$$

Both equations give a very good statistical fit of the observed experimental data. The difference in the slope of the linear equation at pH 5.5 and pH 4.5 derives from the fact that at pH 5.5 the coil state is not protonated and hence **k** is not equal to zero. The observation that the change in $d(T_m)/d(\log[Na^+])$ with base composition for the triple helix to coil transition is perfectly linear , while that of the triple helix to hairpin transition is not (chapter 5.2.2) is a reflection of the different final states of the transition.

The T_m is Proportional to the Global Composition at all Ionic Strength

As a result of the linear change in $d(T_m)/d(\log [Na^+])$ with base composition a plot of T_m versus Log $[Na^+]$ for all oligonucleotides results in a series of straight lines between 50 mM and 1 M Na^+(**Figure 5.3.4**). The line for the oligonucleotide ITS-ATT has a slight greater slope than the others, because of the effect of the protonated C_4 triple helix loop. A plot of T_m with log $[Na^+]$ for B-DNA duplex to coil transitions and for polynucleotides of different composition shows a similar linear dependency (Klump (1988)). In principle it should be possible to extrapolate from the limited data set shown in **Figure 5.3.4** the properties of intramolecular triple helices of any global composition, provided there are no local effects. But intramolecular triple helices with a $\%C^+GC$ content of more than 50% inevitably contain adjacent C^+•G•C triplets resulting in local sequence specific contributions to the overall global composition dependent contributions.

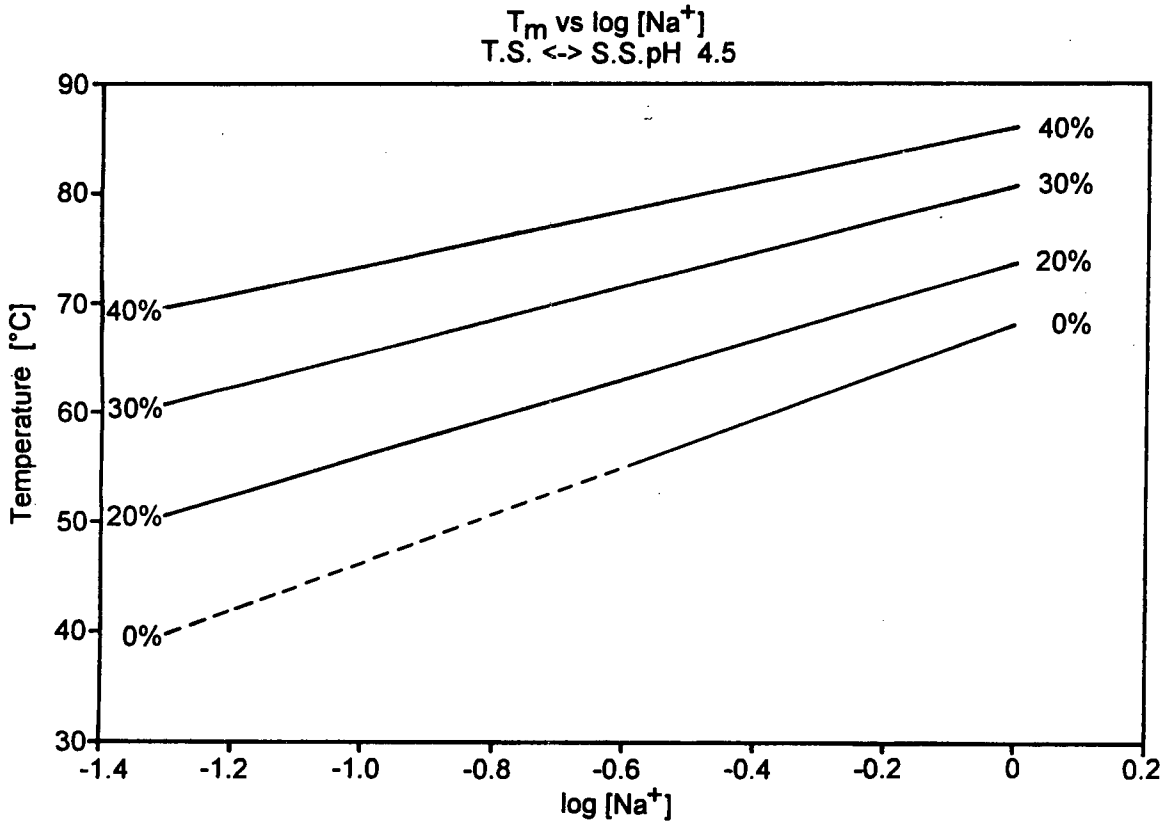


Figure 5.3.4 Phasediagram: Overlay of the T_m versus log [Na^+] plots for the intramolecular triple helix to coil transition of oligonucleotides with different global composition.

5.3.2.3) **Effect of Local Composition on the Counter Ion Condensation**

Local Composition Influences the Degree of Ion Condensation

A comparison of the $d(T_m)/d(\log[Na^+])$ values of the oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ listed in **Table X** reveals the presence of small but noticeable local sequence specific variations in the $d(T_m)/d(\log[Na^+])$ value of the triple helix to coil transition (**Figure 5.3.5**).

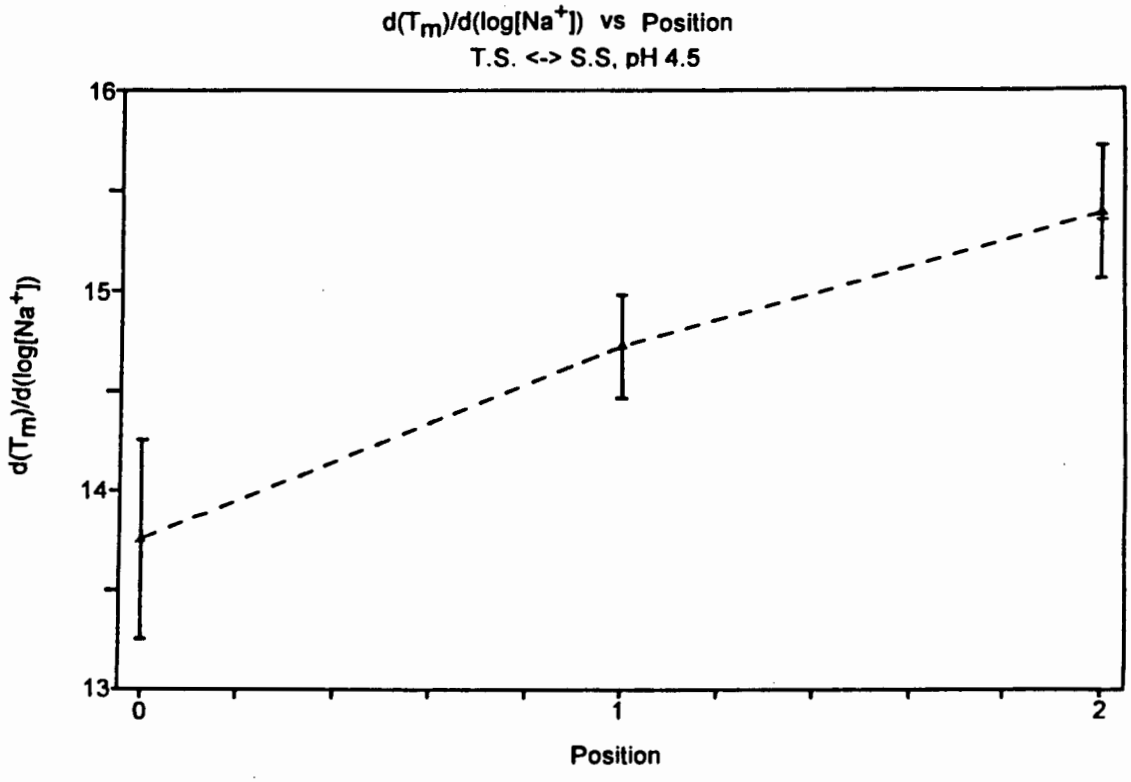


Figure 5.3.5 The effect of local composition on the degree of counterion release ($d(T_m)/d(\log[Na^+])$) of the triple helix to random coil transition at pH 4.5

What is the origin for this local sequence specific change in $d(T_m)/d(\log[Na^+])$? Equation (**2.7**) predicts that any change in $d(T_m)/d(\log[Na^+])$ is proportional to the amount of protons and counterions $(\mathbf{k+\eta/2})$ released on melting. Sequence specific effects as observed in the $d(T_m)/d(\log[Na^+])$ value must be due to either variations in the term $(\mathbf{2.3*R*T_m^2/\Delta H_{obs}})$ or in the amount of counterions $(\eta/2)$ released on melting. The variations reflect differences in the axial charge density. The overall charge density of the three oligonucleotide intramolecular triple helices and random coil states must be the same, since they have the same number of phosphates and protonated cytosines. Variations in the $d(T_m)/d(\log[Na^+])$ value must therefor reflect local variations in the charge density.

Local Changes in Charge Density due to Adjacent Cytosines Cause Local Effects

Since in the random coil state all cytosines are equally exposed to the solvent, the protonation of any cytosine is equally likely, and the charges are as far apart from each other as possible. This will tend to minimize the electrostatic repulsions between adjacent protonated cytosines, and hence it can be assumed that the local charge density in the random coil of all three oligonucleotides is very similar. This does not hold for the intramolecular triple helix state. In the intramolecular triple helix the position of the protons is determined by the position of the third strand cytosines. It is therefor expected that local variations in the charge density of the intramolecular triple helix are reflected in the observed sequence dependent variations of the $d(T_m)/d(\log[Na^+])$ value. The oligonucleotide ITS-$3G_0$, with two adjacent protonated cytosines, has the strongest local variation in charge density, resulting in the smallest $d(T_m)/d(\log[Na^+])$ value. The oligonucleotide ITS-$3G_2$, with all three protonated third strand cytosines well separated, has the least variation in the local charge density resulting in the largest $d(T_m)/d(\log[Na^+])$ value. Finally the oligonucleotide ITS-$3G_1$ has an intermediate local

variation in charge density, and shows an intermediate $d(T_m)/d(\log[Na^+])$ value. It needs to be pointed out that the observed local composition dependent variations in $d(T_m)/d(\log[Na^+])$ are small compared with the global composition dependent variations in $d(T_m)/d(\log[Na^+])$.

5.3.3) The Effect of pH on the Melting Temperature

5.3.3.1) The Effect of Global Composition on $d(T_m)/d(pH)$

The T_m of the Triple Helix to Coil Transition Reaches a Maximum at pH 4.5

Equation (**2.14**) predicts that any change in T_m with solution pH must be proportional to the change in the number of protons bound to the helix and coil states **k**. At pH values more than 0.5 pH units above the pka of free cytosines, the coil state will only be fractionally protonated. Hence at pH values above pH 5.0 **k** can be expected to depend predominantly on the degree of protonation of the triple helix. In accordance with this prediction a linear change in T_m with pH can be observed for the triple helix to coil transition between pH 5.0 and pH 6.0. At pH values below pH 5.0 the degree of protonation of the coil state will increase as the pH approaches the pka of free cytosines. Since about half the cytosines in the intramolecular triple helix are protonated at this pH, this will result in a gradual decrease in **k** and hence a gradual decrease in ΔT_m with pH. At pH 4.5, the pka of free cytosines, 50% of all cytosines in the coil state are protonate, and since 50% of the cytosines in the intramolecular triple helix are also protonated, equation (**2.14**) predicts that the T_m is maximal and that change in T_m (ΔT_m) is zero. This is indeed observed for the intramolecular triple helix to coil transition. At pH values below pH 4.5 the coil state becomes progressively more protonated, and since the intramolecular triple helix now contains fewer protonated cytosines than the coil state any change in pH will favor the coil state and hence result in a lower T_m for the intramolecular triple helix to coil transition. The $d(T_m)/d(pH)$ values of the intramolecular triple helix to coil transition from the linear region between pH 5.0 and

pH 6.0 are given in **Table X.**

The $d(T_m)/d(pH)$ Changes Linearly with Global Composition Between pH 5.0 and pH 6.0

A comparison of the $d(T_m)/d(pH)$ values in the linear region between pH 5.0 and pH 6.0 for the oligonucleotides ITS-$2G_3$, ITS-$3G_2$ and ITS-$4G_1$ listed in **Table X** reveals that the $d(T_m)/d(pH)$ values change linearly with global composition ($\%C^+GC$) (**Figure 5.3.6**). Linear least square curve fitting of the experimental $d(T_m)/d(pH)$ values to the base composition $\%C^+GC$ results in linear equation of reasonable statistical fit (equation (**5.3.4**)):

$$d(T_m)/d(pH) = -0.25 * (\%C^+GC) - 0.41 \qquad (5.3.4)$$

Extrapolation of equation (**5.3.4**) to 0 $\%C^+GC$ results in a $d(T_m)/d(pH)$ value close to that of the hairpin to coil transition of the oligonucleotide ITS-ATT above pH 6.0. Since the intramolecular triple helix of the oligonucleotide ITS-ATT does not contain any protonated cytosines in the triple helix stem, it can be assumed that a $d(T_m)/d(pH)$ value similar to that of the hairpin to coil transition of ITS-ATT would be measured if the triple helix conformation were to exist at 150 mM Na^+.

The correspondence between the experimental $d(T_m)/d(pH)$ for the hairpin to coil transition of ITS-ATT and the extrapolated $d(T_m)/d(pH)$ value for a intramolecular triple helix without any protonated cytosines thus provides confirmation for the assumption that the $d(T_m)/d(pH)$ value is directly proportional to the difference in protonation **k**, and hence that the term $(2.3*R*T_m^2/\Delta H_{obs})$ in equation (**2.14**) is independent of the global composition.

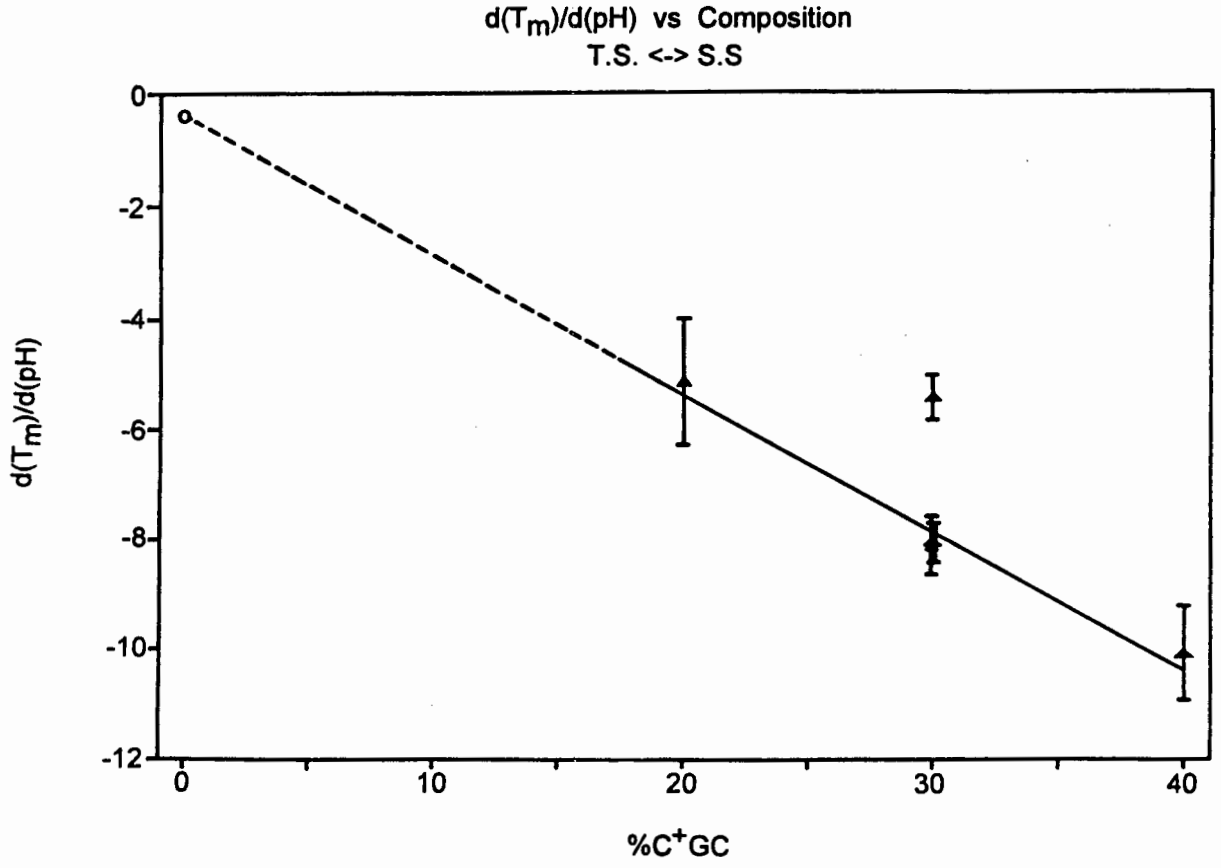


Figure 5.3.6 The effect of base composition on the change in T_m with pH ($d(T_m)/d(pH)$) of the triple helix to coil transition between pH 6.0 and pH 5.0

Similar to the change in T_m and $d(T_m)/d(\log[Na^+])$ with global composition, the slope of equation (**5.3.4**) differs considerably from that of the component triple helix to hairpin and hairpin to coil transitions. Since the hairpin to coil transition is assumed to be independent of the solution pH, only the triple helix component is expected to contribute to the slope of the intramolecular triple helix to coil transition. But the steepness of the slope in equation (**5.3.4**) is less than half that of the corresponding slope of the triple helix to hairpin transition between pH 6.0 and pH 8.0. Hence the hairpin component of the intramolecular triple helix must reduce the pH dependency of the intramolecular triple helix to coil transition.

5.3.3.2) **The Effect of Local Composition on the $d(T_m)/d(pH)$**

The effect of the local composition on the $d(T_m)/d(pH)$ value of the intramolecular triple helix to coil transitions can be estimated by comparing the $d(T_m)/d(pH)$ values of the oligonucleotides ITS-$3G_0$, ITS-$3G_1$ and ITS-$3G_2$ listed in **Table X**. Equation (**2.14**) predicts that the $d(T_m)/d(pH)$ value is directly proportional to **k**. Since all three oligonucleotide sequence variants contain the same number of cytosines in the third strand and in the loop, it can be expected that the $d(T_m)/d(pH)$ value will be independent from the local composition. The same $d(T_m)/d(pH)$ value within the margin of the experimental error is indeed observed for the oligonucleotides ITS-$3G_1$ and ITS-$3G_2$. The only oligonucleotide for which this does not hold is the oligonucleotide ITS-$3G_0$. What can be the reason for this?

The $d(T_m)/d(pH)$ value of ITS-$3G_0$ is nearly the same as that of ITS-$2G_3$, containing one protonated cytosine less. Since the $d(T_m)/d(pH)$ value for the intramolecular triple helix to coil transition is linear only between pH 5.0 and pH 6.0, it is unlikely that the differences in $d(T_m)/d(pH)$ are a reflection of differences in the degree of protonation in the coil state. Hence the similarity in the $d(T_m)/d(pH)$ value between ITS-$3G_0$ and ITS-$2G_3$ indicates that only two cytosines of ITS-$3G_0$ appear to be protonated in the triple helix state. The electrostatic repulsion between adjacent third strand cytosines has previously been shown to reduce the degree of pH dependency of the triple helix to hairpin transition.

CONCLUSION

The results of the spectroscopic and enzymatic probing of the structure of the oligonucleotides listed in **Table V** demonstrate that the single strands fold sequentially into a partially double stranded conformation and finally into a triple stranded complex. Conceptional considerations demonstrates that the random coil first folds into a hairpin with dangling 3'tail and then into an intramolecular triple helix. This sequential folding process depends on the environmental conditions such as pH, ionic strength, temperature, and on the sequence of the single strands. The oligonucleotides investigated can be grouped into two sequence families, namely those which systematically vary in their global composition, and those with different local composition but identical global composition. From a comparison of the effect of the environmental conditions (pH, ionic strength) on the thermal stability of the different conformations the following general conclusions can be drawn about the impact of global and local variation in composition:

1) The Effect of Global Composition on the Triple Helix to Hairpin Transition:
The thermal stability of the intramolecular triple helix increases linearly with the number of cytosines in the third strand at pH 6.75 (100 mM Na^+), provided that the third strand cytosines are sufficiently separated by intervening thymines. A spacer of two thymines can be considered as sufficient. In contrast to double stranded DNA, for which the T_m is also proportional to the GC content, the increase in T_m with increasing %C content for the third strand melting is due to the protonation of the cytosines. This protonation leads to a reduction in the overall electrostatic repulsion between the phosphate backbones. It leads to local attraction between third strand cytosines and the phosphate backbone, and in special cases repulsion of clustered cytosines. All these electrostatic effects depend

strongly on the environmental conditions of pH and ionic strength. The impact of the environmental conditions on members of the first family can be summarized as follows:

a) At pH 7.1 (150 mM Na^+) the thermal stability of T•A•T and C^+•G•C triplets is identical. At pH values above pH 7.1 (150 mM Na^+) cytosines contribute less to the third strand binding as compared to thymines. They destabilize the triple helix relative to a triple helix containing only T•A•T triplets, while at pH below pH 7.1 third strand cytosines stabilize the intramolecular triple helix.

b) At pH 6.75 (400 mM Na^+) the thermal stability of T•A•T and C^+•G•C triplets is identical. Increasing the ionic strength above 400 mM Na^+ (pH 6.75) leads to a destabilization of C^+•G•C triplets relative to T•A•T triplets, while at ionic strength less than 400 mM the opposite holds true.

c) A plot of T_m versus global composition (%C) always results in a straight line relationship, the slope of which depends on the environmental conditions of pH and ionic strength.

2) The Effect of Local Composition on the Triple Helix to Hairpin Transition:

The thermal stability of the triple helix depends also on the local composition. If the third strand cytosines are not at all separated by intervening thymines electrostatic repulsion between adjacent cytosines occurs. This results in a decrease in the T_m for the triple helix to hairpin transition of these oligonucleotides relative to oligonucleotide triplexes of the same global composition but with the cytosines properly spaced. This repulsive electrostatic effect falls off rapidly with the number of thymine bases separating adjacent cytosines. This can be viewed as the local composition effect. The magnitude of this effect depends strongly on the ionic strength and pH of the solution. For high ionic

strength and/or high pH a triple helix with cytosines adjacent to each other will be less stable that a triple helix containing fewer but noninteracting cytosines.

3) The Effect of Global Composition on the Triple Helix to Coil Transition:

At pH values well below pH 6.0 the intramolecular triple helix unfolds completely in a single step to the random coil. The T_m of the triple helix to coil transition changes linearly with the global composition (%C^+GC). The T_m is linearly dependent on the logarithm of the counterion concentration (log [Na^+]), with a slope d(T_m)/d(log[Na^+]) directly proportional to the global composition. At pH 4.5, the pka of free cytosine, the T_m is independent of the pH for all intramolecular triplexes.

4) The Effect of Local Composition on the Triple Helix to Coil Transition:

The thermal stability of the triple helix to coil transition is also dependent on the local composition. The placing of cytosines next to each other results in a destabilization of the triple helix compared to a triple helix of identical global composition but with the cytosines sufficiently separated. Compared to the triple helix to hairpin transition the effect of local composition on the stability of the intramolecular triple helix to coil transition is less pronounced, since it occurs closer to the pka of free cytosines.

BIBLIOGRAPHY

Amaratunga,M., Snowden-Ifft, E. Wemmer, D.E. & Benight, A.S. (1992)
Studies of DNA Dumbells. II. Construction and Characterization of DNA Dumbells with a 16 Base-Pair Duplex Stem and T_n End Loops (n=2,3,4,6,8,10,14)
Biopolymers, **32**, 865-879

Anderson, C.F. & Record, T.M. Jr. (1982)
Polyelectrolyte Theories and their Applications to DNA
Ann. Rev. Phys. Chem., **33**, 191-222

Antao, V.P., Gray, D.M. & Ratliff, R.L. (1988)
CD of six different Conformational Rearrangments of Poly d(A-G) • Poly d(C-T) induced by low pH
Nucleic Acids Res., **16**, 719-738

Arnott, S. & Bond, P.J. (1973)
Structures for Poly(U)•Poly(A)•Poly(U) Triple stranded Polynucleotides
Nature New Biology, **244**, 99-101

Arnott, S. & Bond, P.J. (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)•Poly(dA)•Poly(dT)
J. Mol. Biol., **88**, 509-521

Arnott, S., Bond, P.J., Selsing, E. & Smith, P.J.C. (1976)
Models of Triple-Stranded Polynucleotides with Optimized Stereochemistry
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)
In: Current Protocols in Molecular Biology Unit 3.10.2
J. Wiley & Sons, (1987), New York

Baumann, U., Frank, R. & Blöcker, H. (1986)
Conformational analysis of hairpin oligodeoxyribonucleotides by a single-strand-specific nuclease
Eur. J. Biochem., **161**, 409-413

Beal, P.A. & Dervan, P.B. (1991)
Second Structural Motif for Recognition of DNA by Oligonucleotide-Directed Triple-Helix Formation
Science, **251**, 1360-1363

Beasty, A.M. & Behe, M.J. (1988)
An oligopurine sequence bias occurs in eukaryotic viruses
Nucleic Acids Res., **16**, 1517-1527

Behe, M.J. (1987)
The DNA Sequence of the Human β-Globin Region Is Strongly Biased in Favor of Long Strings of Contiguous Purine or Pyrimidine Residues
Biochemistry, **26**, 7870-7875

Bernues, J., Beltran, R., Cassasnovas, J. M. & Azorin, F. (1989)
Structural Polymorphism of Homopurine-Homopyrimidine Sequences: The Secondary DNA Structure adopted by a $d(GA \bullet CT)_{22}$ sequence in the presence of zinc ions
EMBO, **8**, 2087-2094

Blaho, J.A., Larson, J.E., McLean, M.J. & Wells, R.D. (1988)
Multiple DNA Secondary Structures in Prefect Inverted Repeat Inserts in Plasmids
J. Biol. Chem., **263**, 14446-14455

Blake, R.D. & Haydock, P.V. (1979)
Effect of Sodium Ions on the High Resolution Melting of Lambda DNA
Biopolymers, **18**, 3089-3109

Blake, R.D., Massoulie, J. & Fresco, J. (1967)
Polynucleotides VIII. A Spectral Approach to the Equilibria between Polyadenylate and Polyribouridylate and their Complexes
J. Mol. Biol., **30**, 291-308

Blommers, M.J.J., Van De Ven, F.J.M., Van Der Marel, G.A., Van Boom, J.H. & Hilbers, C.W. (1991)
The three-dimensional structure of a DNA hairpin in solution
Eur. J. Biochem., **201**, 33-51

Blommers, M.J.J., Walters, J.A.L.I., Haasnoot, C.A.G., Aelen, J.M.A., van der Marel, G.A., van Boom, J.H. & Hilbers, C.W. (1989)
Effects of Base Sequence on Loop Folding in DNA Hairpins
Biochemistry, **28**, 7491-7498

Boulard. Y., Gabarro-Arpa, J., Cognet, J.A.H., Le Bret, M., Guy, A., Teoule, R., Guschlbauer, W. & Fazakerley, G.V. (1991)
The solution structure of a DNA hairpin containing a loop of three thymidines determined by nuclear magnetic resonance and molecular mechanics
Nucleic Acid Res., **19**, 5159-5167

Brahms, J. (1964)
Circular Dichroism Investigations of the Two Conformations of Polyriboadenylic Acid
Nature, **202**, 797-798

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

Broitmann, S.L., Im, D.D. & Fresco, J.R. (1987)
Formation of the triple stranded polynucleotide helix, Poly (A•A•U)
Proc. Natl. Acad. Sci. USA, **84**, 5120-5124

Brown, T., Leonard, G.A., Booth, D.E. & Kneale, G. (1990)
Influence of pH on the Conformation and Stability of Mismatch Base-pairs in DNA
J. Mol. Biol., **212**, 437-440

Byrd, C., Ohtsuka, E., Moon, M.W. & Khorana, H.G. (1965)
Synthetic Deoxyribo-Oligonucleotides as Templates for the DNA Polymerase of *Escherichia Coli*: New DNA-like Polymers containing Repeating Nuceotide Sequences
Proc. Natl. Acad. Sci. USA, **53**, 79-86

Callahan, D.E., Trapane, T.L., Miller, P.S., Ts'o, P.O.P. & Kan, L-S. (1991)
Comparative Circular Dichroism and Fluorescence Studies of Oligodeoxyribonucleotide and Oligodeoxyribonucleoside Methylphosphonate Pyrimidine Strands in Duplex and Triplex Formation
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)
Physical Studies of DNA Premelting Equilibria in Duplexes with and without Homo dA•dT Tracts: Correlations with DNA Bending
Biochemistry, **29**, 6161-6171

Chang, R.
in: Physical Chemistry with Applications to Biological Systems (2nd Ed.), Chp. 17,pp 494
McMillan Publishing Co., Inc., New York, (1981)

Charles, D. (1991)
A triple helix to cripple viruses
New Scientist, **130**, 19

Chen, F-M. (1991)
Intramolecular Triplex Formation of the Purine • Purine • Pyrimidine Type
Biochemistry, **30**, 4427-4479

Cheng Y-K. & Pettitt, B.M. (1992)
Stabilities of Double- and Triple-Helical Nucleic Acids
Prog. Biophys. Molec. Biol., **58**, 225-257

Chomilier, J., Sun, J-S., Collier, D.A., Garestier, Th., Helene, C. & Lavery, R. (1992)
A Computational and Experimental Study of the Bending Induced at a Double-Triple Helix Junction
Biophys. Chem., **42**, 143-152

Coll, M., Frederick, C.A., Wang, A.H-J. & Rich, A. (1987)
A bifurcated hydrogen-bonded conformation in the d(A•T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin
Proc. Natl. Acad. Sci. USA, **84**, 8385-8389

Collier, D.A., Mergny, J-L., Thuong, N.T. & Helene, C. (1991)
Site-Specific Intercalation at the Triplex-Duplex Junction Induces a Conformational Change which is Detectable by Hypersensitivity to Diethylpyrocarbonate.
Nucleic Acids Res., **19**, 4219-4224

Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J & Hogan, M.E. (1988)
Site-Specific Oligonucleotide Binding Represses Transcription of the Human c-*myc* Gene *in Vitro*
Science, **241**, 456-459

Cristhophe, D., Cabrer, B., Bacolla, Targovnik, H., Pohl, V. & Vassart, G. (1985)
An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene
Nucleic Acids Res., **13**, 5127-5144

de los Santos, C., Rosen, M. & Patel, D. (1989)
NMR Studies of DNA $(R^+)_n \bullet (Y^-)_n \bullet (Y^+)_n$ Triple Helices in Solution: Imino and Amino Proton Markers of T•A•T and C•G•C^+ Base-Triple Formation
Biochemistry, **28**, 7282-7289

Dervan, P.B.
Reagents for the site-specific cleavage of megabase DNA
Nature, **359**, 87-88

DeVoe, H. & Tinoco, I. Jr. (1962)
The Stability of Helical Polynucleotides: Base Contributions.
J. Mol. Biol., **4**, 500-517

Distefano, M. & Dervan, P. B. (1992)
Ligand-Promoted Dimerisation of Oligonucleotides Binding Cooperatively to DNA
J. Am. Chem. Soc., **114**, 11006-11007

Distefano, M. & Dervan, P. B. (1993)
Energetics of cooperative binding of oligonucleotides with discrete dimerisation domains to DNA by triple helix formation
Proc. Natl. Acad. Sci. USA., **90**, 1179-1183

Distefano, M.D., Shin, J.A. & Dervan, P.B. (1991)
Cooperative Binding of Oligonucleotides to DNA by Triple Helix Formation: Dimerization via Watso & Crick Hydrogen Bonds
J. Am. Chem. Soc., **113**, 5901-5902

Doktycz, M.J., Goldstein, R.F., Paner, T.M., Gallo, F.J. & Benight, A.S. (1992)
Studies of DNA Dumbells. I. Melting Curves of 17 DNA Dumbells with Different Duplex Stem Sequences Linked by T_4 Endloops: Evaluation of the Nearest Neighbor Stacking Interactions in DNA
Biopolymers, **32**, 849-864

Dolinnaya, N.G., Braswell, E.H., Fossella, J., Klump, H.H. & Fresco, J.R. (1993)
Molecular and Thermodynamic Properties of $d(A^+\text{-}G)_{10}$, a Single Stranded Helix without Base Pairing
Biochemistry, ***in press***

Durand, M., Peloille, S., Thuong, N.T. & Maurizot, J.C., (1992)
Triple Helix Formation by an Oligonucleotide Containing One $(dA)_{12}$ and Two $(dT)_{12}$ Sequences Bridged by Two Hexaethylene Glycol Chains
Biochemistry, **31**, 9197-9204

Durand, M., S., Thuong, N.T. & Maurizot, J.C. (1992)
Binding of Netropsin to a DNA Triple Helix
Biochemistry, **31**, 9197-9204

Egholm, M., Nielsen, P.E., Buchard, O. & Berg, R.H. (1992)
Recognition of Guanine and Adenine in DNA by Cytosine and Thymine Containing Peptide Nucleic Acids (PNA)
J. Am. Chem. Soc., **114**, 9677-9678

Elson, E.L., Scheffler, I.E. & Baldwin, R.L. (1970)
Helix Formation by d(TA) Oligomers. III. Electrostatic Effects
J. Mol. Biol., **54**,401-415

Erie, D., Sinha, N, Olson, W., Jones, R. & Breslauer, K. (1987)
A dumbbell-shaped, double-hairpin structure of DNA: a thermodynamic investigation.
Biochemistry, **26**, 7150-7159

Erie, D.A., Jones, R.A., Olson, W.K., Sinha, N.K. & Breslauer, K.J. (1989)
Melting behaviour of a covalent closed single-stranded, circular DNA
Biochemistry, **28**, 268-273

Erie, D.A., Suri, A.K., Bresslauer, K.J., Jones, R.A. & Olson, W.K (1993)
Theoretical Predictions of DNA Hairpin Loop Conformations: Correlation with Thermodynamic and Spectroscopic Data
Biochemistry, **32**, 436-454

Felsenfeld, G. & Miles, H.T. (1967)
The Physical and Chemical Properties of Nucleic Acids
Ann. Rev. Biochem., **36**, 407-448

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., Barbier, C., Chassignol, M., Thuong, N.T. & Helene, C. (1989)
Sequence-specific recognition and cleavage of duplex DNA via triple-helix formation by oligonucleotides covalently linked to a phenanthroline-copper chelate
Proc. Natl. Acad. Sci. USA., **86**, 9702-9706

Francois, J-C., Saison-Behmoaras, T., Chassignol, M., Thuong, N.T. & Helene, C. (1989)
Sequence-targeted Cleavage of single- and double-stranded DNA by Oligothymidylates Covalently Linked to 1,10-Phenanthroline
J. Biol. Chem., **264**, 5891-5898

Francois, J-C., Saison-Behmoaras, T.,Thuong, N.T. & Helene, C. (1989)
Inhibition of Restriction Endonuclease Cleavage via Triple Helix Formation by Homopyrimidine Oligonucleotides
Biochemistry, **28**, 9617-96-19

Fresco, J.R. & Klemperer, E. (1959)
Polyriboadenylic Acid, A Molecular Analogue of Ribonucleic Acid and Desoxyribonucleic Acid
Ann. N.Y. Acad. Sci., **81**, 730-741

Fresco, J.R. & Massoulie, J. (1963)
Polynucleotides V: Helix-Coil Transition of Polyriboguanylic Acid
J. Am. Chem. Soc., **85**, 1352-1353

Froehler, B.C. & Ricca, D.J. (1992)
Triple-Helix Formation by Oligodeoxynucleotides Containing the Carbolic Analogs of Thymidine and 5-Methyl-2'-deoxycytidine
J. Am. Chem. Soc., **114**, 8320-8322

Giovannangeli, C., Rougée, M., Garestier, Thuong, N.T. & Hélène, C. (1992)
Triple-helix formation by oligonucleotides containing the three bases thymine, cytosine and guanine
Proc. Natl. Acad. Sci. USA, **89**, 8631-8635

Giovannangeli, C., Thuong, N.T. & Helene C. (1992)
Oligodeoxynucleotide-directed photo-induced cross-linking of HIV proviral DNA via triple-helix formation
Nucleic Acids Res., **20**, 4275-4281

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

Griffin, L.C., Kiessling, L.L., Beal, P.A., Gillespie, P. & Dervan P.B. (1992)
Recognition of All Four Base Pairs of Double-Helical DNA by Triple Helix Formation: Design of Nonnatural Deoxyribonucleosides for Pyrimidine• Purine Base Pair Binding
J. Am. Chem. Soc., **114**, 7976-7982

Griffin, L.C., Kiessling, L.L., Beal, P.A., Gillespie, P. & Dervan, P.B (1992)
Recognition of All Four Base Pairs of Double Helical DNA by Triple-Helix Formation: Design of Nonnatural Deoxyribonucleosides for Pyrimidine•Purine Base Pair Binding
Biochemistry, **31**, 2829-2834

Grunberg-Manago, M., Oritz, P. & Ochoa, S. (1956)
Enzymatic Synthesis of Polynucleotides I. Polynucleotide Phosphorylase of *Azotobacter Vinelandii*
Biochem. Biophys. Acta., **20**, 269-285

Guschlbauer, W.
Melting Temperatures of Polynucleotide Complexes
in: Lanold Börnstein,New Series VII/Ic, Saenger W.(ed), Springer Verlag Berlin (1990)

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

Hanvey, J.C., Shimizu, M. & Wells, R.D. (1988)
Intramolecular DNA triplexes in supercoiled plasmids
Proc. Natl. Acad. Sci. USA, **85**, 6292-6296

Hanvey, J.C., Shimizu, M. & Wells, R.D. (1990)
Site-specific inhibition of *EcoR1* restriction/modification enzymes by a DNA triple helix
Nucleic Acids Res., **18**, 157-161

Hare, D.R.. & Reid, B.R. (1986)
Three-dimensional Structure of a DNA Hairpin in Solution: Two-dimensional NMR Studies and Distance Geometry Calculations
Biochemistry, **25**, 5341-5350,

Hattori, M. Frazier, J. & Miles, T.D. (1976)
The Structure of the Triple Stranded G-2C Polynucleotide Helices
Biopolymers, **15**, 523-531

Helene, C. & Toulme, J-J. (1990)
Specific regulation of gene expresion by antisense, sense and antigene nucleic acids
Biochem. Biophys. Act., **1049**, 99-125

Helene, C. Thuong, N.T., Saison-Behmoaras & Francois, J-C. (1989)
Sequence-specific artificial endonucleases
TIBTECH, **7**, 310-315

Hoogsteen, K., (19959
The structure of crystals containing a hydrogenbonded complex of 1-methylthymine and 9-methyladenine
Acta Cryst., **12**, 822-823

Horne, D.A. & Dervan, P.B. (1991)
Effects of an abasic site on the triple helix formation characterized by affinity cleaving
Nucleic Acids Res., **19**, 4963-4965

Howard, F.B. & Miles, H.T. (1977)
Interaction of Poly(A) and Poly(I). A Reinvestigation
Biochemistry, **16**, 4647-4650

Howard, F.B., Miles, H.T., Liu, K., Frazier, J., Raghunathan, G., Sasisekharan, V. (1992)
Structure of $d(T)_n \bullet d(A)_n \bullet d(T)_n$: The DNA TripleHelix Has B-Form Geometry with *C2'-Endo* Sugar Pucker
Biochemistry, **31**, 10671-10677

Huang, C-C., Nguyen, D., Martinez, R. & Edwards, C.A.(1992)
Triple-Helix Formation is Compatible with an Adjacent DNA-Protein Complex
Biochemistry, **31**, 993-998

Inman, R.B. (1964)
Multistranded DNA Homopolymer Interactions
J. Mol. Biol., **10** , 137-146

Ito, T., Smith, C.L. & Cantor, C.R.
Sequence-specific DNA Purification by Triplex Affinity Capture
Proc. Natl. Acad. Sci. USA., **89**, 495-498

Jayasena, S.D. & Johnston, B.H. (1992)
Oligonucleotide directed Triple Helix Formation at Adjacent Oligopurine and Oligopyrimidine DNA Tracts by Alternate Strand Recognition.
Nucleic Acids Res., **20**, 5279-5288

Jayasena, S.D. & Johnston, B.H. (1992)
Intramolecular Triple Helix Formation at $(Pu_nPy_n) \bullet (Pu_nPy_n)$ Tracts: Recognition of Alternate Strands via Pu•PuPy and Py•PuPy Base triplets
Biochemistry, **31**, 320-327

Jenkins, Y. & Barton, J.K. (1992)
A Sequence-Specific Molecular Light Switch: Tethering of an Oligonucleotide to a Dipyridophenazine Complex of Ruthenium(II)
J. Am. Chem. Soc., **114**, 8736-8738

Johnson, W.C.
Electronic Circular Dichroism (CD) Spectroscopy of Nucleic Acids
in: Lanold Börnstein,New Series VII/Ic, Saenger W.(ed), Springer Verlag Berlin (1990)

Johnson, D. & Morgan, A.R. (1978)
Unique Structures Formed by Pyrimidine•Purine DNAs which may be Four Stranded
Proc. Natl. Acad. Sci. USA., **75**, 1637-1641

Johnson, K.H., Gray, 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

Kan, L-S., Callahan, D.E., Trapane, T.L., Miller, P.S., Ts'o, P.O.P. & Huang, D.H. (1991)
Proton NMR and Optical Spectroscopic Studies on the DNA Triplex formed by d-A-(G-A)7-G and d-C-(T-C)7-T
J. Biomol. Struc. Dyn., **8**, 911-933

Kiessling, L.L., Griffin, L.C. & Dervan, P.B (1992)
Flanking Sequence Effects within the Pyrimidine Triple-Helix Motif Characterized by Affinity Cleaving
Biochemistry, **31**, 2829-2834

Klump H.H (1988)
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 (2nd Ed.) by M.N.Jones (Ed.), Elsevier, New York

Kool, E.T. (1991)
Molecular Recognition by Circular Oligonucleotides: Increasing the Selectivity of DNA Binding
J. Am. Chem. Soc., **113**, 6265-6266

Krakauer, H. & Sturtevant, J.M. (1968)
Heats of the Helix-Coil Transition of the Poly A - Poly U Complexes
Biopolymers, **6**,491-512

Le Pecq, J.B. & Paoletti, C. (1967)
A Fluorescent Complex between Ethidium Bromide and Nucleic Acids
J. Mol. Biol., **27**, 87-106.

Lee, J.S., Johnon, D.A. & Morgan, R.A. (1979)
Complexes formed by $(pyrimidine)_n \bullet (purine)_n$ DNA's on lowering the pH are Three-stranded.
Nucleic Acids Res., **6**, 3073-3091,

Lee, J.S., Woodsworth, M.L., Latimer, L.J.P. & Morgan, A.R. (1984)
Poly(pyrimidine)•Poly(purine) Synthetic DNA's Containing 5-Methylcytosine form Stable Triplexes at Neutral pH
Nucleic Acids Res., **12**, 6603 - 6614

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

Lin, C.H. & Patel, D.J. (1992)
Site-Specific Covalent Duocarmycin A-Intramolecular DNA Triplex Complex
J. Am. Chem. Soc., **114**, 110658-10660

Lipsett, M.N. (1964)
Complex Formation between Polycytidylic Acid and Guanine Oligonucleotides
J. Biol. Chem., **239**, 1256-1260

Liquier, J., Coffinier, P., Firon, M. & Taillandier, E. (1991)
Triple Helical Polynucleotidic Structures: Sugar Conformations Determined by FTIR Spectroscopy
J. Biomol. Struct. Dynam., **9**, 437-445

Luebke, K.J. & Dervan, P.B. (1992)
Nonenzymatic Ligation of Double-helical DNA by Alternate Strand Triple Helix Formation
Nucleic Acids Res., **20**, 3005-3009

Luebke, K.J. & Dervan, P.B. (1992)
Nonenzymatic Sequence-specific Ligation of Double-Helical DNA
J. Am. Chem. Soc., **113**, 7447-7448

Luebke, K.J. & Dervan, P.B. (1992)
Nonenzymatic Ligation of Oligodeoxyribonucleotides on a Duplex DNA Template by Triple-Helix Formation
J. Am. Chem. Soc., **113**, 7447-7448

Lyamichev, V.I, Mirkin, S.M. & Frank-Kamenetskii, M.D. (1985)
A pH dependent Structural Transition in the Homopurine-Homopyrimidine Tract in Superhelical DNA
J. Biomolec. Struct. Dyn., **3**,327-338,

Lyamichev, V.I, Mirkin, S.M. & Frank-Kamenetskii, M.D. (1986)
Structures of Homopurine-Homopyrimidine Tract in Superhelical DNA
J. Biomolec. Struct. Dyn., **3**, 667-669,

Macaya, R., Schultz, P. & Feigon, J. (1992)
Sugar Conformation in Intramolecular DNA Triplexes Determined by Coupling Constants Obtained by Automated Simulation of P.Cossy Cross Peaks
J. Am. Chem. Soc., **114**, 781-783

Macaya, R., Wang, E., Schultz, P., Sklenar, V. & Feigon, J. (1992)
Proton Nuclear Magnetic Resonance Assignment and Structural Chararcterisation of an Intramolecular DNA Triplex
J. Mol. Biol., **225**, 755-773

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 Intramolecular Triplexes
Science, **254**, 270-274

Maher III, L.J. (1992)
DNA Triple-Helix Formation: An Aproach to Artificial Gene Repressors
BioEssays, **14**, 807-815

Maher III, L.J., Dervan P.B. & Wold, B. (1992)
Analysis of Promotor-Specific Repression by Triple-Helical DNA Complexes in a Eucaryotic Cell-Free Transcription System
Biochemistry, **31**, 70-81

Maher III, L.J., Dervan, P.B. & Wold, B.J. (1990)
Kinetic Analysis of Oligodeoxyribonucleotide-Directed Triple Helix Formation on DNA
Biochemistry, **29**, 8820-8826

Maher III, L.J., Wold, B. & Dervan P.B. (1989)
Inhibition of DNA Binding Proteins by Oligonucleotide-Directed Triple Helix Formation
Science, **245**, 725-730

Manning, G.S. (1978)
The Molecular Theory of Polyelectrolyte Solutions with Applications to the Electrostatic Properties of Polynucleotides
Quart.Rev.Biophys., **11**, 179-246

Manzini, G., Xodo, L.E., Gasparotto, D., Quadrifoglio, F., van der Marel, G.A. & van Boom, J.H. (1990)
Triple Helix Formation by Oligopurine-Oligopyrimidine DNA Fragments
J. Mol. Biol., **213**, 833-843

Marck, C. & Thiele, D. (1978)
Poly(dG)•poly(dC) at neutral and alkaline pH: The formation of triple stranded poly(dG)•poly(dG)•poly(dC)
Nucleic Acid Res., **5**, 1017-1028

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

Marmur, J. & Doty, P. (1962)
Determination of the Base Composition of Deoxyribonucleic Acid from its Thermal Denaturation Temperature
J. Mol. Biol., **5**, 109-118

Massoulie (1968a)
Thermodynamique des associations de poly A et poly U en milieu neutre et alcaline
Eur. J. Biochem., **3**, 428-438

Massoulie (1968b)
Associations de poly A et poly U en milieu acide
Eur. J. Biochem., **3**, 439-447

McGhee, J.D. & von Hippel, P.H. (1974)
Theoretical Aspects of DNA-Protein Interactions: Co-operative and Non-co-operative Binding of Large Ligands to a One-dimensional Homogeneous Lattice
J. Mol. Biol., **86**, 469-489

Mergny, J-L., Collier, D., Rougee, M., Montenay-Garestier, Th. & Helene, C. (1991)
Intercalation of Ethidium Bromide into a Triple-Stranded Oligonucleotide
Nucleic Acids Res., **19**, 1521-1526

Mergny, J.L., Duval-Valentin, G., Nguyen, C.H., Perrouault, L., Faucon, B., Rougee, M., Montenay-Garestier, T., Bisagni, E. & Helene, C. (1992)
Triple Helix-Specific Ligands
Science, **256**, 1681-1684

Mergny, J.L., Sun, J-S., Rougée, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J. & Hélène, C (1991)
Sequence Specificity in Triple-Helix Formation: Experimental and Theoretical Studies of the Effect of Mismatches on Triplex Stability
Biochemistry, **30**, 9791-9798

Michelson, A.M., Massoulié, J. & Guschlbauer, W (1967)
Synthetic Polynucleotides
Prog in Nucleic Acid Res & Molec. Biol., **6**, 83-141

Miles, H.T. & Frazier, J. (1964a)
Infrared Study of Helix Strandedness in the Poly A-Poly U System
Biochem. Biophys Res. Com., **14**, 21-28,

Miles, H.T. & Frazier, J. (1964b)
A Strand Dispropotionation Reaction in a Helical Polynucelotide System
Biochem. Biophys Res. Com., **14**, 129-136,

Miles, H.T. (1960)
Infrared Spectra of the Three-Stranded Helices formed by Polyuridilic Acid with Polyadenylic Acid and with Tetraadenylic Acid
Biochim, Biophys. Acta., 45, 196-198

Miles, H.T. (1964)
The Structure of the Three-Stranded Helix, Poly (A+2U)
Proc. Natl. Acad. Sci. USA, **51**, 1104-1109

Milligan, J.F., Krawczyk, S.H., Wadwani, S. & Matteucci, M.D. (1993)
An anti-parallel triple helix motif with oligodeoxynucleotides containing 2'-deoxyguaninosine and 7-deaza-2'-deoxy-xanthosine
Nucleic Acids Res., **21**, 327-333

Mooren, M.M.W., Pulleyblank, D.E., Wijmenga S.S., Blommers, M.J.J. & Hilbers, C.W. (1990)
Polypurine/Polypyrimidine Hairpins Form a Triple Helix Structure at Low pH
Nucleic Acids Res., **18**, 6523-6529

Morgan., A.R. & Wells, R.D. (1968)
Specificity of the Three-stranded Complex Formation between Double-stranded DNA and Single-stranded RNA containing Repeating Nucleotide Sequences
J. Mol. Biol., **37**, 63-80

Moser, H.E. & Dervan, P.B. (1987)
Sequence-Specific Cleavage of Double Helical DNA by Triple Helix Formation
Science, **238**, 645-650

Nadeau, J.G & Crothers D.M. (1989)
Structural basis for DNA bending
Proc. Natl. Acad. Sci. USA, **86**, 2622-2626

Narang, S.A.
Synthesis and Applications of DNA and RNA
Academic Press Inc., New York, 1987

Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A (1987)
The structure of an oligo(dA)•oligo(dT) tract and its biological implications
Nature (London), **330**, 221-226

Neumann, E. & Ackermann, Th. (1969)
Thermodynamic Investigation of the Helix-Coil Transition of a Polyribonucleotide System
J. Phys. Chem., **73**, 2170-2178

Ohms, J.. & Ackermann, Th (1990)
Thermodynamics of Double and Triple-Helical Aggrgates Formed by Self-Complementary Oligoribonucleotides of the Type rA_XU_y
Biochemistry, **29**, 5237-5243

Olmsted, M.C., Anderson, C.F. & Record, T.M.Jr. (1989)
Monte Carlo Description of Oligoelectrolyte Properties of DNA Oligomers: Range of End effects and the Approach of Molecular and Thermodynamic Properties to the Polyelectrolyte Limits
Proc. Natl. Acad. Sci. USA, **86**, 7766-7770

Olmsted, M.C., Anderson, C.F. & Record, T.M.Jr. (1991)
Importance of Oligoelectrolyte End Effects for the Thermodynamics of Conformational Transitions of Nucleic Acid Oligomers: A Grand Canonical Monte Carlo Analysis
Biopolymers, **31**, 1593-1604

Ono, A., Ts'o, P.O.P. & Kan, L-S. (1991)
Triplex Formation of Oligonucleotides Containing 2'-O-Methylpseudoisocytidine in substitution for 2'-Deoxycytidine
J. Am. Chem. Soc., **113**, 4032-4033

Orson, F.M., Thomas, D.W, McShan, W.M., Kessler, D.J. & Hogan, M.E. (1991)
Oligonucleotide Inhibition of IL2Rα mRNA Transcription by Promoter Region Collinear Triplex Formation in Lymphocytes
Nucleic Acids Res., **19**, 3435-3441

Owen, R.J., Hill, L.R. & Lapage, S.P. (1969)
Determination of DNA Base Composition from Melting Profiles in Dilute Buffers
Biopolymers, **7**, 503-516

Paner, T.M., Amaratunga,M. & Benight, A.S. (1992)
Studies of DNA Dumbells. III. Theoretical Analysis of Optical Melting Curves of Dumbells with a 16 Base-Pair Duplex Stem and T_n End Loops (n=2,3,4,6,8,10,14)
Biopolymers, **32**, 865-879

Park, Y-W. & Breslauer, K.J. (1992)
Drug binding to Higher Ordered DNA Structures: Netropsin Complexation with a Nucleic Acid Triple Helix
Proc. Natl. Acad. Sci. USA., **89**, 6653-6657

Pei, D., Corey, D.R. & Schultz, P.G. (1990)
Site-specific Cleavage of Duplex DNA by Semisynthetic Nucleases via Triple-Helix Formation
Proc. Natl. Acad. Sci. USA., **87**, 9858-9862

Pei, D., Ulrich, H.D. & Schultz, P.G. (1991)
A Combinatorial Approach Toward DNA Recognition
Science, **253**, 1408-1411

Perrouault, L., Asseline, U., Rivalle, C., Thuong, N.T., Bisagni, E., Giovannangeli, C., Le Doan, T. & Helene, C. (1990)
Sequence-Specific Artificial Photoinduced Endonucleases based on Triple Helix-Forming Oligonucleotides
Nature, **344**, 358-360

Pettijohn, D.E. & Hecht, R. (1973)
RNA Molecules Bound to the Folded Bacterial Genom Stabilize DNA Folds and Segregate Domains of Supercoiling
Symp. on Quantitative Biology, **38**, 31-41

Pilch, D.S., Brousseau, R. & Shafer, R.H. (1990)
Thermodynamics of Triple Helix Formation: Spectrophotometric studies on the $d(A)_{10}$•$d(T)_{10}$ and $d(C^{+}_{3}T_4C^{+}_{3})$•$d(G_3A_4G_3)$•$d(C_3T_4C_3)$ triple helices
Nucleic Acids Res., **18**, 5743--5750

Pilch, D.S., Levenson, C. & Shafer, R.H. (1990)
Structural Analysis of the $(dA)_{10}$•$2(dT)_{10}$ Triple Helix
Proc. Natl. Acad. Sci. USA, **87**, 1942-1946

Pilch, D.S., Levenson, C. & Shafer, R.H. (1991)
Structure, Stability, and Thermodynamics of a Short Intermolecular Purine-Purine-Pyrimidine Triple Helix
Biochemistry, **30**, 6081-6087

Plum, E.G., Park, Y-W., Singleton, S.F., Dervan, P.B. & Breslauer, K.J. (1990)
Thermodynamic characterisation of the stability and the melting behaviour of a DNA triplex: A spectroscopic and calorimetric study
Proc. Natl. Acad. Sci. USA, **87**, 9436-9440

Povsic, T.J. & Dervan, P.B (1990)
Sequence-Specific Alkylation of Double-Helical DNA by oligonucleotide-directed Triple-Helix Formation
J. Am. Chem. Soc., **112**, 9428-9430

Povsic, T.J. & Dervan, P.B. (1989)
Triple Helix Formation by Oligonucleotides on DNA Extended to the Physiological pH Range
J. Am. Chem. Soc., **111**, 3059-3061

Povsic, T.J., Strobel, S.A. & Dervan, P.B (1992)
Sequence-Specific Double-Strand Alkylation and Cleavage of DNA mediated by Triple Helix Formation
J. Am. Chem. Soc., **114**, 5934-5941

Prakash, G. & Kool, E.T (1991)
Molecular Recognition by Circular Oligonucleotides. Strong Binding of Single-stranded DNA and RNA
J. Chem. Soc. Chem. Commun., **1991**, 1161-1163

Prakash, G. & Kool, E.T (1992)
Structural Effects in the Recognition of DNA by Circular Oligonucleotides
J. Am. Chem. Soc., **114**, 3523-3527

Privalov, P.L. & Potekin, S.A. (1986)
Scanning Microcalorimetry in Studying Temperature Induced Changes in Proteins
Methods Enzymol., **131**, 4-51

Pulleyblank, D.E, Glover, M., Farah, C. & Haniford, D.B. (1988)
The Specificity of "Single Strand Specific Endonucleases": Probes of Phosphordiester Conformation in Double Stranded Nucleic Acids. Lefthanded Polypurine • Polypyrimidine Structures. Long Range Transmission of Conformational Information in DNA.
In: *Unusual DNA Structures*, (Wells, R. & Harvey, S.C., eds.) pp. 23-44, Springer Verlag, New York

Radding, C.M., Josse, J. & Kornberg, A. (1962)
Enzymatic Synthesis of Deoxyribonucleic Acid
XII. A Polymer of Deoxyguanylate and Deoxycytidylate
J. Biol. Chem., **237**, 2869-2876

Radhakrishnan, I. & Patel, D.J. (1991)
NMR Assignment Strategy for DNA Protons through Three-Dimensional Proton-Proton Conectivities. Application to an Intramolecular DNA Triplex
J. Am. Chem. Soc., **113**, 8542-8544

Radhakrishnan, I. & Patel, D.J. (1992)
Solution Conformation of a G•TA Triple in an Intramolecular Pyrimidine•Purine•Pyrimidine DNA Triplex
J. Am. Chem. Soc., **114**, 6913-6915

Radhakrishnan, I., de los Santos, C. & Patel, D.J. (1991)
Nuclear Magnetic Resonance Structural Studies of Intramolecular Purine•Purine•Pyrimidine DNA Triplexes in Solution
J. Mol. Biol., **221**, 1403-1418

Radhakrishnan, I., Gao, X. & Patel, D.J. (1991)
Three-DimensionalHomonuclear NOSEY-TOCSY of an Intramolecular Pyrimidine•Purine•Pyrimidine DNA Triplex Containing a Central G•TA Triplet: Nonexchangable Proton Assignments and Structural Implications
Biochemistry, **31**, 2514-2523

Radhakrishnan, I., Gao, X., de los Santos, C., Live, D. & patel, D.J. (1991)
NMR Structural Studies of Intramolecular $(Y^+)_n \bullet (R^+)_n \bullet (Y^-)_n$ DNA Triplexes in Solution: Imino and Amino Proton and Nitrogen Markers of G•TA Base Triple Formation
Biochemistry, **30**, 9022-9030

Radhakrishnan, I., Patel, D. & Gao, X. (1991)
NMR Assignment Strategy for DNA Protons through Three-Dimensional Proton-Proton Connectiveties. Application to an Intramolecular DNA Triplex
J. Am. Chem. Soc., **113**, 8542-8544

Raghunathan, G., Miles, H.T., Sasisekharan, V. (1993)
Symmetry and Molecular Structure of a DNA Triple Helix: $d(T)_n \bullet d(A)_n \bullet d(T)_n$
Biochemistry, **32**, 455-462

Rajagopal, P. & Feign, J. (1989)
NMR Studies of Triple-Strand Formation from the Homopurine-Homopyrimidine Deoxyribonucleotide $d(GA)_4$ and $d(TC)_4$
Biochemistry, **28**, 7859-7870

Rajagopal, P. & Feign, J. (1989)
Triple Strand Formation in the Homopurrine-Homopyrimidine DNA oligonucleotides $d(G\text{-}A)_4$ and $d(T\text{-}C)_4$
Nature, **339**, 637-640

Rawitcher, M.A., Ross, P.D. & Sturtevant, J.M. (1963)
The Heat of the Reaction between Polyriboadenylic Acid and Polyribouridylic Acid
J. Am. Chem. Soc., **85**, 1915-1918

Record, M.T. Jr. (1967)
Elecectrostatic Effects on Polynucleotide Transitions. II. Behaviour of Titrated Systems
Biopolymers, **5**, 993-1008

Record, M.T. Jr., Mazur, S.J., Melancon, P., Roe, J-H., Shaner, .L & Unger, L. (1981)
Double Helical DNA: Conformations, Physical Properties, and Interactions with Ligands
Ann. Rev. Biochem., **50**, 997-1024

Record, T.M. Jr.& Lohman, T.M., (1978)
A semiempirical Extension of Polyelectrolyte theory to the Treatment of Oligoelectrolytes: Application to Oligonucleotide Helix-Coil Transitions
Biopolymers, **17**,159-166

Record, T.M. Jr., Lohman, T.M. & de Haseth, P. (1976)
Ion Effects on Ligand-Nucleic Acid Interactions
J. Mol. Biol. **107**, 145.158

Rich, A (1960)
A Hybrid Helix Containing both Deoxyribose and Ribose Polynucleotides and its Relation to the Transfer of Information between Nucleic Acids
Proc. Natl. Acad. Sci. USA, **46**, 1044-1053

Rich, A. (1958)
Formation of Two- and Three-Stranded Helical Molecules by Polyinosinic Acid and Polyadenylic Acid
Nature, **181**, 521-525

Riley, M., Maling, B. & Chamberlin, M.J. (1966)
Physical and Chemical Characterization of Two and Three stranded Adenine-Thymine and Adenine-Uracil Homopolymer Complexes
J. Mol. Biol., **20**, 359-389

Riordan,M.L. & Martin, J.C. (1991)
Oligonucleotide-based Therapeutics
Nature, **350**, 442-443

Roberts, R.W. & Crothers, D.M. (1991)
Specificity and Stringency in DNA Triple Formation
Proc. Natl. Acad. Sci. USA, **88**, 9397-9401

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,

Ross, C., Samuel, M. & Broitman, S.L. (1992)
transcriptional Inhibition of the Bacteriophage T_7 Early Promoter Region by Oligonucleotide Triple Helix Formation
Biochem. Biophys. Res. Comm., **189**, 1674-1680

Ross, P.D. & Scruggs, R.L. (1965)
Heat of the Reaction Forming the Three-Stranded Poly (A+2U) Complex
Biopolymers, **3**, 491-496

Rougee, M., Faucon, B., Mergny, J.L., Barcelo, F., Giovannangeli, C.,Garestier, T. & Helene, C. (1992)
Kinetics and Thermodynamics of Triple-Helix Formation: Effects of Ionic Strength and Mismatches
Biochemistry, **31**, 9269-9278

Rougée, M., Faucon, B., Mergny, J.L., Barcelo, F., Giovannangeli, C., Garestier, T. & Hélène, C. (1992)
Kinetics and Thermodynamics of Triple-Helix Formation: Effects of Ionic Strength and Mismatches
Biochemistry, 31, 9269-9278

Sänger, W.
Principle of Nucleic Acid Structure
Springer-Verlag, New York (1984)

Scaria, P.V. & Shafer, R.H. (1991)
Binding of Ethidium Bromide to a DNA Triple Helix
J. Biol. Chem., **2666**, 5417 - 5423

Schneider, B., Cohen, D. & Berman, H.M (1992)
Hydration of DNA Bases: Analysis of Crystalographic Data
Biopolymers, **32**, 725-750

Senior, M.M., Jones, R.A. & Breslauer, K.J. (1988)
Influence of Loop Residues on the Relatve Stabilities of DNA Hairpin Structures
Proc. Natl. Acad. Sci. USA, **85**, 6242-6246

Shaw, J-P., Milligan, J.F., Krawczyk, S.H. & Metteucci, M. (1991)
Specific, High-Efficiency, Triple-Helix-Mediated Cross-linking to Duplex DNA
J. Am. Chem. Soc., **113**, 7765-7766

Shea, R.G., Ng, P. & Bischofberger, N. (1990)
Thermal Denaturation Profiles and Gel Mobility Shift Analysis of Oligodeoxynucleotide Triplexes
Nucleic Acids Res., **18**, 4859-4866

Singleton, S.F. & Dervan, P.B. (1992)
Thermodynamics of Oligodeoxyribonucleotide-Directed Triple Helix Formation: An Analysis Using Quantitive Affinity Cleavage Titration
J. Am. Chem. Soc., **114**, 6957-6965

Singleton, S.F. & Dervan, P.B. (1992)
Influence of pH on the Equilibrium Association Constants for Oligodeoxyribonucleotide-Directed Triple Helix Formation at Single DNA Sites
Biochemistry, **31**, 10995-11003

Sklenar, V., & Feigon, J. (1990)
Formation of a Stable Triplex from a Single DNA Strand
Nature, **345**, 836-838

Steely, H.T., Gray, D.M. & Ratliff, R.L. (1986)
CD of Homopolymer DNA-RNA hybrid duplexes and triplexes containing A•T or A•U Basepairs
Nucleic Acids Res., **14**, 10071-10090

Stevens, C.L. & Felsenfeld, G. (1964)
The Conversion of Two-Stranded Poly (A+U) to Three-Stranded Poly (A+2U) and Poly A by Heat
Biopolymers, **2**, 293-314

Stilz, H.U. & Dervan, P.B. (1993)
Specific recognition of CG Basepairs by 2-Deoxynebularine within the Purine•Purine•Pyrimidine Triple-Helix Motif
Biochemistry, **32**, 2177-2185

Strobel, S.A. & Dervan, P.B. (1990)
Site-Specific Cleavage of a Yeast Chromosome by Oligonucleotide-Directed Triple-Helix Formation
Science, **249**, 73-75

Strobel, S.A. & Dervan, P.B. (1991)
Single-site Enzymatic Cleavage of Yeast Genomic DNA mediated by Triple-Helix Formation
Nature, **350**, 172-174

Strobel, S.A., Doucette-Stamm, L.A., Riba, L., Houseman, D.E. & Dervan, P.B. (1991)
Site-Specific Cleavage of Human Chromosome 4 mediated by Triple-Helix Formation
Science, **254**, 1639-1642

Strobel, S.A., Moser, H.E. & Dervan, P.B. (1988)
Double-Strand Cleavage of Genomic DNA at a Single Site by Triple-Helix Formation
J. Am. Chem. Soc., **110**, 7927-7929

Sun, J-S., Lavery, R., Chomilier, J., Zakrzewska, K., Montenay-Garestier, Th. & Helene, C. (1991)
Theoretical Study of Ethidium Intercalation in Triple Stranded DNA and at Triplex-duplex Junctions
J. Biomol. Struc. Dyn., **9**, 425-436

Sun, J-S., Mergny, J.L.,Lavery, R., Montenay-Garestier, T., & Hélène, C (1991)
Triple Helix Structures: Sequence dependence, Flexibility and Mismatch effects
J. Biomolec. Struc. Dyn., **9**, 411-424

Takasugi, M., Guendouz, A., Chassignol, M., Decout, J.L., Lhomme, J., Thuong, N.T. & Helene C. (1992)
Sequence-specific Photo-induced Cross-linking of two Strands of Double-helical DNA by Psoralen Covalently Linked to a Triple Helix Forming Oligonucleotide
Proc. Natl. Acad. Sci. USA., **88** , 5602-5606

Thiele, D. & Guschlbauer, W., (1968)
Evidence for a Three Stranded Complex between Poly I and Poly C
FEBS Letters, **1**, 173-175

Thiele, D. & Guschlbauer, W., (1969)
Polynucléotides Protonés. VII. Transitions Thermiques entre Differents Complexes de l'Acide Polyinosinique et de l'Acide Polycytidylique en Milieu Acide
Biopolymers, **8**, 361-378

Thiele, D.,, Marck, C., Schneider, C.& Guschlbauer, W. (1978)
Protonated polynucleotides structures - 23. The Acid - Base Hysteresis of poly(dG)•poly(dC)
Nucleic Acids Res., **5**, 1997-2012

Umemoto, K., Sarma, M.H., Gupta, G., Luo, J. & Sarma, R.H. (1990)
Structure and Stability of a DNA Triple Helix in Solution: NMR Studies on $d(T)_6$ • $d(A)_6$ • $d(T)_6$ and its Complex witha Minor Groove Binding Drug
J. Am. Chem. Soc., **112**, 4539-4545

van de Sande, J.H., Ramsing, N.B., Germann, M.W., Elhorst, W., Kalisch, B.W., von Kitzing, E., Pon, R.T., Clegg, R.C. & Jovin, T.M. (1988)
Parallel Stranded DNA
Science, **241**, 551-557

Völker, J., Botes, D.P., Lindsey, G.G & Klump, H.H. (1993)
Energetics of a Stable Intramolecular DNA Triple Helix Formation
J. Mol. Biol., **230**, 1278-1290

Wang, E., Malek, S. & Feigon, J. (1992)
Structure of a G•T•A Triplet in an Intramolecular DNA Triplex
Biochemistry, **31**, 4838-4846

Warner, R.C. (1957)
Studies on Polynucleotides Synthesized by Polynucleotide Phosphorylase III. Interaction and Ultraviolet Absorption
J. Biol. Chem., **229**, 711-724

Watson, J.D. & Crick, F.H.C. (1953)
Genetical Implications of the Structure of Deoxyribonucleic Acids
Nature, **171**, 964-967,

Wells, R.D. (1988)
Unusual DNA Structures
J. Biol. Chem., **263**, 1095-1098

Wells. R.D., Larson, J.E., Grant, R.C., Shortle, B.E. & Cantor, C.R.
Physicochemical Studies on Polyribonucleotides Containing Defined Repeating Nucleotide Sequences
J. Mol. Biol., **54**, 465-497,

Xodo, L.E., Manzini, G. & Quadrifoglio, F. (1990)
Spectroscopic and Calorimetric Investigation on the DNA Triplex Formed by d(CTCTTCTTTCTTTTCTTTCTTCTC) and d(GAGAAGAAAGA) at Acidic pH
Nucleic Acids Res., **18**, 3557-3564

Xodo, L.E., Manzini, G., Quadrifoglio, F., van der Marel, G.A. & van Boom, J.H. (1991)
Effect of 5-Methylcytosine on the Stability of Triple-Stranded DNA - a Thermodynamic Study
Nucleic Acids Res., **19**, 5625-5631

Yoon, K., Hobbs, C.A., Walter, A.E. & Turner, D.H. (1993)
Effect of a 5'-Phosphate on the Stability of a Triple Helix
Nucleic Acids Res., **21**, 601-606

Zimmer, Ch. & Wähnert, U. (1986)
Nonintercalating DNA-Binding Ligands: Specificity of the Interaction and their Use as Tools in Biophysical, Biochemical and Biological Investigations of the genetic Material
Prog. Biophys. Molec. Biol., **47**, 31-112

Zubay, G. (1962)
A Theory on the Mechanism of Messenger-RNA Synthesis
Proc. Natl. Acad. Sci. USA, **48**, 456-461

APPENDIX

APPENDIX A

T_m versus Ionic Strength at pH 8.0

[1] Phasediagram: T_m vs log $[Na^+]$ for ITS-ATT at pH 8.0

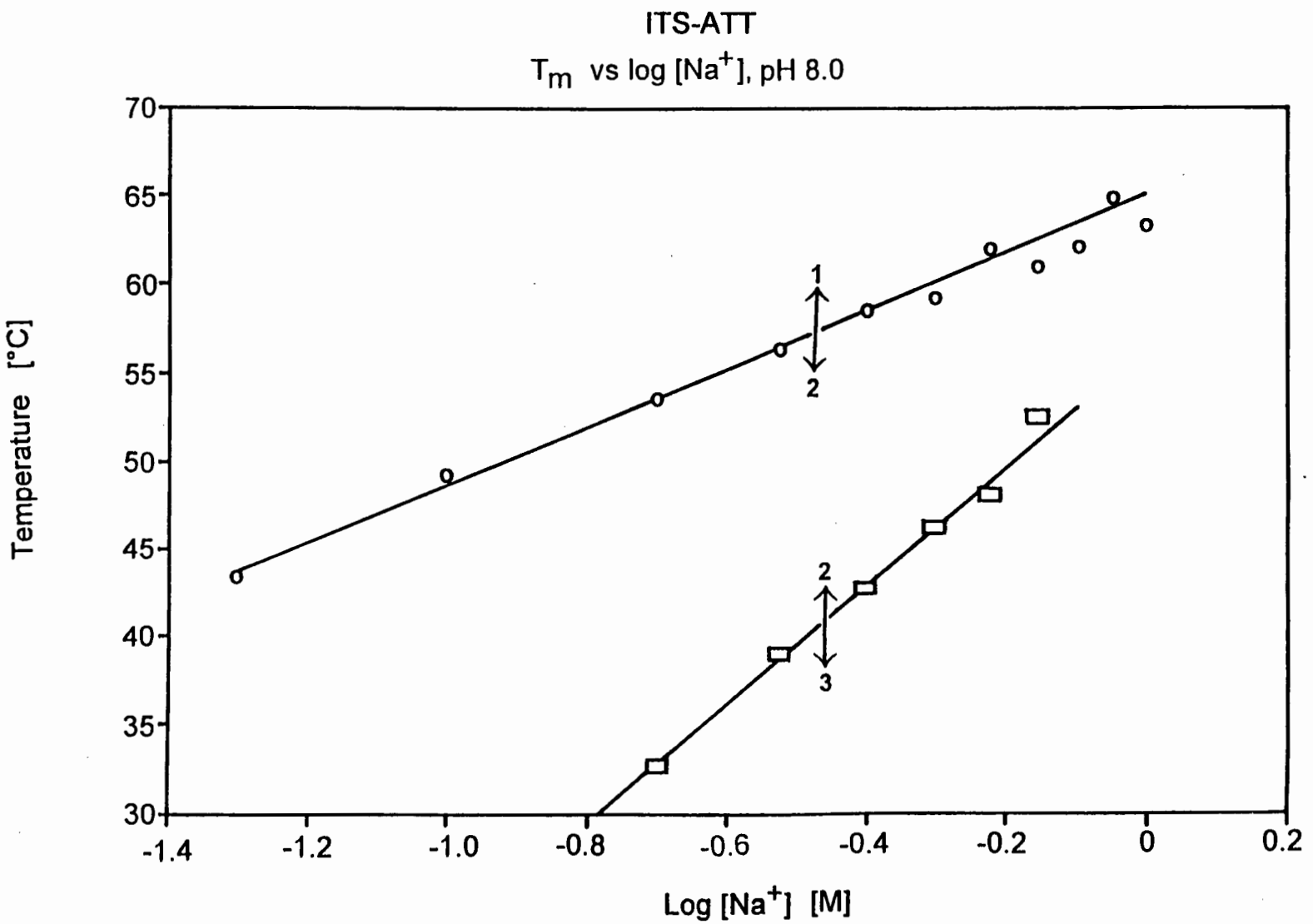
[2] Phasediagram: T_m vs log $[Na^+]$ for ITS-$2G_3$ at pH 8.0

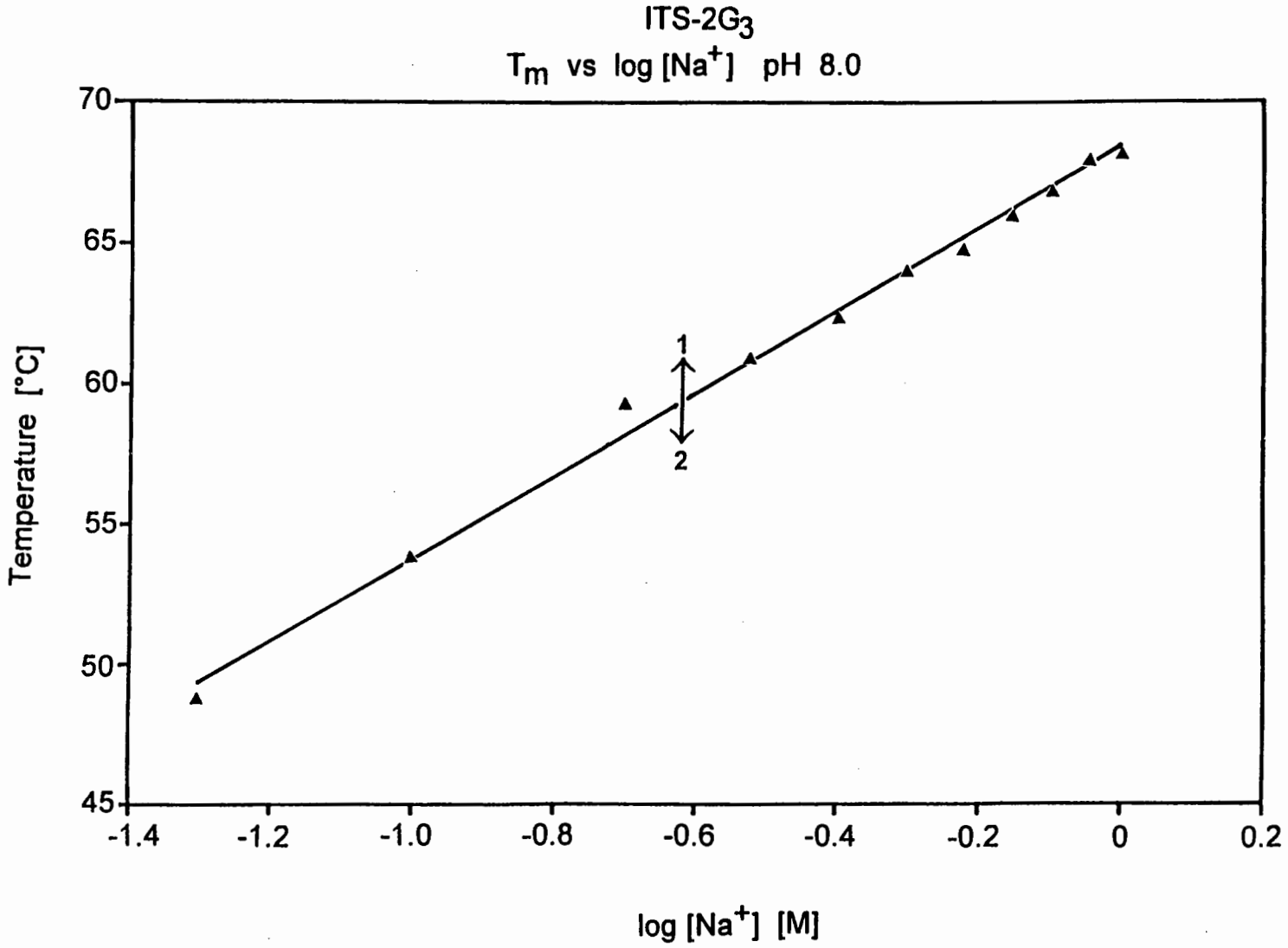
[3] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_0$ at pH 8.0

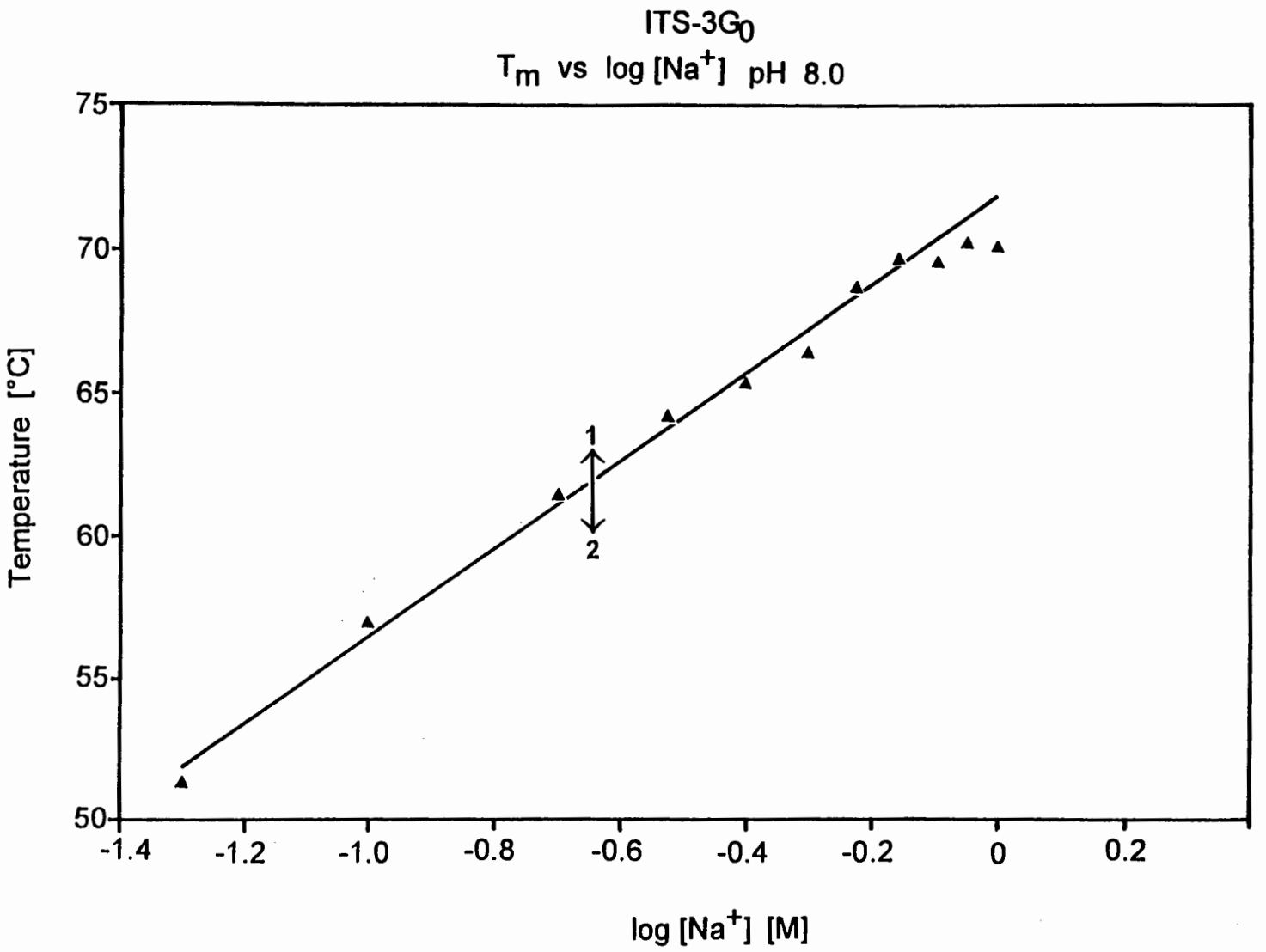
[4] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_1$ at pH 8.0

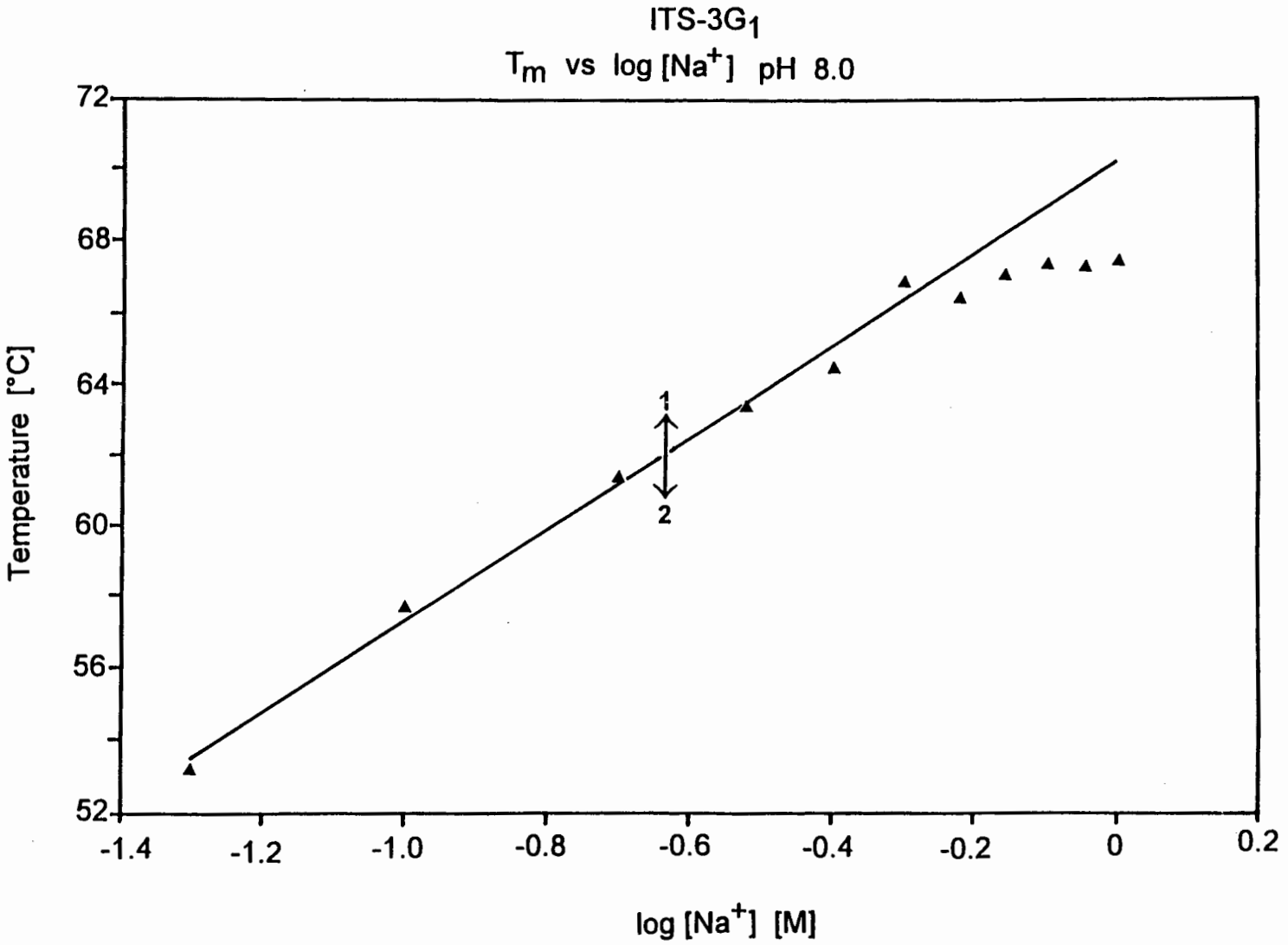
[5] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_2$ at pH 8.0

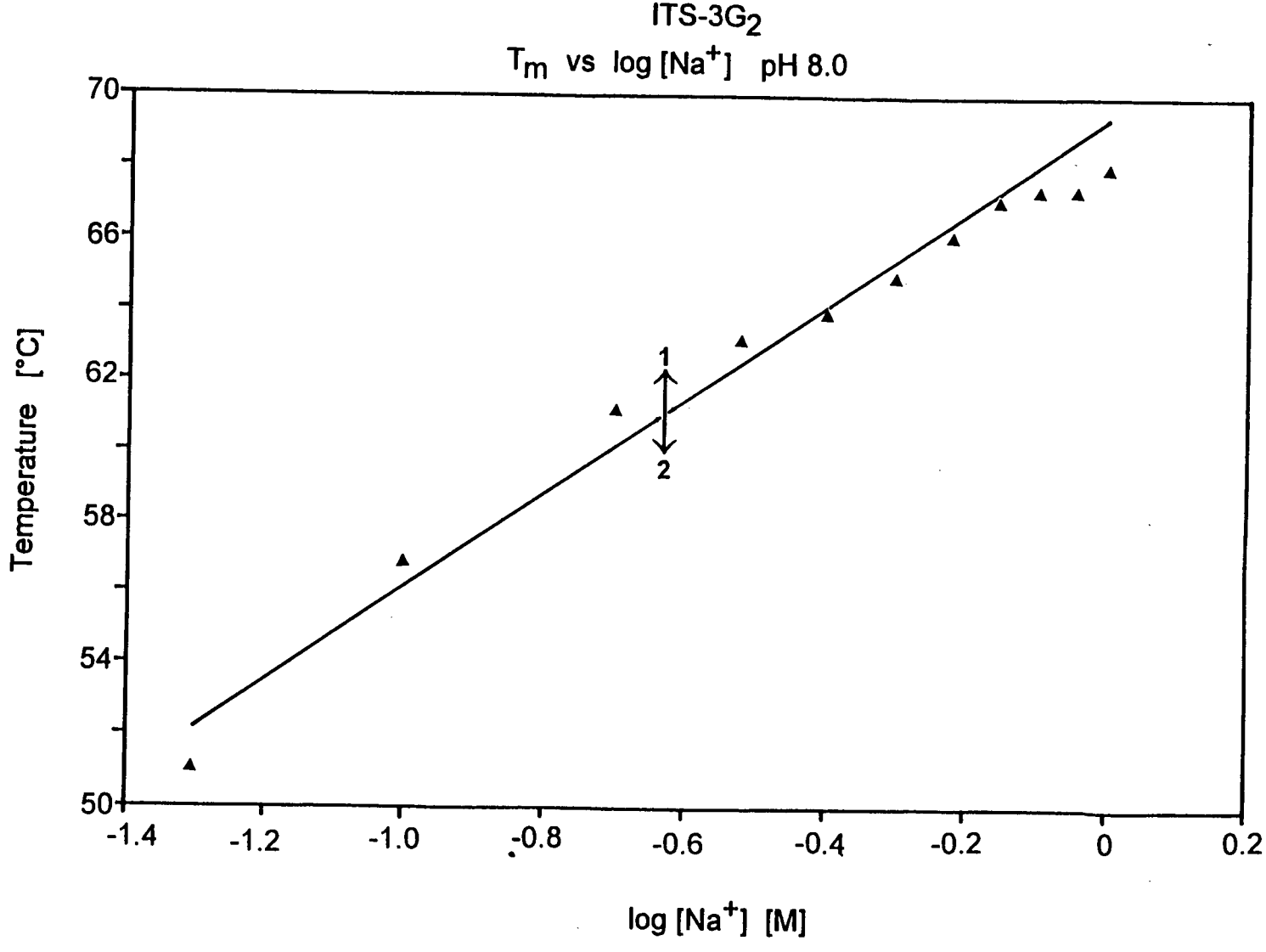
[6] Phasediagram: T_m vs log $[Na^+]$ for ITS-$4G_1$ at pH 8.0

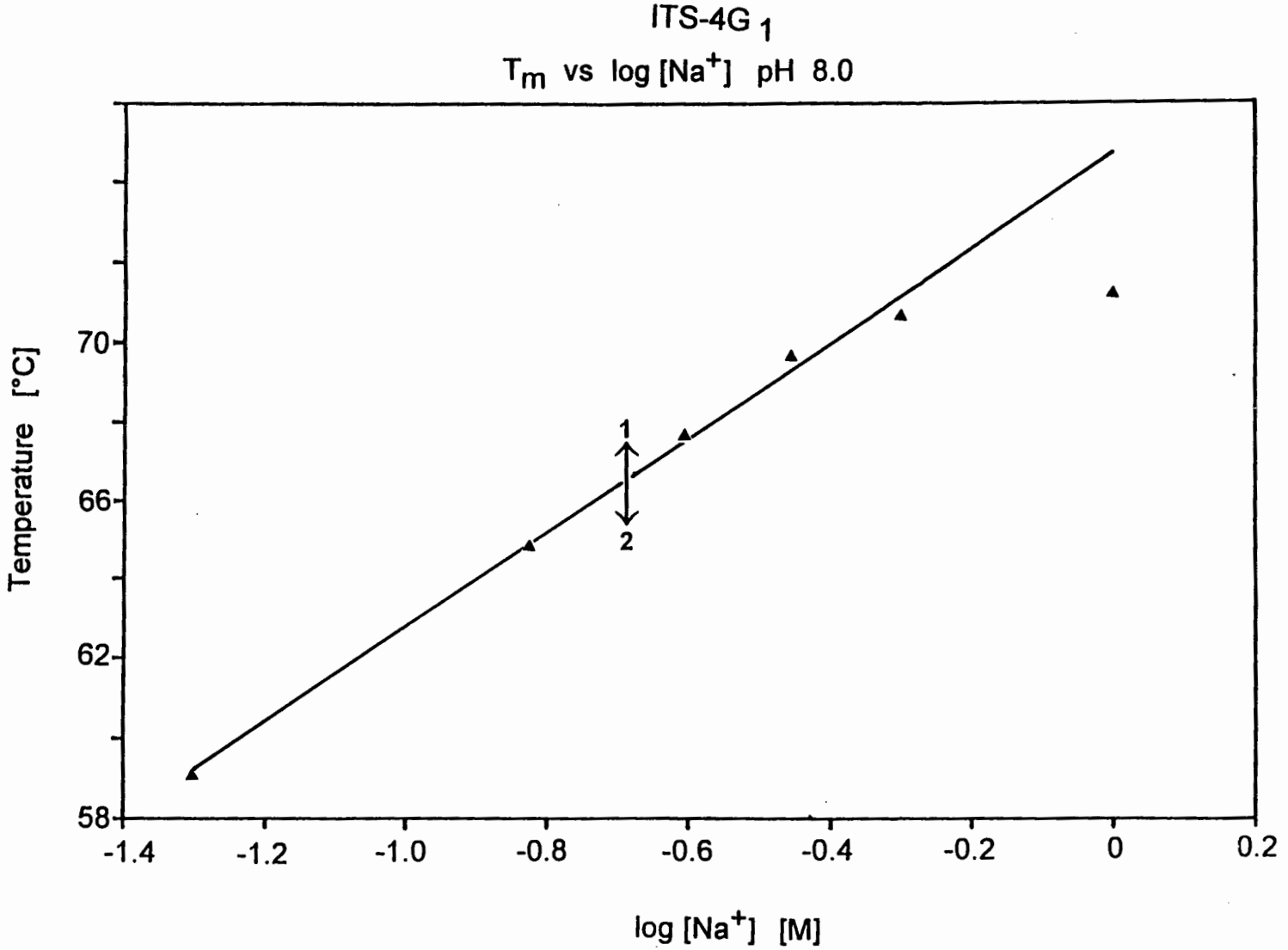

ITS-ATT
T_m vs log [Na^+], pH 8.0
Temperature [°C]
70
65
60
55
50
45
40
35
30
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
Log [Na^+] [M]
1
2
2
3


ITS-2G3
Tm vs log [Na+] pH 8.0
Temperature [°C]
70
65
60
55
50
45
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]
1
2


ITS-3G0
Tm vs log [Na+] pH 8.0
Temperature [°C]
75
70
65
60
55
50
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]
1
2


ITS-3G1
Tm vs log [Na+] pH 8.0
Temperature [°C]
72
68
64
60
56
52
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]
1
2


ITS-3G2
Tm vs log [Na+] pH 8.0
Temperature [°C]
70
66
62
58
54
50
1
2
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-4G 1
T_m vs log [Na^+] pH 8.0
Temperature [°C]
70
66
62
58
1
2
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na^+] [M]

APPENDIX B

T_m versus Ionic Strength at pH 6.75

[1] Phasediagram: T_m vs log $[Na^+]$ for ITS-ATT at pH 6.75

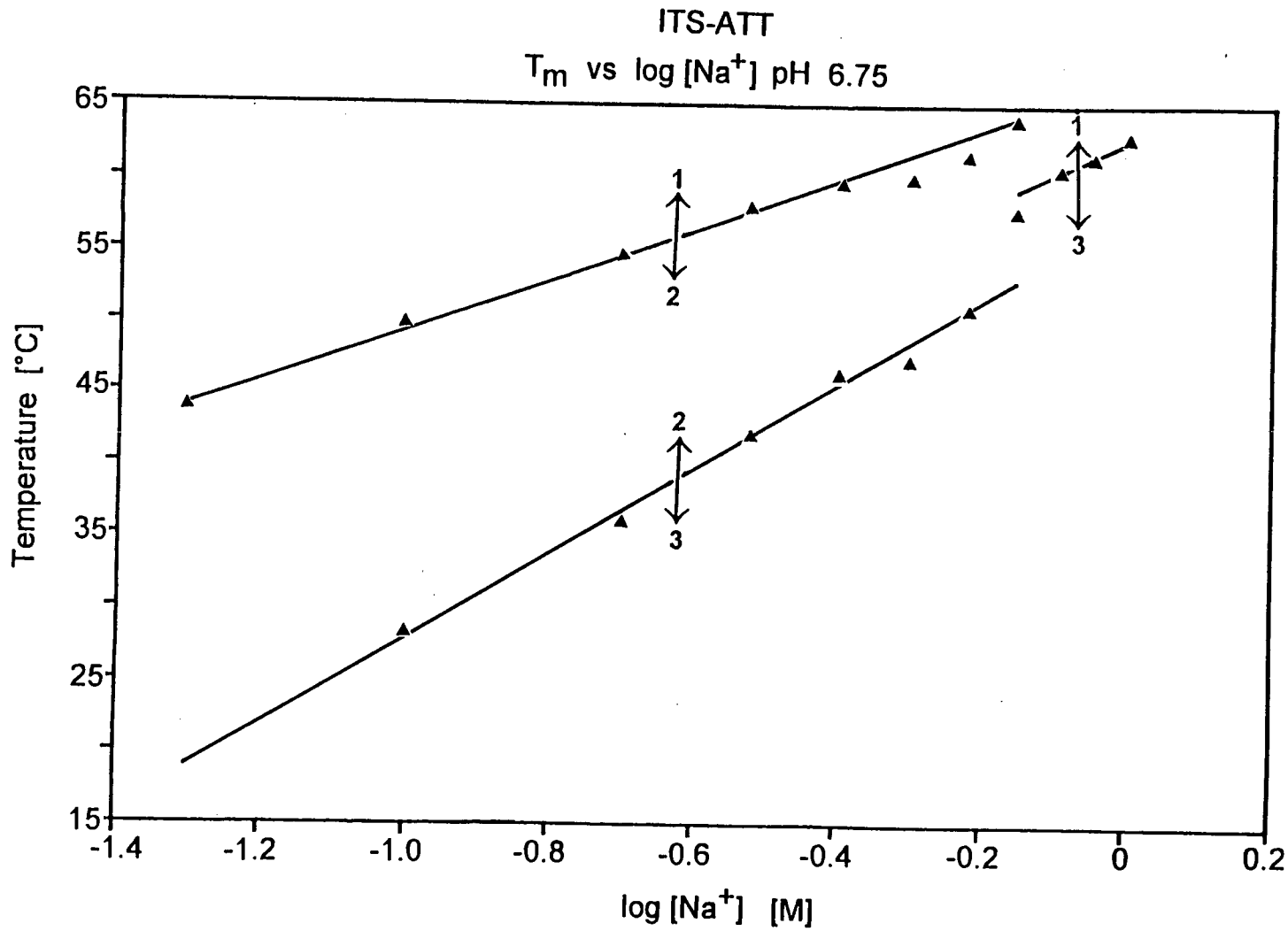
[2] Phasediagram: T_m vs log $[Na^+]$ for ITS-$2G_3$ at pH 6.75

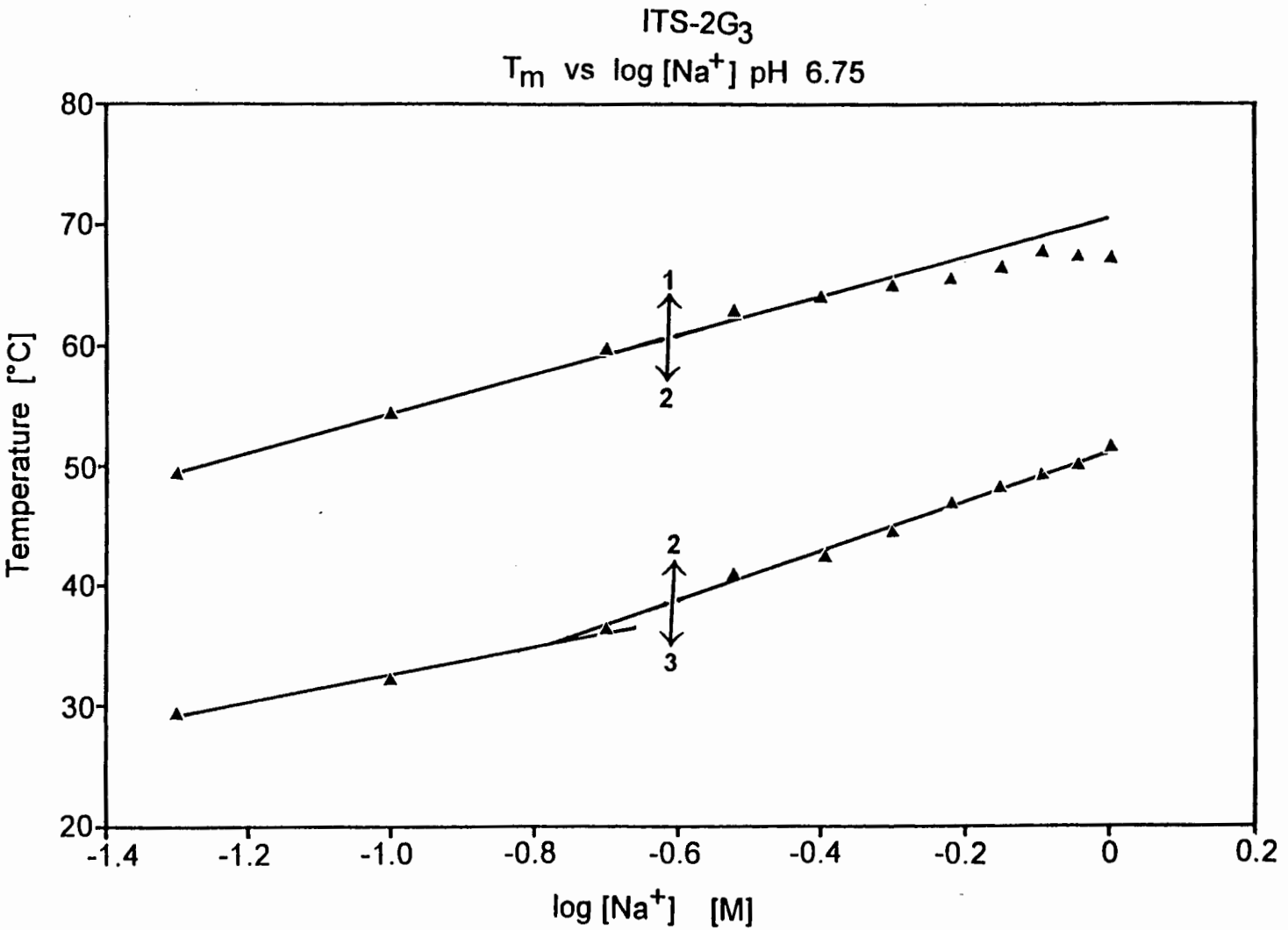
[3] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_0$ at pH 6.75

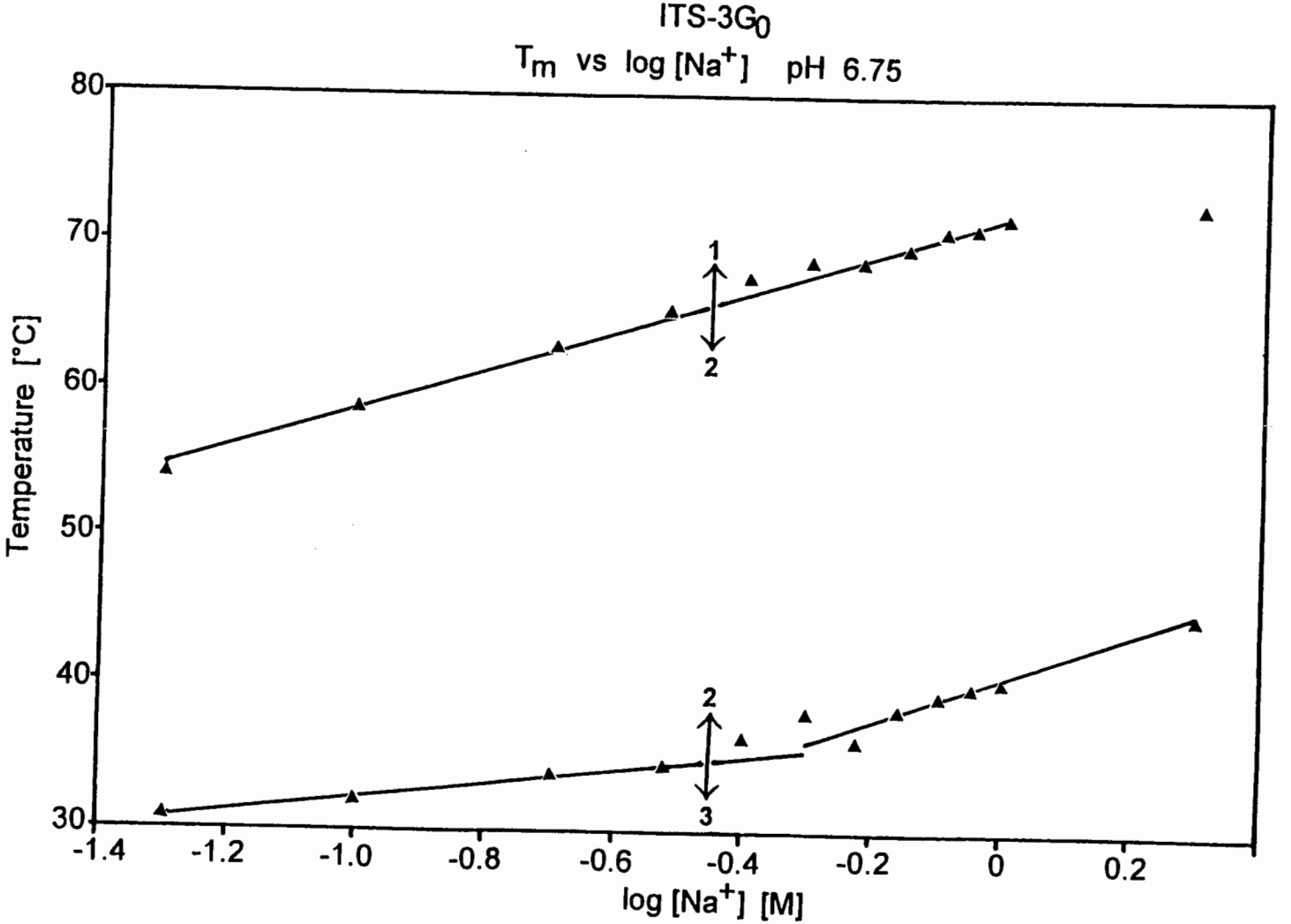
[4] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_1$ at pH 6.75

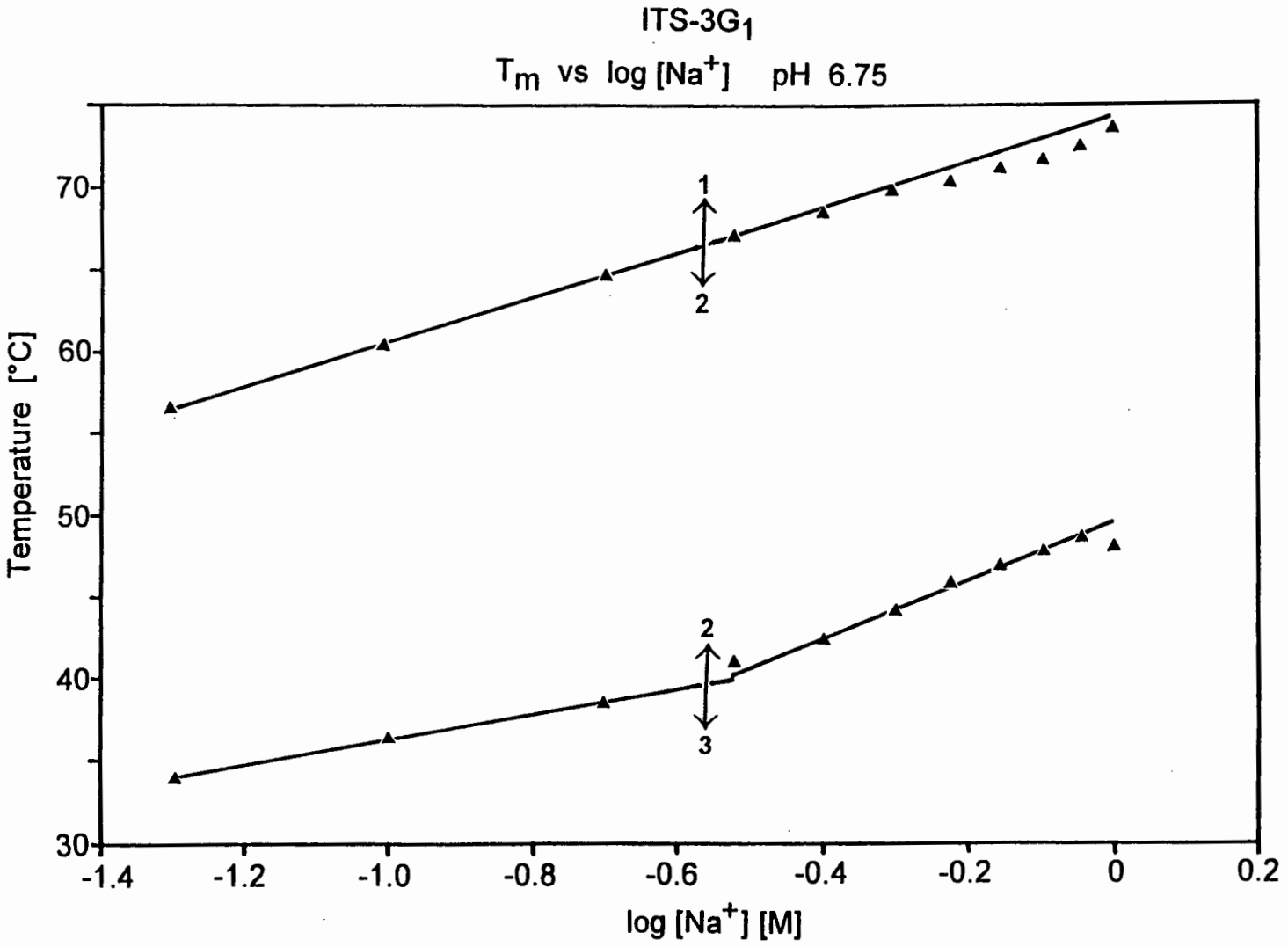
[5] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_2$ at pH 6.75

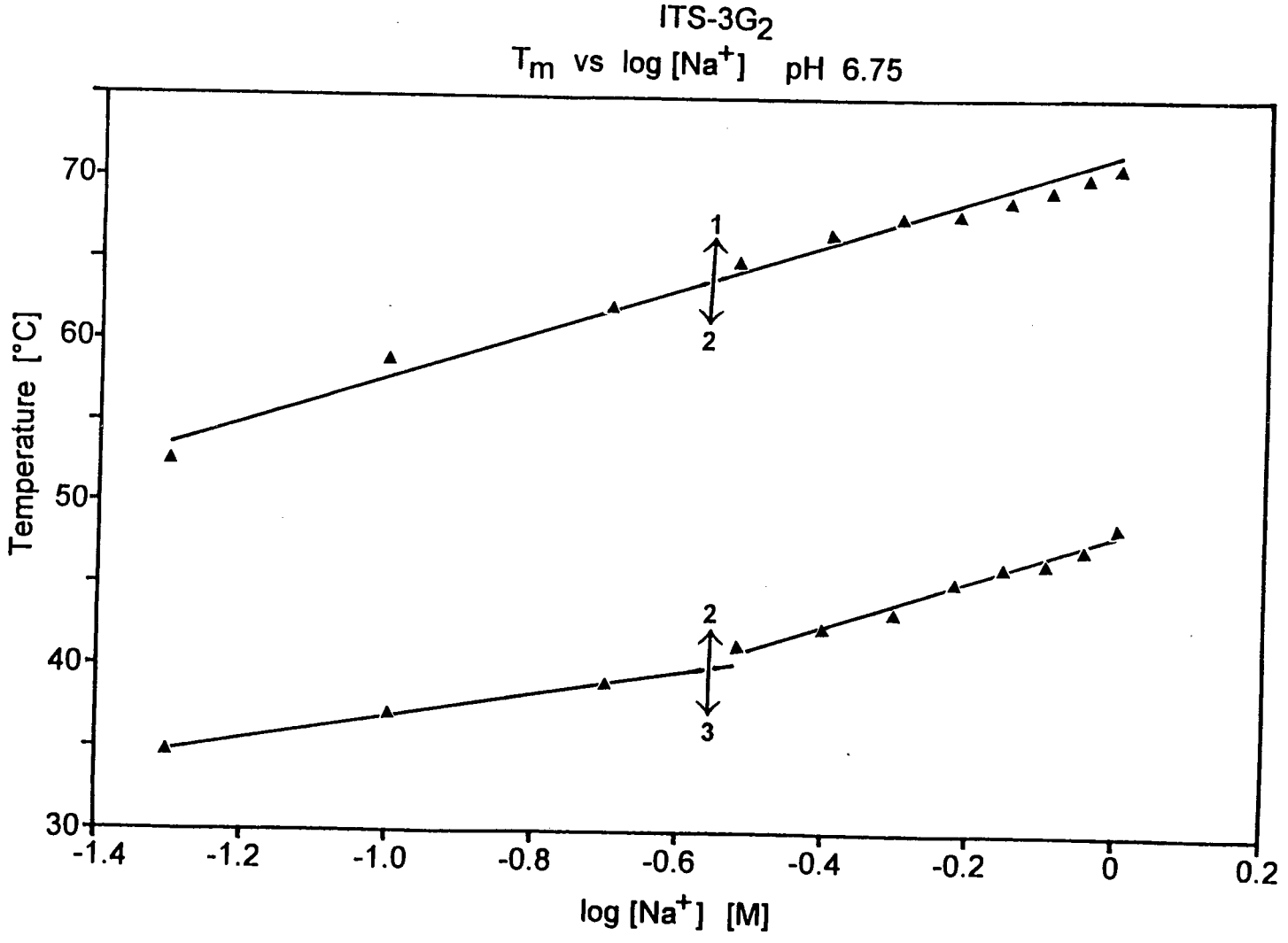
[6] Phasediagram: T_m vs log $[Na^+]$ for ITS-$4G_1$ at pH 6.9

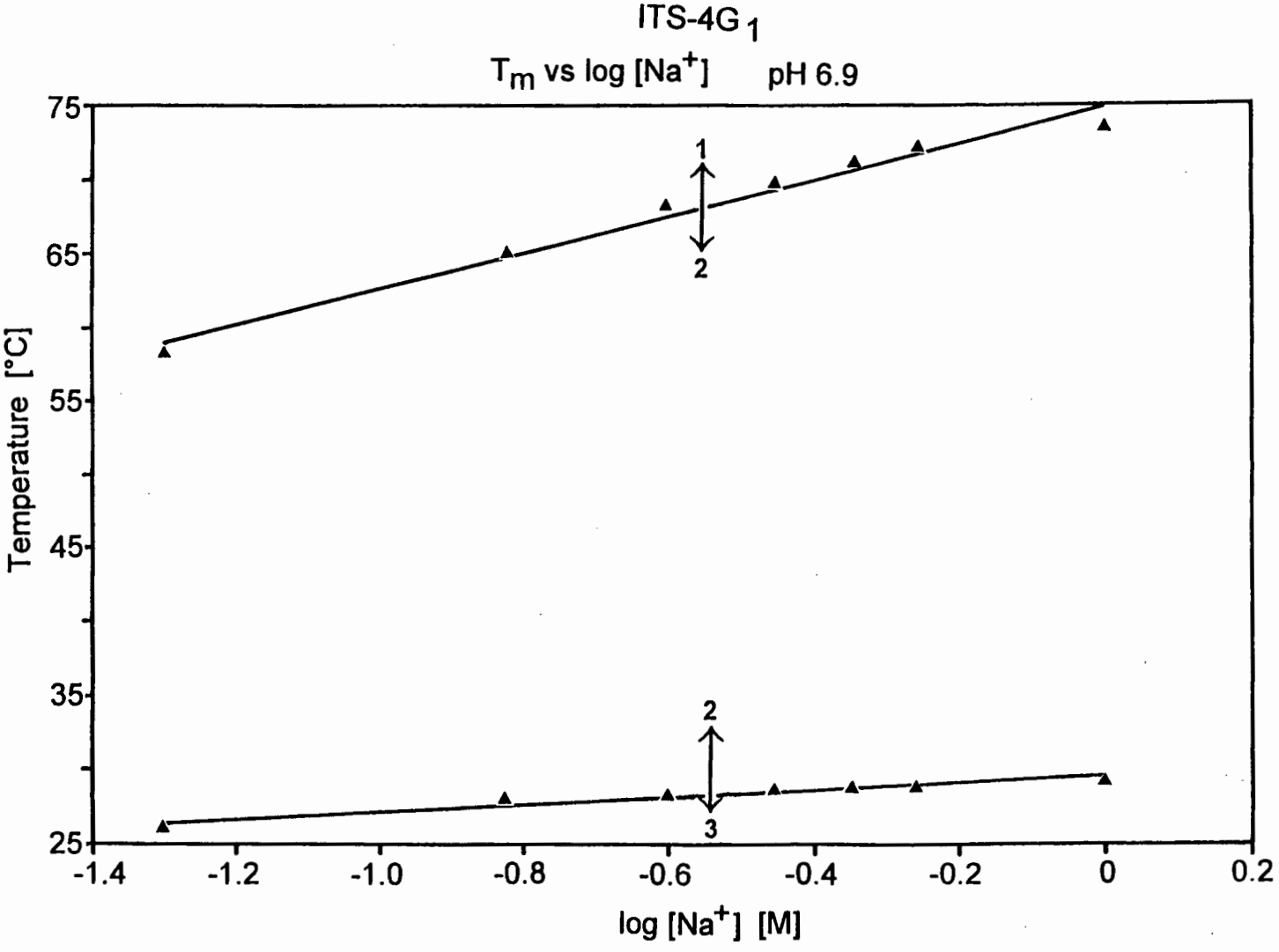

ITS-ATT
T_m vs log [Na^+] pH 6.75
Temperature [°C]
65
55
45
35
25
15
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na^+] [M]
1
2
3
1
2
3


ITS-2G3
Tm vs log [Na+] pH 6.75
Temperature [°C]
80
70
60
50
40
30
20
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]
1
2
2
3


ITS-3G0
Tm vs log [Na+] pH 6.75
Temperature [°C]
80
70
60
50
40
30
1
2
2
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-3G1
Tm vs log [Na+] pH 6.75
Temperature [°C]
70
60
50
40
30
1
2
2
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-3G2
Tm vs log [Na+] pH 6.75
Temperature [°C]
70
60
50
40
30
1
2
2
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-4G 1
T_m vs log [Na^+] pH 6.9
Temperature [°C]
75
65
55
45
35
25
1
2
2
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na^+] [M]

APPENDIX C

T_m versus Ionic Strength at pH 4.5

[1] Phasediagram: T_m vs log $[Na^+]$ for ITS-ATT at pH 4.5

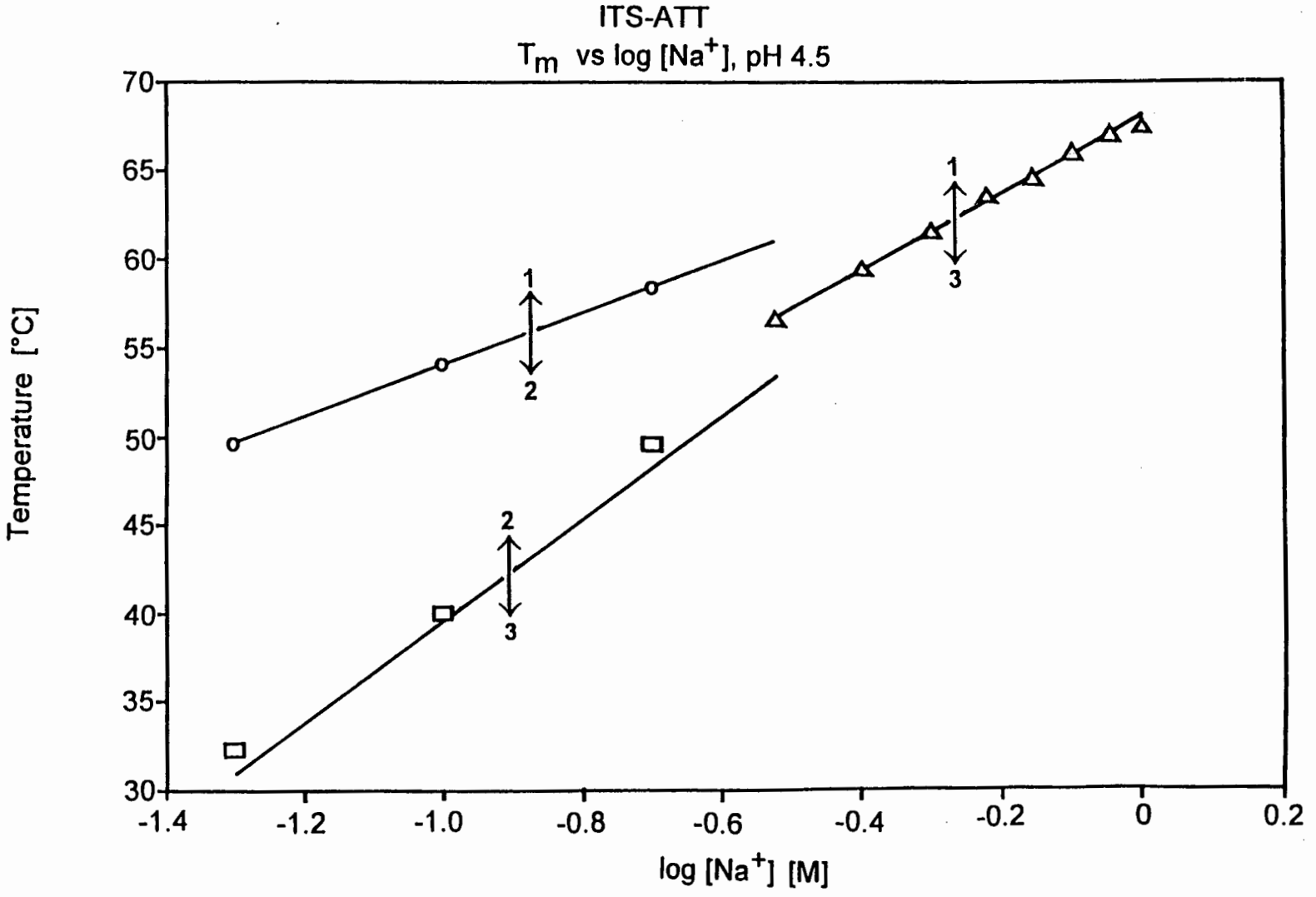
[2] Phasediagram: T_m vs log $[Na^+]$ for ITS-$2G_3$ at pH 4.5

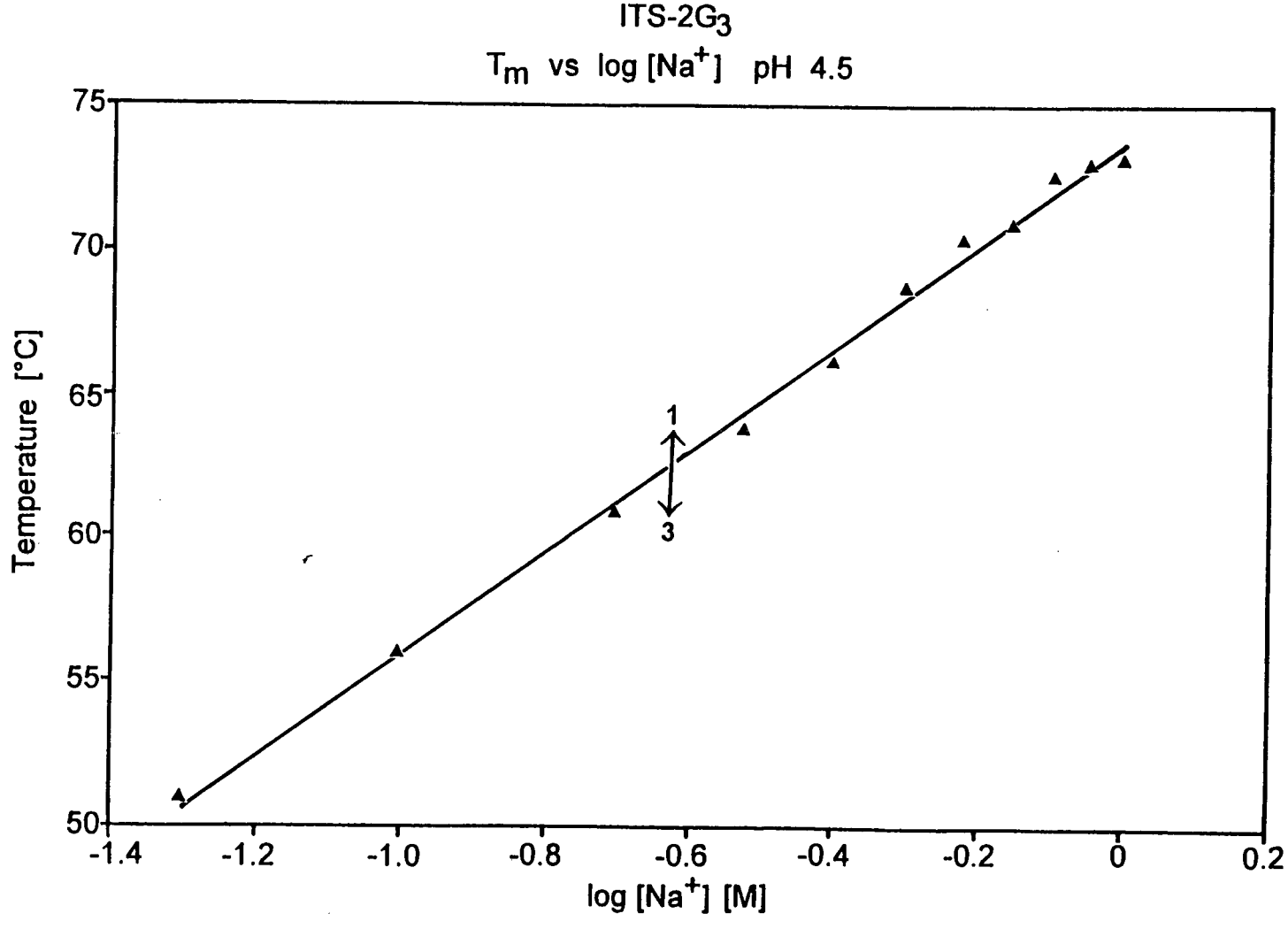
[3] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_0$ at pH 4.5

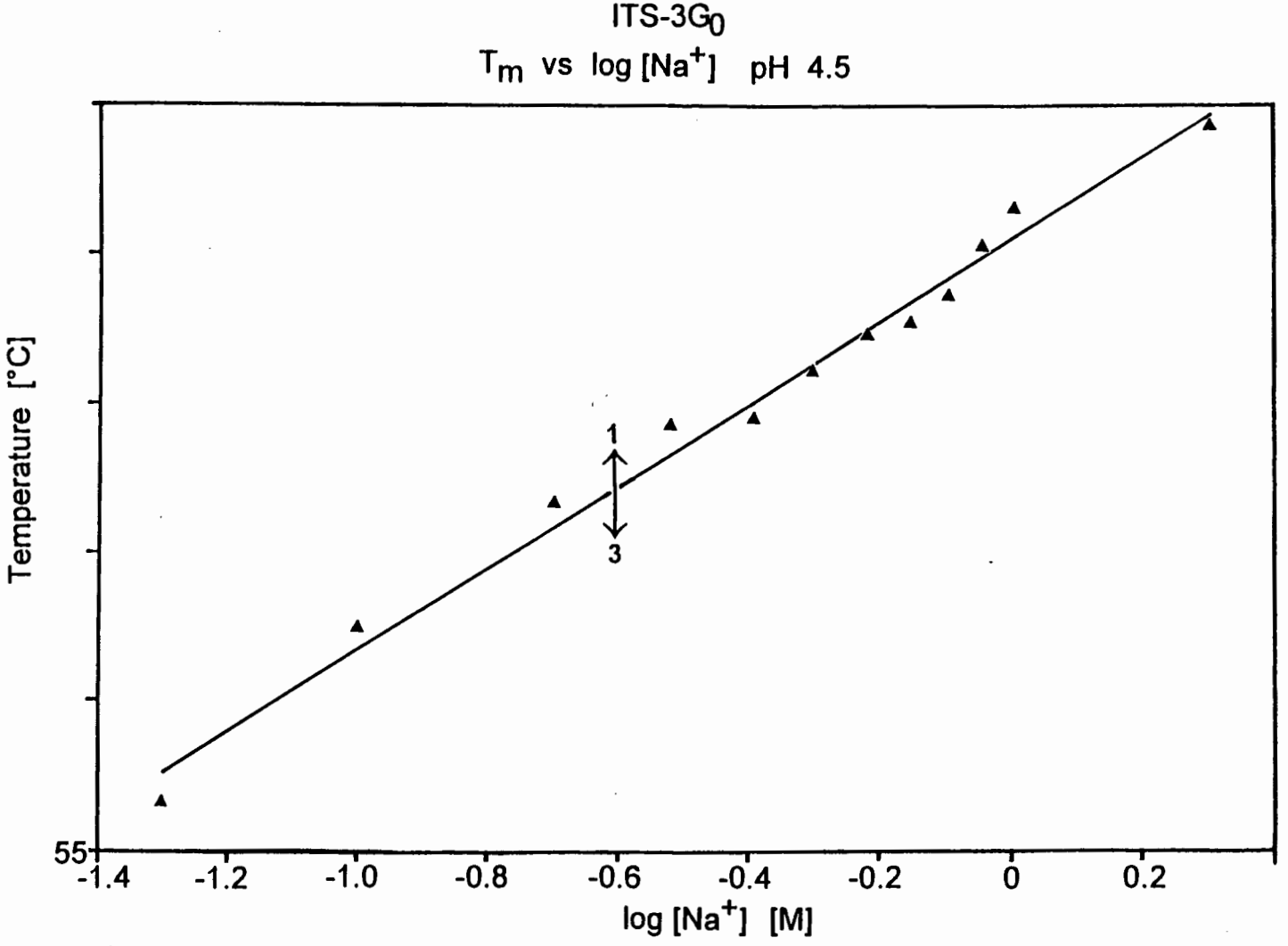
[4] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_1$ at pH 4.5

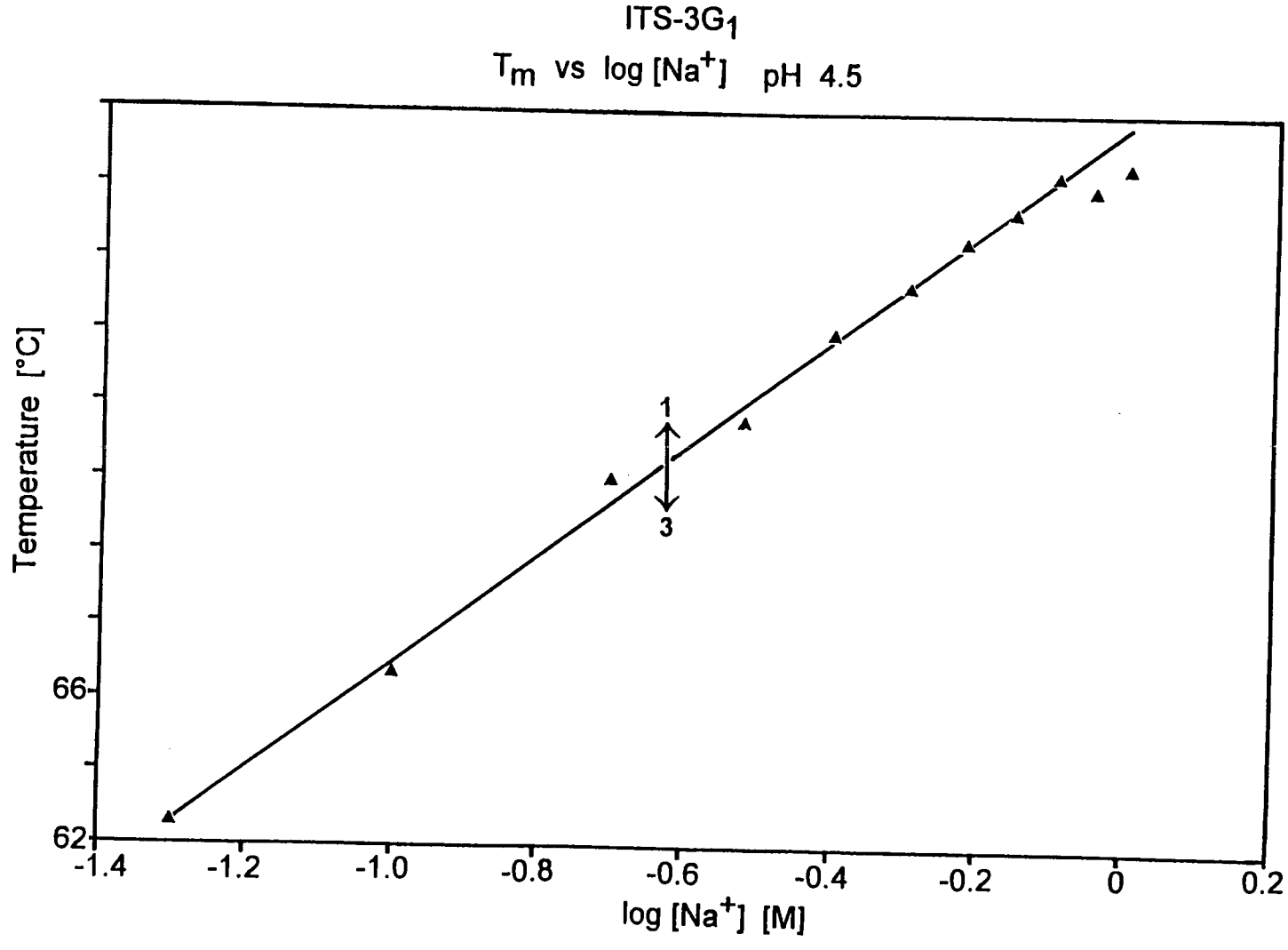
[5] Phasediagram: T_m vs log $[Na^+]$ for ITS-$3G_2$ at pH 4.5

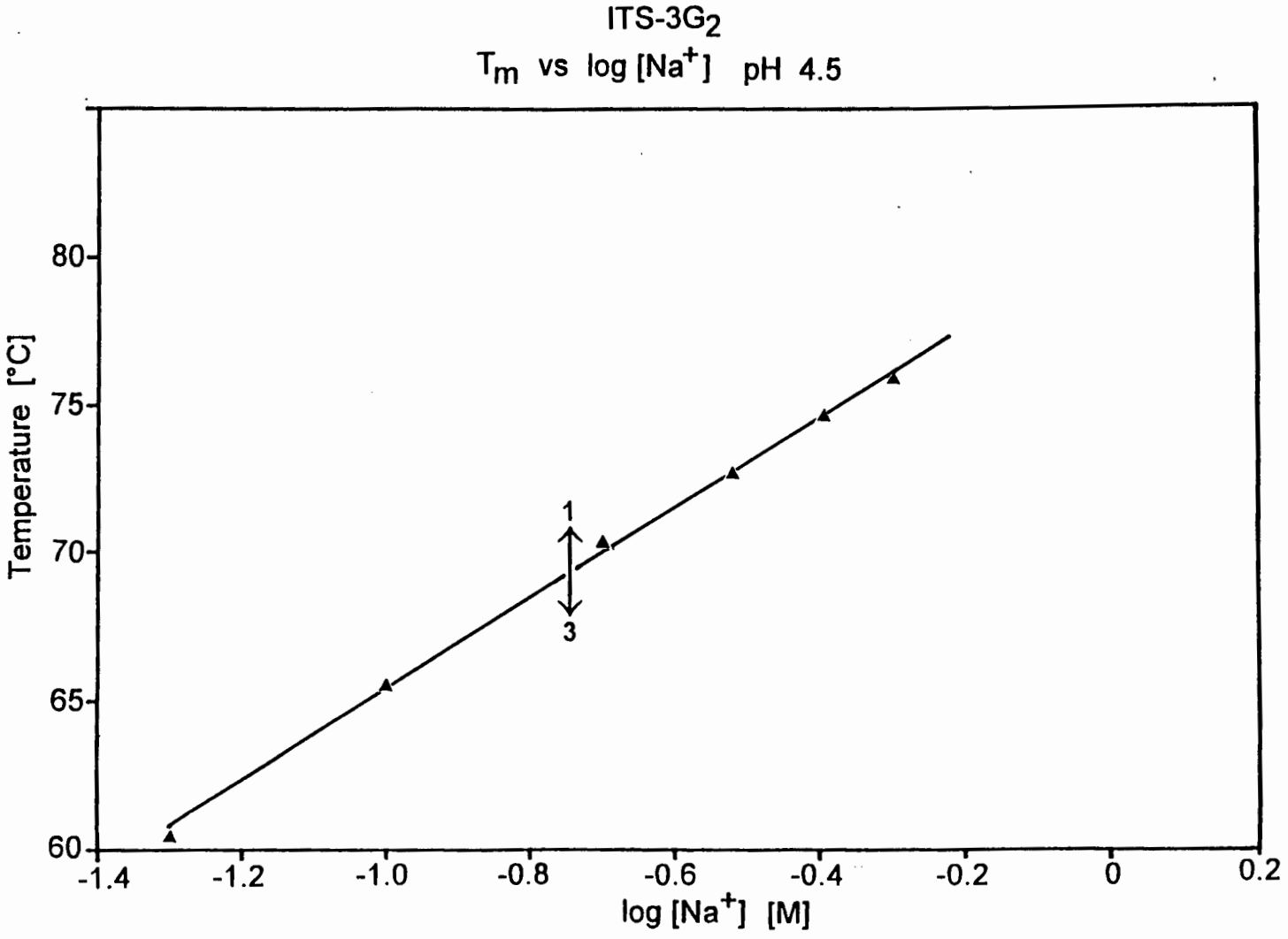
[6] Phasediagram: T_m vs log $[Na^+]$ for ITS-$4G_1$ at pH 4.5

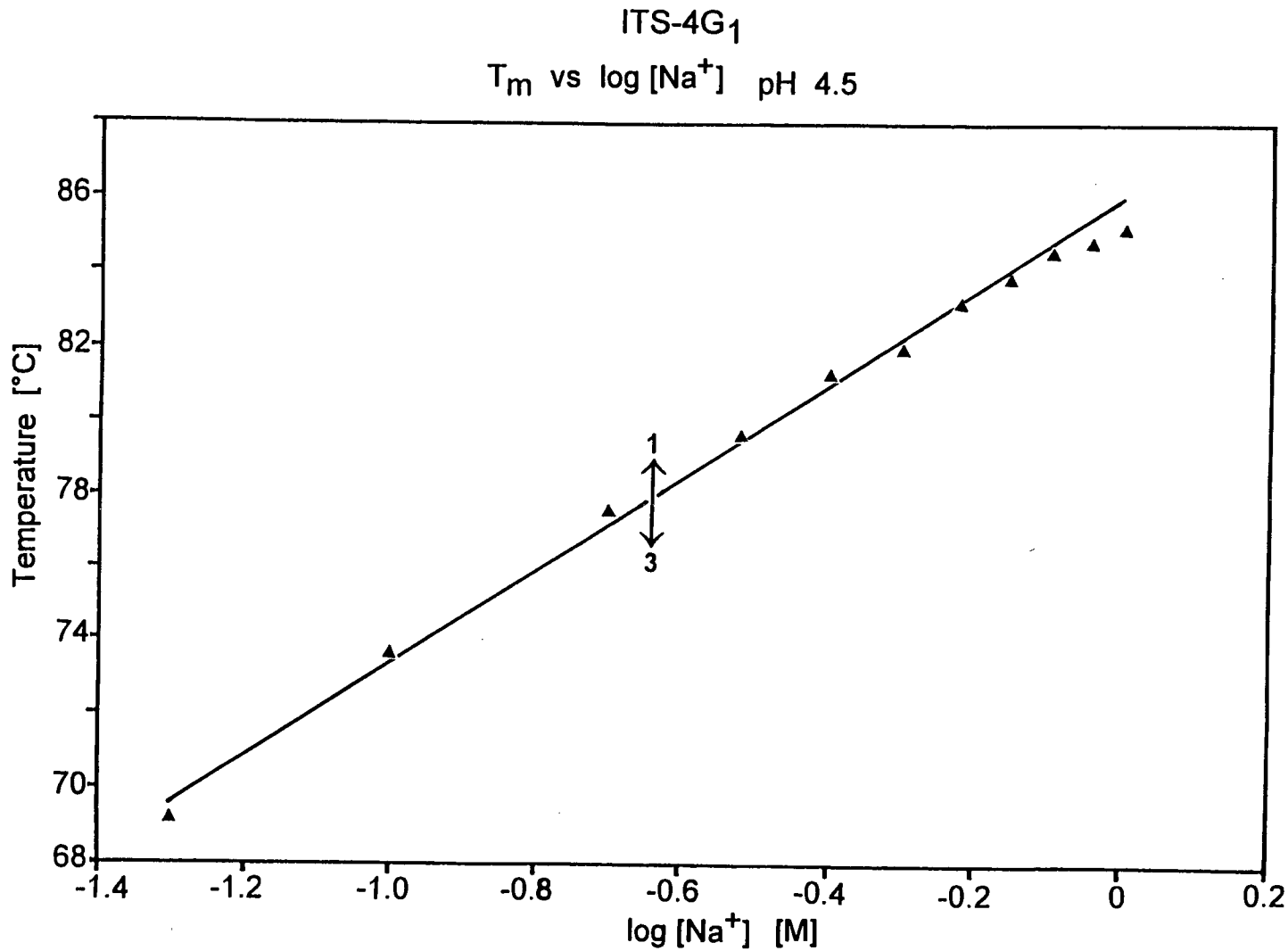

ITS-ATT
T_m vs log [Na^+], pH 4.5
Temperature [°C]
70
65
60
55
50
45
40
35
30
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na^+] [M]
1
2
1
3
2
3


ITS-2G3
Tm vs log [Na+] pH 4.5
Temperature [°C]
75
70
65
60
55
50
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]
1
3


ITS-3G0
Tm vs log [Na+] pH 4.5
Temperature [°C]
55
1
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-3G1
Tm vs log [Na+] pH 4.5
Temperature [°C]
66
62
1
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-3G2
Tm vs log [Na+] pH 4.5
Temperature [°C]
80
75
70
65
60
1
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]


ITS-4G1
Tm vs log [Na+] pH 4.5
Temperature [°C]
86
82
78
74
70
68
1
3
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
0.2
log [Na+] [M]

APPENDIX D

T_m versus pH in 150 mM Na^+

[1] Phasediagram: T_m vs pH for ITS-ATT in 150 mM Na^+

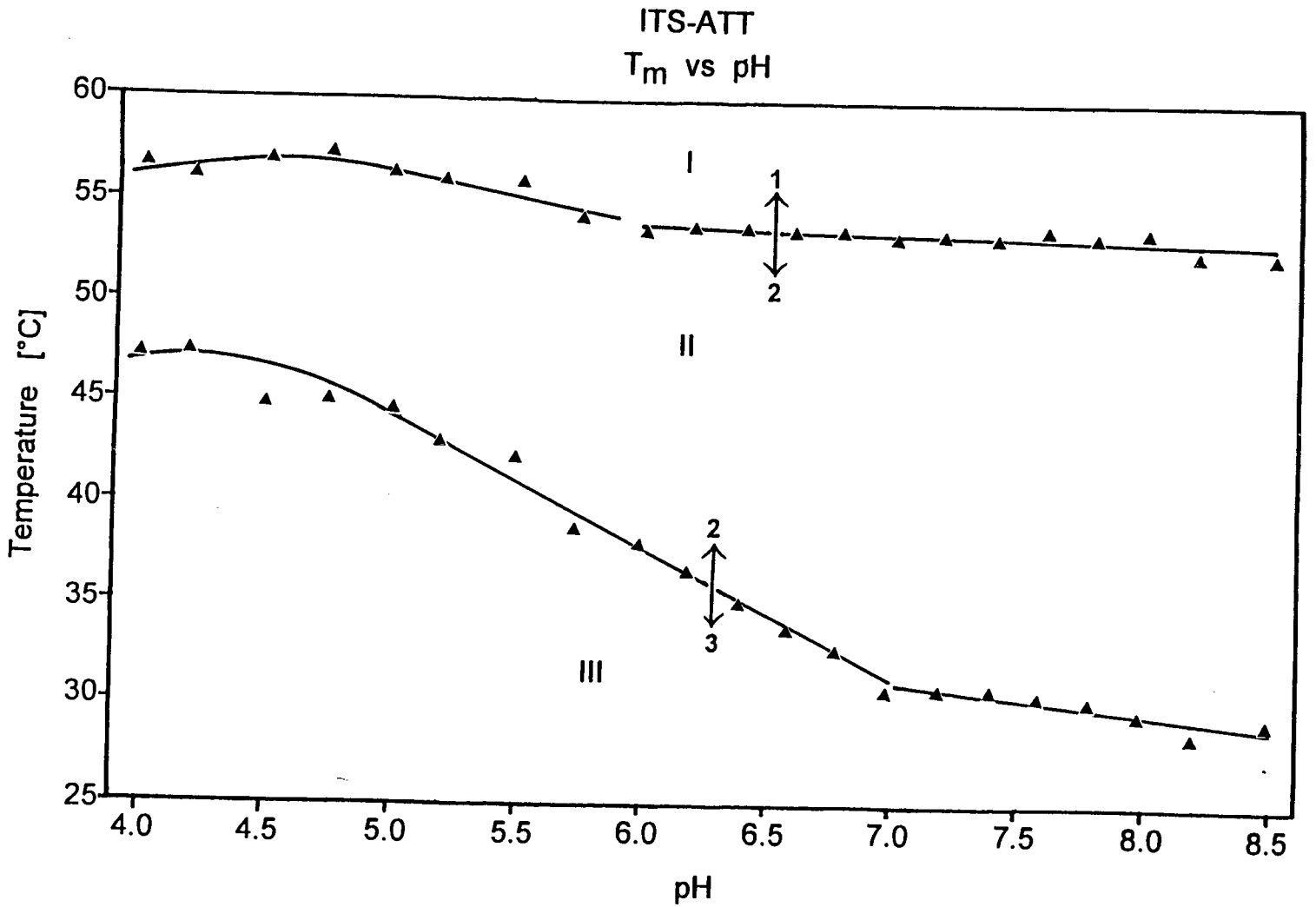
[2] Phasediagram: T_m vs pH for ITS-$2G_3$ in 150 mM Na^+

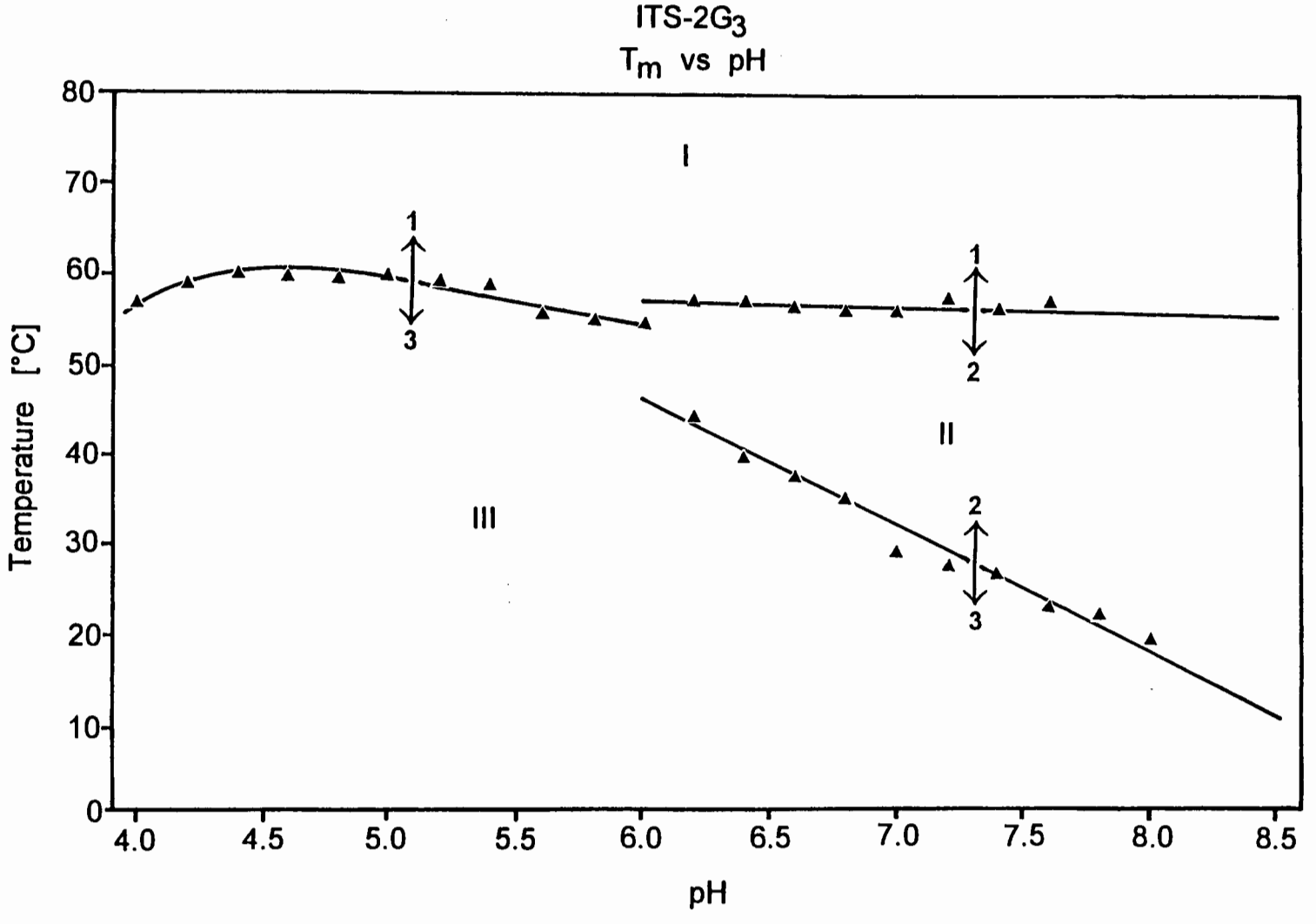
[3] Phasediagram: T_m vs pH for ITS-$3G_0$ in 150 mM Na^+

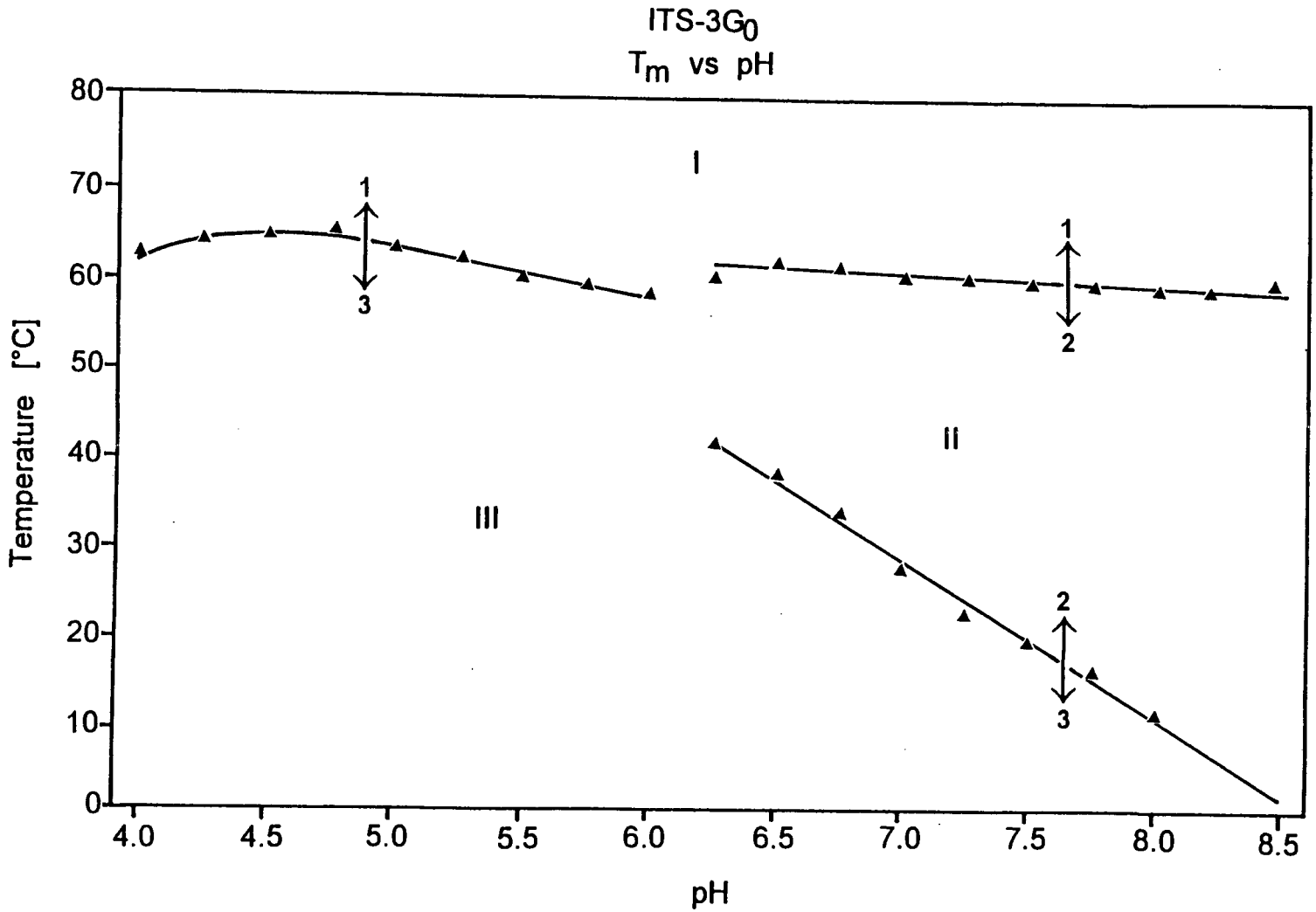
[4] Phasediagram: T_m vs pH for ITS-$3G_1$ in 150 mM Na^+

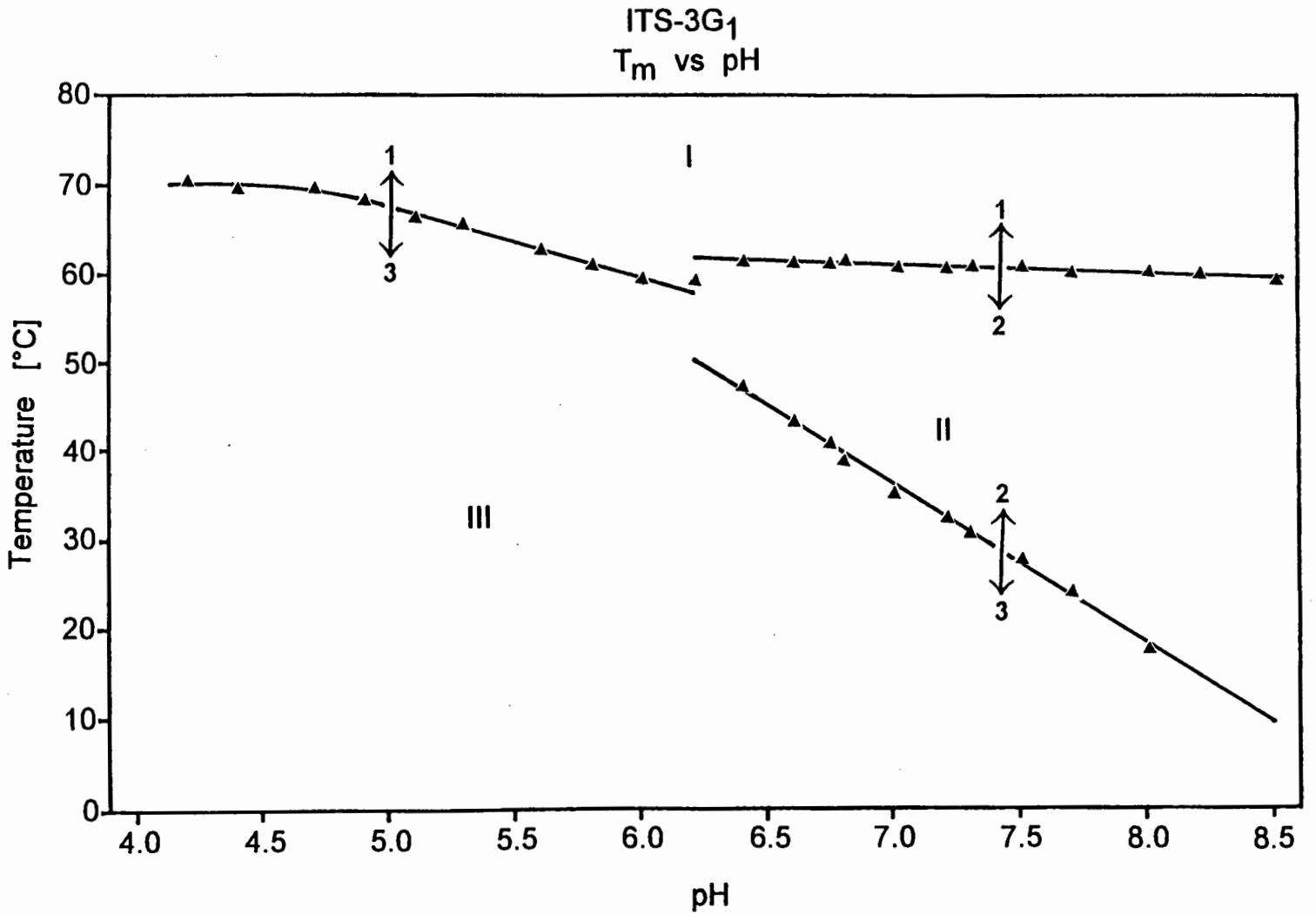
[5] Phasediagram: T_m vs pH for ITS-$3G_2$ in 150 mM Na^+

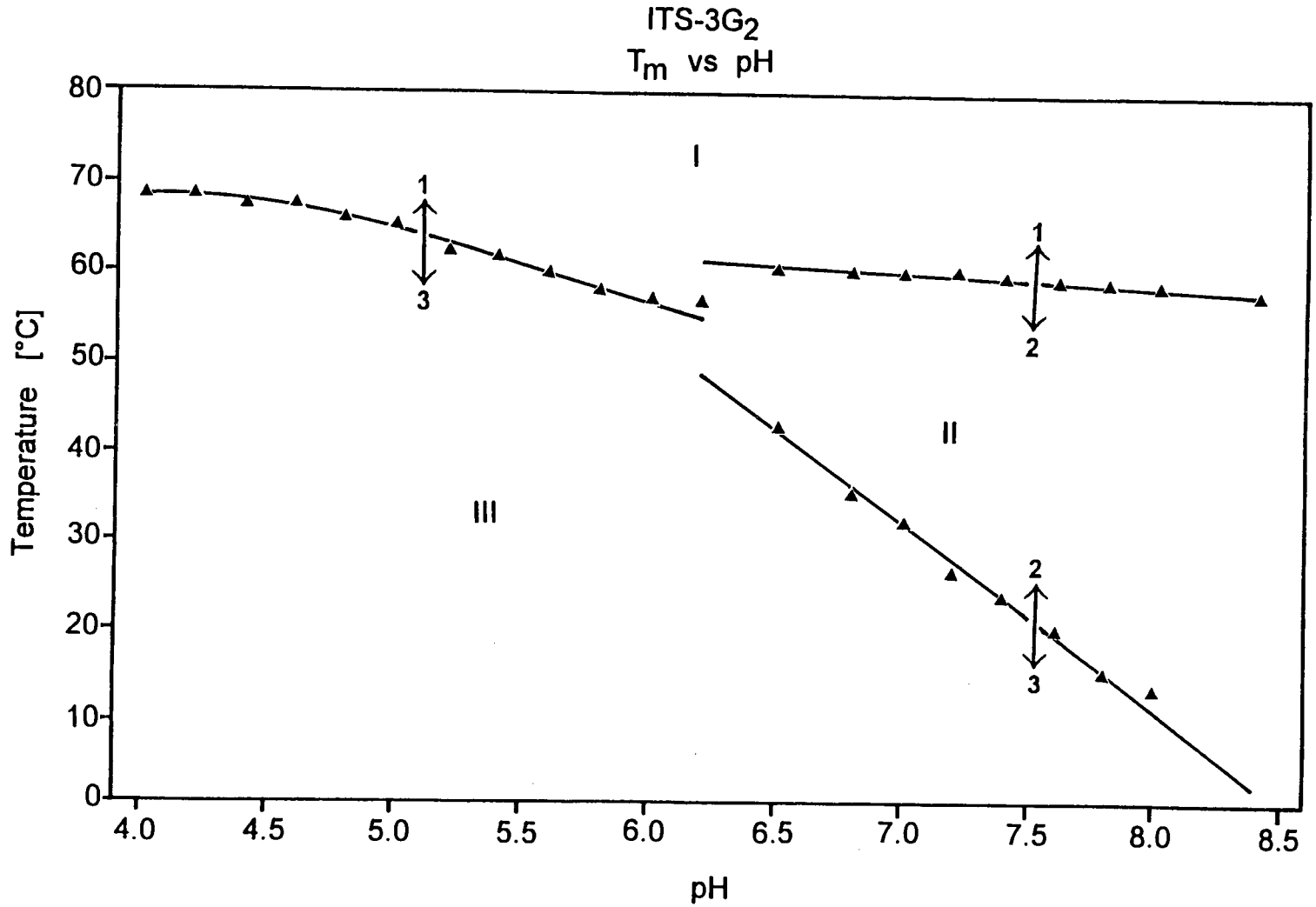
[6] Phasediagram: T_m vs pH for ITS-$4G_1$ in 150 mM Na^+

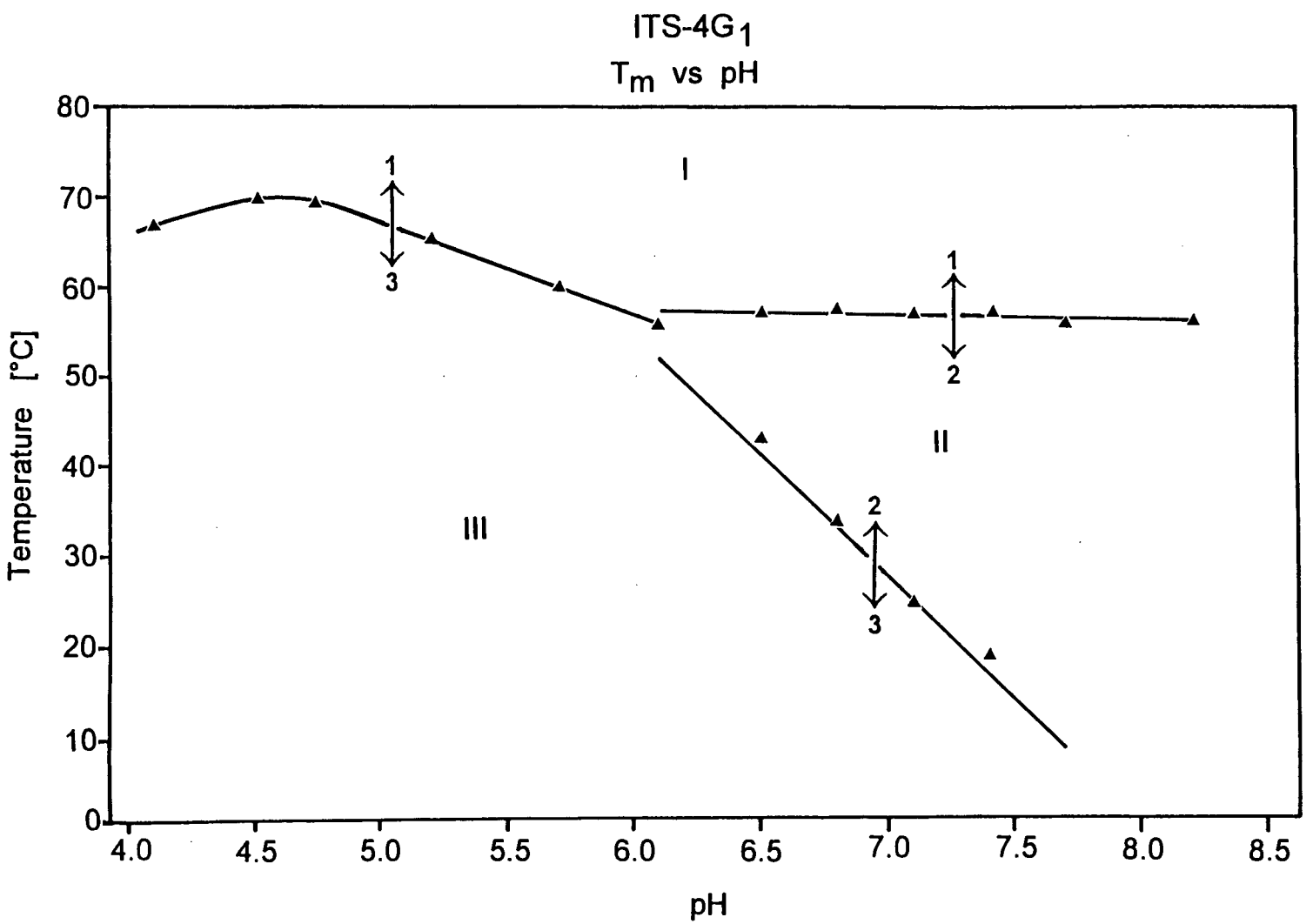

ITS-ATT
T_m vs pH
Temperature [°C]
60
55
50
45
40
35
30
25
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
pH
I
II
III
1
2
2
3


ITS-2G3
Tm vs pH
Temperature [°C]
pH
I
II
III
1
2
3
1
2
3
80
70
60
50
40
30
20
10
0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5


ITS-3G0
Tm vs pH
Temperature [°C]
80
70
60
50
40
30
20
10
0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
pH
I
II
III
1
2
3
1
3
2


ITS-3G1
Tm vs pH
Temperature [°C]
pH
I
II
III
1
2
3
1
2
3
80
70
60
50
40
30
20
10
0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5


ITS-3G2
Tm vs pH
Temperature [°C]
pH
80
70
60
50
40
30
20
10
0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
I
II
III
1
2
3
1
3
2


ITS-4G1
Tm vs pH
Temperature [°C]
80
70
60
50
40
30
20
10
0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
pH
I
II
III
1
3
1
2
2
3